Regional Information Report No. 5J21-02

Genetic Population Structure of Chinook Salmon from the Middle and Upper Susitna River: A Report to Alaska Energy Authority, Susitna-Watana Hydroelectric Project (submitted July 25, 2017)

by

Andrew W. Barclay

Heather A. Hoyt

and

Christopher Habicht

March 2021

Alaska Department of Fish and Game **Division Commercial Fisheries**

Symbols and Abbreviations

The following symbols and abbreviations, and others approved for the Système International d'Unités (SI), are used without definition in the following reports by the Divisions of Sport Fish and of Commercial Fisheries: Fishery Manuscripts, Fishery Data Series Reports, Fishery Management Reports, and Special Publications. All others, including deviations from definitions listed below, are noted in the text at first mention, as well as in the titles or footnotes of tables, and in figure or figure captions.

Weights and measures (metric)

centimeter	cm
deciliter	dL
gram	g
hectare	ha
kilogram	kg
kilometer	km
liter	L
meter	m
milliliter	mL
millimeter	mm

Weights and measures (English)

cubic feet per second	ft^3/s
foot	ft
gallon	gal
inch	in
mile	mi
nautical mile	nmi
ounce	oz
pound	lb
quart	qt
yard	yd

Time and temperature

day	d
degrees Celsius	°C
degrees Fahrenheit	°F
degrees kelvin	K
hour	h
minute	min
second	s

Physics and chemistry

all atomic symbols	
alternating current	AC
ampere	A
calorie	cal
direct current	DC
hertz	Hz
horsepower	hp
hydrogen ion activity (negative log of)	pH
parts per million	ppm
parts per thousand	ppt, ‰
volts	V
watts	W

General

Alaska Administrative Code	AAC
all commonly accepted abbreviations	e.g., Mr., Mrs., AM, PM, etc.
all commonly accepted professional titles	e.g., Dr., Ph.D., R.N., etc.
at	@
compass directions:	
east	E
north	N
south	S
west	W
copyright	©
corporate suffixes:	
Company	Co.
Corporation	Corp.
Incorporated	Inc.
Limited	Ltd.
District of Columbia	D.C.
et alii (and others)	et al.
et cetera (and so forth)	etc.
exempli gratia (for example)	e.g.
Federal Information Code	FIC
id est (that is)	i.e.
latitude or longitude	lat or long
monetary symbols (U.S.)	$, ¢
months (tables and figures): first three letters	Jan,...,Dec
registered trademark	®
trademark	™
United States (adjective)	U.S.
United States of America (noun)	USA
U.S.C.	United States Code
U.S. state	use two-letter abbreviations (e.g., AK, WA)

Mathematics, statistics

all standard mathematical signs, symbols and abbreviations	
alternate hypothesis	H_A
base of natural logarithm	*e*
catch per unit effort	CPUE
coefficient of variation	CV
common test statistics	(F, t, χ^2, etc.)
confidence interval	CI
correlation coefficient (multiple)	R
correlation coefficient (simple)	r
covariance	cov
degree (angular)	°
degrees of freedom	df
expected value	*E*
greater than	>
greater than or equal to	≥
harvest per unit effort	HPUE
less than	<
less than or equal to	≤
logarithm (natural)	ln
logarithm (base 10)	log
logarithm (specify base)	$\log_2$, etc.
minute (angular)	'
not significant	NS
null hypothesis	H_O
percent	%
probability	P
probability of a type I error (rejection of the null hypothesis when true)	α
probability of a type II error (acceptance of the null hypothesis when false)	β
second (angular)	"
standard deviation	SD
standard error	SE
variance	
population	Var
sample	var

REGIONAL INFORMATION REPORT NO. 5J21-02

GENETIC POPULATION STRUCTURE OF CHINOOK SALMON FROM THE MIDDLE AND UPPER SUSITNA RIVER: A REPORT TO ALASKA ENERGY AUTHORITY, SUSITNA-WATANA HYDROELECTRIC PROJECT (SUBMITTED JULY 25, 2017)

by

Andrew W. Barclay, Heather A. Hoyt, and Christopher Habicht
Alaska Department of Fish and Game, Division of Commercial Fisheries, Gene Conservation Laboratory, Anchorage

Alaska Department of Fish and Game
Division of Commercial Fisheries
333 Raspberry Road, Anchorage, Alaska, 99518-1565

March 2021

The Regional Information Report Series was established in 1987 and was redefined in 2007 to meet the Division of Commercial Fisheries regional need for publishing and archiving information such as area management plans, budgetary information, staff comments and opinions to Alaska Board of Fisheries proposals, interim or preliminary data and grant agency reports, special meeting or minor workshop results and other regional information not generally reported elsewhere. Reports in this series may contain raw data and preliminary results. Reports in this series receive varying degrees of regional, biometric and editorial review; information in this series may be subsequently finalized and published in a different department reporting series or in the formal literature. Please contact the author or the Division of Commercial Fisheries if in doubt of the level of review or preliminary nature of the data reported. Regional Information Reports are available through the Alaska State Library and on the Internet at: http://www.adfg.alaska.gov/sf/publications/.

Product names used in this publication are included for completeness and do not constitute product endorsement. The Alaska Department of Fish and Game does not endorse or recommend any specific company or their products.

Andrew W. Barclay, Heather A. Hoyt, and Christopher Habicht
Alaska Department of Fish and Game, Division of Commercial Fisheries, Gene Conservation Laboratory
333 Raspberry Road, Anchorage, AK 99518, USA

This document should be cited as follows:
Barclay, A. W., H. A. Hoyt, and C. Habicht. 2021. Genetic population structure of Chinook salmon from Middle and Upper Susitna River: A report to Alaska Energy Authority, Susitna-Watana hydroelectric project (submitted July 25, 2017). Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report No. 5J21-02, Anchorage.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

ABSTRACT

In 2012, the Alaska Energy Authority proposed a hydroelectric dam on the Susitna River upstream of Devils Canyon. Chinook salmon are the only anadromous species known to spawn within and above Devils Canyon. Policymakers need to know if Chinook salmon spawning above the canyon constitute separate self-sustaining population(s) or are a collection of strays from other populations. Here we analyzed genotypes for 12 microsatellite loci from 322 spawning adults and 408 juvenile Chinook salmon collected in the Middle and Upper Susitna Rivers. We determined that the juvenile collections were not appropriate for representing populations within and above Devils Canyon because they were highly related (resulting in upwardly biased genetic distances) and they did not conform to Hardy-Weinberg expectations. Annual estimates of effective population size for within and above Devils Canyon based on juvenile relatedness, sonar, and mark–recapture data were on the order of 10 to 100 fish for each section after accounting for biases and uncertainties in effective population size estimates. Tests for homogeneity of allele frequencies were significantly different between adults captured above and within Devils Canyon and between these fish and populations from below Devils Canyon (Indian River and Portage Creek). Allelic richness and private allele richness for adults captured within and above Devils Canyon was not significantly different from populations below Devils Canyon. Based on these lines of evidence, we conclude that spawning aggregates within and above Devils Canyon are neither a collection of strays from Indian River and Portage Creek nor self-sustaining populations, but likely a combination of both.

Key words: Chinook salmon, Northern Cook Inlet, Susitna River, *Oncorhynchus tshawytscha*, microsatellite, population structure, migration, drift, Upper Susitna River, Middle Susitna River, Su Hydro, Alaska Energy Authority.

INTRODUCTION

In 2012, the Alaska Energy Authority (AEA) initiated the process to license a hydroelectric project on the Susitna River (Su Hydro), which would involve construction of a dam and reservoir approximately 55 kilometers upstream of Devils Canyon (Figure 1). Among prelicensing studies were 13 studies on fish and aquatic resources. One of these studies was Study 9.14 *Genetic Baseline Study for Selected Fish Species*. One objective of this project was to characterize the genetic population structure of Chinook salmon (*Oncorhynchus tshawytscha*) from Upper Cook Inlet, with emphasis on spawning aggregates in the Middle and Upper Susitna River.

The impact of Su Hydro on Chinook salmon was of particular interest to the project because they are the only anadromous species known to pass the Devils Canyon impediments in the Middle River and spawn in areas both below and above (Upper River) the proposed dam site. Estimated numbers of Chinook salmon that ascended into and above Devils Canyon during the study years was low. Mark–recapture estimates were 15 fish (2013) to 111 fish (2014); sonar estimated 24 fish passing the dam site in 2014, and aerial surveys estimate 63 fish above Devils Canyon in 2013 (Table 1; AEA 2015). Understanding the population structure of Chinook salmon collected above and below Devils Canyon would therefore inform policymakers on whether Chinook salmon spawning above the canyon constitute a separate population.

Barclay and Habicht (*In prep.*) analyzed the population structure of Chinook salmon from the Susitna River in the context of populations from Upper Cook Inlet using single nucleotide polymorphism (SNP) markers. In their analysis, 3 collections from the Middle River were included (Indian River, Portage Creek, and Cheechako Creek). Portage Creek and Indian River were genetically similar to Lower Susitna River populations. Cheechako Creek appeared more similar to lower Chulitna River populations. However, the genetic distances between Cheechako Creek (F_{ST} = 0.006) and Portage and Indian Creeks (F_{ST} = 0.009) was small.

Analyses testing hypotheses of gene flow and estimating effective population sizes for Chinook salmon captured within and above Devils Canyon have not been completed. The original project design called for collecting samples from juvenile and adult Chinook salmon over multiple years (minimum of 3 to 4) from multiple locations both within and above Devils Canyon (AEA 2012). This distribution and number of samples was anticipated to allow for the estimation of effective population size using multiple methods, including temporal methods. The original sampling design was also set up to enable tests of 3 primary hypotheses to explain population structure of Chinook Salmon above and within Devils Canyon: 1) they represent self-sustaining population(s) (Hypothesis 1a; Figure 2); 2) they are individuals originating from other geographic spawning aggregates (e.g., Portage Creek; Hypothesis 1b); or 3) they are some combination of a local population and a nearby stock(s) (Hypothesis 2). To distinguish between hypotheses 1a and b, the original experimental design called for measuring interannual stability in allele frequencies within locations both within and above Devils Canyon. However, this project was funded to collect samples for 2 years, and in the first year, access to sampling within and above Devils Canyon below the proposed dam site was restricted. As a result, different analyses had to be performed using the available samples suitable for genetic analysis from both juvenile and adult Chinook salmon collected under the Su Hydro project. These samples provide insight into the forces of genetic migration (straying) and genetic drift on Chinook salmon in this portion of the drainage but are not sufficient to fully test hypotheses of gene flow between populations below, within and above Devils Canyon. Here we present an analysis for Middle and Upper Susitna River Chinook salmon populations that demonstrates gene flow among spawning aggregates, and low levels of genetic divergence among these spawning aggregates and populations below Devils Canyon.

DEFINITIONS

Definitions of commonly used genetic terms are provided here to better understand the methods, results, and interpretation of this study.

Allele. Alternative form of a given gene or DNA sequence.

Bottleneck. A sharp reduction in effective population size reducing the genetic variation within a population.

F_{IS}. The inbreeding coefficient of an individual with respect to the local subpopulation.

F_{ST}. Fixation index is an estimate of the proportion of the variation at a locus attributable to divergence among populations.

Linkage disequilibrium. A state that exists in a population when alleles at different loci are not distributed independently in the population's gamete pool, sometimes because the loci are physically linked.

Gene flow. The introduction of genes to a population, through genetic migration and mating from another population of the same species, thereby altering the allele frequencies of the population.

Genetic drift. The change in allele frequencies in a population through time due to random sampling at each generation. The effect of genetic drift increases with smaller population size and shorter number of generations.

Genetic marker. A known DNA sequence that can be identified by a simple assay.

Genotype. The set of alleles for 1 or more loci for a fish.

Hardy-Weinberg expectations (*HWE*). Genotype frequencies expected from a given set of allele frequencies for a locus. Fit to HWE genotypic proportions assumes random mating, no mutation (the alleles remain unchanged), no genetic migration or emigration (no exchange of alleles between populations), infinitely large population size, and no selective pressure for or against the alleles.

Heterozygosity. The proportion of individuals in a population that have 2 different allele forms (are heterozygous) at a particular marker. Average heterozygosity can be used as measure of variability in a sample.

Locus (plural, loci). A fixed position or region on a chromosome.

Microsatellite. A locus made up of short repeated sequences of DNA. The number of repeats determines the allele size.

Genetic migration. The movement of genes from 1 population into another accomplished by fish that stray and spawn in non-natal spawning habitat.

Linked markers. Genetic markers showing linkage disequilibrium, or physical linkage on a chromosome.

Population. A locally interbreeding group of spawning individuals that do not interbreed with individuals in other spawning aggregations, and that may be uniquely adapted to a particular spawning habitat. This produces isolation among populations and may lead to the appearance of unique attributes (Ricker 1958) that result in different productivity rates (Pearcy 1992; NRC 1996). This population definition is analogous to *spawning aggregations* described by Baker et al. (1996) and *demes* described by the NRC (1996).

METHODS

TISSUE SAMPLING

Tissue samples suitable for genetic analyses (genetic samples) were collected from spawning aggregates of Chinook salmon by ADF&G using hook-and-line gear in the tributaries in the Middle and Upper Susitna River below, within, and above Devils Canyon (Figure 3). Genetic samples were also collected from juvenile Chinook salmon from tributaries and the mainstem within and above Devils Canyon using minnow traps, electrofishers, rotary screw traps, and fyke nets. Additionally, some genetic samples were collected from radiotagged Chinook salmon at fish wheels operated at Curry and the lower Susitna River and then tracked to locations in the Middle and Upper Susitna River. We considered the final locations for Chinook salmon radio-tag samples as their collection location in subsequent tables and figures. Target sample size for each spawning location was 95 individuals across all years to achieve acceptable precision to estimate allele frequency (Waples 1990; Kalinowski 2004).

Genetic samples from adults were collected and preserved in 95% ethanol (axillary process; Table 2). Genetic tissues from juveniles were either preserved in 95% ethanol or sampled onto Omni swabs (Whatman FTA Product No. WHAWB100035; Sigma-Aldrich, Inc., St. Louis, MO) and dried. Ethanol-preserved tissues were either placed into individual vials or collectively into 125–500 ml containers, with 1 or more containers for each collection site for each year.

Collection information including location name, latitude, longitude, and collection year were recorded for each sample.

LABORATORY ANALYSIS

Assaying Genotypes

Genomic DNA was extracted from tissue samples using the DNeasy 96 Blood and Tissue Kit by QIAGEN (Valencia, CA). Samples were genotyped for 12 of the 13 microsatellite (µSAT) loci used in the standardized GAPS baseline (Seeb et al. 2007; Table 3). The µSAT loci were amplified via polymerase chain reaction (PCR) on Dual 384-Well GeneAmp PCR System 9700s (Applied Biosystems). The 12 loci were multiplexed into 7 PCR reactions. Each reaction consisted of a 10 µL mixture of 1X colorless GoTaq Flexi buffer, 0.2 mM each dNTP, and 0.5 units Taq DNA polymerase (Promega, Madison, WI) Primer concentrations, MgCl2 concentrations, and the corresponding annealing temperature for each primer are available in Seeb et al. (2007). The 7 PCR reactions were combined into 3 plexes for electrophoresis. Reaction plates were loaded with 0.5 µl PCR product, 0.4 µl of GeneScan 500 LIZ (AB) internal lane size standard, and 9.0 µl of Hi-Di (AB). PCR fragment analysis was completed on an Applied Biosystems 3730 capillary DNA sequencer. PCR bands were visualized, separated into bin sets, and genotypes were scored using GeneMapper (AB) software.

A genotype for a given locus and DNA sample was considered a failure if the genotype was ambiguous. Failures could be due to low quantity or low quality DNA or sample contamination. Genotypes produced were imported and archived in the Gene Conservation Laboratory Oracle database, LOKI.

Laboratory Failure Rates and Quality Control

Quality control (QC) analyses were conducted to identify laboratory errors and to measure the background discrepancy rate of the genotyping process. These analyses were performed as a separate genotyping event from the original genotyping, with staff duties altered to reduce the likelihood of repeated human errors. The QC protocol consisted of re-extracting 8% of project fish and genotyping them for the same µSATs assayed in the original project. Laboratory errors found during the QC process were corrected, and genotypes were corrected in the database. Inconsistencies not attributable to laboratory error were recorded, but original genotype scores were retained in the database. Discrepancy rates were calculated as the number of conflicting genotypes divided by the total number of genotypes compared. These rates describe the difference between original project data and QC data for all µSAT loci, and are capable of identifying extraction, assay plate, and genotyping errors. The overall failure rate was calculated by dividing the number of failed single-locus genotypes by the number of assayed single-locus genotypes. Assuming that the discrepancies among analyses were due equally to errors during original genotyping and during QC genotyping and that these analyses are unbiased, the error rate in the original genotyping was estimated as half the overall rate of discrepancies. This QC method is the best representation of the error rate of the Gene Conservation Laboratory's current genotype production.

STATISTICAL ANALYSIS

Data Retrieval and Quality Control

We retrieved genotypes from LOKI and imported them into *R*[1] with the *RJDBC* package (Urbanek 2014).[2] All subsequent analyses were performed in *R*, unless otherwise noted.

Prior to statistical analysis, we performed 2 analyses to confirm the quality of the data. First, we identified individuals that were missing substantial genotypic data because they likely had poor quality DNA. We used the 80% rule (missing data at 20% or more of loci; Dann et al. 2009) to identify individuals missing substantial genotypic data. We removed these individuals from further analyses. The inclusion of individuals with poor quality DNA might introduce genotyping errors into the baseline. Second, we identified individuals with duplicate genotypes and removed 1 of them from further analyses. Duplicate genotypes can occur as a result of sampling or extracting the same individual twice, and were defined as pairs of individuals sharing the same alleles in 100% of screened loci with genotypic data. The sample with the most missing genotypic data from each duplicate pair was removed from further analyses. If both samples had the same amount of genotypic data, the first sample was removed from further analyses.

Establishing Areas

We used available sample sizes by geographic area, and impediments associated with Devils Canyon (Table 2) to establish areas to be used to provide structure to the population genetic analyses. Samples from within these areas were combined into area collections for further testing.

Locus Statistics

We calculated the observed number of alleles, observed heterozygosity (H_o), and F_{ST} for the 12 µSAT markers based using the package *hierfstat*.[3]

Testing for Genetic Bottlenecks

Excess heterozygotes in a population can arise from genetic bottlenecks. We tested juvenile area collections for signs of genetic bottlenecks using the program BOTTLENECK (Cornuet and Luikart 1997; Piry et al. 1999). BOTTLENECK computes the distribution of expected heterozygosity (H_e) for each population and locus given the observed number of alleles and sample size for the population with the assumption of mutation-drift equilibrium. The distributions are obtained though simulating the coalescent process of genes for the Infinite Allele Model (IAM; Kimura and Crow 1964) and Stepwise Mutation Model (SMM; Kimura and Ohta 1975). Because each locus can have mutation behavior for either the IAM or SSM, we report the sign test heterozygote excess probabilities for both models (Piry et al. 1999).

---

[1] The R project for statistical computing, Vienna, Austria. Available from https://www.R-project.org/.

[2] Urbanek, S. 2014. RJDBC: Provides access to databases through the JDBC interface. R package version 0.2-5. Available from http://CRAN.R-project.org/package=RJDBC.

[3] A package for the statistical software R. HIERFSTAT: the latest version is available at http://www.unil.ch/popgen/softwares/hierfstat.htm

Assessing Validity of Using Juveniles to Represent Populations

While generally not preferred, juvenile fish can be used in population genetic analyses under certain conditions (Waples and Anderson 2017). Under some circumstances, use of juveniles may bias allele frequency estimates, affecting measurement of heterozygosity and F_{ST}. Before we used juveniles to supplement the adult collections in the population genetic analyses, we tested to see if these collections were appropriate to include. We weighted multiple lines of evidence to determine whether or not to include juvenile samples to augment adult area samples. Evidence that would move to exclude these samples included 1) the samples represented few families relative to the number of juveniles; 2) genetic differences between the juveniles and the adults from the same areas; and 3) juvenile area collections not conforming to Hardy-Weinberg expectations (HWE).

Relatedness

COLONY (Version 2.0.6.3, Jones and Wang 2010; Wang and Santure 2009) was used to determine full- and half-sibling relationships among collected samples. COLONY implements a likelihood method to infer sibling relationship (sibship) from genotypic data. We used the pure pairwise likelihood output from COLONY for inferring full- and half-sibling relationships among juvenile samples and analyzed juvenile samples from each sampling year separately. This information provided an estimate of the number of families represented by the juveniles collected at different sites for each year. See Figure 4 for settings used in COLONY.

Genetic distance between juvenile and adult collections

Pairwise F_{ST} values (Weir and Cockerham 1984) were calculated for each area between adult and juvenile collections with the package *hierfstat*. These values were compared to F_{ST} values among adult area collections to determine relative genetic distances. High relative genetic distances would signal that the juvenile Chinook salmon collections may not be appropriately used in population structure analyses.

Hardy-Weinberg expectations

For each locus within each juvenile area collection, we tested for conformance to HWE using the program *Genepop* version 4.1.4 (Rousset 2008). We combined probabilities for each collection across loci and for each locus across collections using Fisher's method (Sokal and Rohlf 1995). We assessed significance by correcting for multiple tests with Bonferroni's method (Rice 1989; $\alpha = 0.05$ / no. of collections or loci).

Population Analyses

Hardy-Weinberg expectations

We tested all area collections for HWE using the same methods we employed on juvenile area collections. If samples met HWE, we considered these area collections to be populations. If a collection did not meet HWE, we compared the expected number of heterozygotes to the observed number of heterozygotes to determine whether the deviation from HWE was biologically significant (Waples 2015). If the observed number of heterozygotes (H_o) exceeded 2 fish, we determined that the deviation was biologically significant and the collection was not considered a population for subsequent analyses. If the deviation from HWE was not deemed biologically significant (less than 2 fish), we considered the collection a population for subsequent analyses.

Removal of loci from the baseline

When testing populations for conformance to HWE we combined probabilities for each locus across populations using Fisher's method (Sokal and Rolf 1995) and examined the frequency of departures from HWE to identify loci that exhibited substantially more departures than others. If a locus had significant departures from HWE in over half of populations and a significant departure from HWE across populations after correcting for multiple tests with Bonferroni's method ($\alpha = 0.05$ / no. of loci), the locus was not used in subsequent analyses.

Allele richness

Allele richness and private allelic richness was calculated for each population using the program HP-RARE 1.1 (Kalinowski 2005). Statistical differences in allele richness and private allelic richness for all combinations of populations were tested using the Student's *t*-tests (2-tailed, paired) in Excel. We assessed significance by correcting for multiple tests with Bonferroni's method (Rice 1989; $\alpha = 0.05$ / no. of tests). We also noted the number of alleles that were private to each population.

Testing of homogeneity among populations

We tested for differentiation between populations with an exact genotypic G-test (Sokal and Rohlf 1995) using the package *hierfstat*.

Genetic distance among populations

Pairwise F_{ST} was calculated among populations using the same methods as those described for juvenile area collections.

Estimating Effective Population Sizes Within and Above Devils Canyon

Multiple lines of evidence are available to determine the magnitude of the effective population (N_e) size of Chinook salmon within and above Devils Canyon. We used COLONY to estimate the effective population size using juvenile area collections. COLONY infers N_e from sibship frequencies estimated from a sibship assignment analysis using multilocus genotypes of a sample of offspring from a single cohort in a population (Wang 2009). This method yields a much more accurate estimate of N_e than other methods for estimating N_e. See Figure 4 for settings used in COLONY.

We obtained census estimates for above and within Devils Canyon based on mark–recapture studies in 2013 and 2014, peak aerial survey counts (2013) and sonar estimates (2014) for above Devils Canyon (Study 9.7, AEA 2015). We adjusted the census estimates to estimate N_e by multiplying these estimates of N by 0.57 (Waples et al. 2010; estimate for steam-type Chinook salmon).

RESULTS

TISSUE SAMPLING

A total of 322 genetic samples were collected from spawning populations of Chinook salmon from the Middle and Upper Susitna River (Table 2; Figure 3). Target sample sizes of 95 fish were met at the 2 tributaries to the Lower River (Indian River and Portage Creek), but fewer adults were collected in many areas within and above Devils Canyon. In addition, a total of 408

genetic samples were collected from juvenile Chinook salmon in tributaries and the mainstem of the Middle and Upper Susitna River within and upstream of Devils Canyon.

LABORATORY ANALYSIS

Assaying Genotypes

A total of 301 adult and 405 juvenile Chinook salmon were selected for analysis and assayed for 12 µSAT markers (Table 2 and 3).

Laboratory Failure Rates and Quality Control

For all samples selected for analysis, the overall failure rate for genotypes was 5.45%. The overall discrepancy rate was 0.00%; therefore, the overall estimated error rate was 0.00%.

STATISTICAL ANALYSIS

Data Retrieval and Quality Control

Using the 80% rule for sufficiently complete genotypes, 81 individuals were removed from the collections for poor DNA quality. Based on the criterion for detecting duplicate individuals, 8 individuals were removed from collections as duplicate individuals.

Establishing Areas

Adequate numbers of Chinook salmon were collected in both Indian River and Portage Creek to consider these 2 sites as separate areas for population genetic analyses. Impediments associated with Devils Canyon and available samples were the primary variables used to determine areas above the first impediment in Devils Canyon. We divided this river reach into 2 areas: 1 within Devils Canyon (between the lower and upper impediments) and 1 above the last impediment in Devils Canyon.

Locus Statistics

Locus-specific statistics were calculated for adult area collections, juvenile area collections, Portage Creek, and Indian River (Table 3). The observed number of alleles ranged from 4 to 44 for adult collections and 4 to 40 for juveniles, Portage Creek, and Indian River. For adult area collections, the average H_o was 0.77 and overall F_{ST} was 0.01. For juvenile collections, Portage Creek, and Indian River, the average H_o was 0.74 and overall F_{ST} was 0.04.

Testing for Genetic Bottlenecks

Sign test results detected a significant (P<0.05) heterozygote excess in the above Devils Canyon juvenile collection (IAM: 0.001; SMM: 0.005). For the within Devils Canyon collection the test failed to reject the null hypothesis that (IAM: 0.063; SMM: 0.383). These results indicate that the juvenile collection from above Devils Canyon may show signs of genetic bottlenecks.

Assessing Validity of Using Juveniles to Represent Populations

Relatedness

We analyzed 34 (2012), 188 (2013), and 139 (2014) juveniles for sibling relationships in COLONY (Figures 5–7). All juveniles were collected within and above Devils Canyon. COLONY inferred at least 1 full-sibling relationship for 88.2% (2012), 98.8% (2013), and 90.6% (2014) of samples in each year. These full-sibling relationships represented 12 (2012), 43 (2013),

and 64 (2014) families. Sires and dams may have contributed to multiple families. In each year, 100% of samples had at least 1 half-sibling relationship. Family groups were closely associated with geographic area (Figures 5–7). These relationships were most evident from the 2013 and 2014 samples, when multiple areas were sampled (Figures 6 and 7).

Genetic distance between juvenile and adult collections

F_{ST} between juvenile collections was the highest (0.07) followed by F_{ST} between juvenile and adult collections (Table 4). F_{ST} between juvenile and adult collections ranged from 0.03 to 0.04 and between adult collections F_{ST} ranged from 0.00 to 0.01.

Hardy-Weinberg expectations

Over the 12 loci and 2 juvenile area collections, 13 of 24 tests deviated significantly from HWE (P<0.05; Table 5). Of the 13 tests that deviated from HWE, 10 occurred in the collection above Devils Canyon and 3 occurred in the collection within Devils Canyon. After adjusting for multiple tests, the collection from above Devils Canyon deviated significantly from HWE and the collection within Devils Canyon did not deviate significantly.

Decision on using juveniles to represent populations

Based on high degree of relatedness among collected juveniles, the disparity in genetic distances between juvenile and adult collections, and the lack of conformance with HWE for the juvenile collections, we determined that it is not appropriate to use juveniles to augment adult collections to represent allele frequencies for area collections.

Population Analyses

Hardy-Weinberg expectations

Over the 12 loci and 4 area collections, 8 of 48 tests deviated significantly from HWE (P<0.05; Table 6). All significant tests occurred in the collections from within and above Devils Canyon and were spread over 6 loci. After adjusting for multiple tests, only the within Devils Canyon collection deviated significantly from HWE and significant deviations (P=0.05) from HWE occurred for 5 loci. Differences in H_e and H_o were less than 2 fish for all loci. We therefore maintained all adult area collection as populations for future analyses.

Removal of loci from the baseline

No loci departed from HWE for more than 2 populations and all markers were retained in further analyses.

Allele richness

Overall allelic richness among collections for the 12 loci ranged from 2.43 (*Ots9*) to 23.88 (*Omm1080*) and was based on a minimum sample size of 28 (Table 7). The results of a Student's *t*-test revealed significant differences in allelic richness between Portage Creek and within Devils Canyon populations (P=0.05; Table 8). After adjusting for multiple tests, there were no significant differences in allelic richness between populations within and above Devils Canyon relative to populations below Devils Canyon. Private alleles occurred in all populations: 3 occurred above Devils Canyon (3 loci), 5 occurred within Devils Canyon (2 loci), 9 occurred in Portage Creek (6 loci), and 8 occurred in Indian River (4 loci). The average private allelic richness for these populations was as follows: above Devils Canyon, 0.79; within Devils Canyon, 0.54; Portage Creek, 0.82; and Indian River, 0.72. The results of a Student's *t*-test revealed no

significant difference in private allelic richness between populations within and above Devils Canyon relative to populations below Devils Canyon (Table 8).

Testing of homogeneity among populations

Tests for homogeneity between adult populations were significant (P=0.05) in 5 out of 6 tests (Table 9). The test for homogeneity for Indian River and Portage Creek was not significant.

Genetic distance among populations

F_{ST} values among adult populations were smaller than among juvenile collections (Table 4). F_{ST} ranged from 0.005 to 0.009 between populations from within and above Devils Canyon and between them and Portage Creek and Indian River. F_{ST} between Portage Creek and Indian River was less than 0.001.

Estimating effective population sizes within and above Devils Canyon

Census estimates from mark–recapture, sonar, and aerial survey studies ranged from 15 to 63 above Devils Canyon and 15 to 111 within Devils Canyon (Table 1). N_e estimates based on mark–recapture, sonar, and aerial survey estimates and the juvenile relatedness analysis ranged from 9 to 36 fish above Devils Canyon and 8 to 63 fish within Devils Canyon.

DISCUSSION

DECISION ON USING JUVENILES TO REPRESENT POPULATIONS

Using juvenile samples to represent populations can be problematic if the juveniles overrepresent some families and underrepresent others (Allendorf and Phelps 1981). In these cases, estimated allele frequencies can be biased and overly precise. This can occur if family groups remain together resulting in higher likelihoods that related individuals are sampled during sampling events. Here we review the evidence leading to our decision that juveniles should not be used in the population structure analyses.

Relatedness

The extremely high proportions of sibling and half-sibling relationships among the juvenile fish showed that either the juveniles within and above Devils Canyon school in family groups during their first year in fresh water and/or that the sampling effort was intense within locations where they occurred. The relatedness data provides evidence of both. Within sampling areas, but not among sampling locations, relatedness was high, supporting the hypothesis that family groups mix rarely among areas above Devils Canyon (Figures 5–7). The only evidence of family mixing among areas above Devils Canyon was 1 individual collected in Cheechako Creek that was related to many individuals (full and half siblings) from Chinook Creek, an upstream spawning area (Figure 7). We also see evidence that at least 1 adult spawned in 2 locations above Devils Canyon (half siblings, but no full siblings, found in between Devil Creek and Cheechako Creek; Figure 7). The collection of siblings and half siblings across multiple collection efforts (i.e., multiple minnow traps over multiple days) within areas above Devils Canyon support the hypothesis that the sampling was intense. The high levels of related individuals could bias estimates of the allele frequencies toward allele frequencies of most-sampled families.

Genetic distance between juvenile and adult collections

F_{ST} values between juvenile and adult collections were 3 to 4 times higher than the same comparisons between adult collections. F_{ST} between juvenile and adult collections ranged from 0.031 to 0.045 and between adult collections F_{ST} ranged from 0.000 to 0.009 (Table 4). This might be due to juvenile collections biased by unequal representation among families, or it might be due to large random annual fluctuation in allele frequencies (genetic drift) resulting from low numbers of spawners in the area. Bias or annual fluctuations in the juvenile allele frequencies may lead to spurious relationships among populations and, therefore, are not appropriate to use in population analyses.

Hardy-Weinberg equilibrium expectations

Genotype frequencies in the collection of juveniles from above Devils Canyon were significantly out of HWE expectations. In 11 of 12 loci, significant deviations from HWE indicated that there were more heterozygotes than expected in this collection. In populations at equilibrium, random mating should lead to similar numbers of loci with more or fewer heterozygotes than expected, but in this collection almost all loci indicted excess. This might happen above Devils Canyon by 3 processes: 1) if the parents of the sampled juveniles are the offspring from multiple cohorts of very small breeding aggregates that experience genetic bottlenecks each year, resulting in genetically divergent cohorts; 2) if the parents are new genetic migrants from multiple genetically distinct outside populations; or 3) some combination of processes 1 and 2. In our tests for bottlenecks, we detected significant heterozygote excess in the juvenile collection from above Devils Canyon; however, these results are inconclusive because we cannot distinguish between processes 1 and 2. Because of this, it would not be appropriate to use these juveniles for calculating allele frequencies for a population structure analysis.

ESTIMATING EFFECTIVE POPULATION SIZE FOR POPULATIONS WITHIN AND ABOVE DEVILS CANYON

Accurate estimates of effective population sizes are important for interpreting the results from this study. We used 3 sources of information to estimate effective population sizes and they all suffer from potential biases, mostly negative (estimates are likely lower than the true value). In addition, it should be kept in mind that the estimates only cover 2 to 3 years, so estimating effective population size over the long term is not possible. Long-term estimates of effective population sizes are calculated as the harmonic mean. The harmonic mean is disproportionately influenced downward by small numbers of spawners, as compared to the standard mean. On the other hand, the overlapping generation structure of Chinook salmon makes the effective population size larger than the annual estimates.

Estimates based on the COLONY results for the juveniles placed the effective population sizes for collections from within and above Devils Canyon in the order of 10 to 30 fish, depending on location and year. These estimates are likely biased low because of the nonrandom sampling and high proportions of related individuals (Waples and Anderson 2017). Excluding some proportion of related individuals may result in a less biased sample but removing too many yields an upwardly biased estimate. Finding this “Goldilocks number” is challenging (Waples and Anderson 2017).

Estimates based on the mark–recapture numbers are highly uncertain, due to the small sample sizes (2 fish in 2013 and 4 fish in 2014; AEA 2015).

Estimates based on the sonar data are biased low for the section of the Susitna River above Devils Canyon, because the sonar was deployed at the proposed dam site, approximately 22 miles above the upper impediment in Devils Canyon; a number of tributaries (Devil, Fog and Tsusena Creeks) where Chinook salmon have been observed drain into this stretch of the Susitna River.

Taking all of these biases and uncertainties into account, it seems reasonable that the annual effective population sizes for Chinook salmon spawning within and spawning above Devils Canyon is in the order of 10 to 100 fish for each area.

EVIDENCE FOR AND AGAINST GENE FLOW TO DEVILS CANYON

Genetic Distinctiveness

Genetic differentiation between populations from within and above Devils Canyon and the 2 most geographically proximate populations below the canyon (Indian River and Portage Creek) indicate that the within and above Devils Canyon populations are not entirely strays from these populations. Although F_{ST} values between the 2 populations below Devils Canyon and the within and above Devils Canyon populations were low, test for homogeneity of allele frequencies indicate that there is a significant difference between these populations. If the populations within and above Devils Canyon were made up solely of strays from Indian River and Portage Creek, F_{ST} between them would be lower and tests for homogeneity would not have been significant.

Allele Richness

High allele richness can only be maintained with genetic migration, given the likely low N_e of spawners within and above Devils Canyon. All measures of N_e indicate that these spawning aggregates have effective population sizes in the 10s of fish (as measured juveniles collected in this study and estimated from mark–recapture and sonar estimates of census size; Table 1). Even if the years measured in this study reflect the low end of migration of Chinook salmon into tributaries within and above Devils Canyon, it is likely that N_e is small because the effective population size is the harmonic mean of all years and would be heavily influenced by years with low numbers of spawners.

Consistent low N_e over time leads to rapid allele losses, especially for low frequency alleles. A high proportion (71%) of the µSAT alleles screened in this study had allele frequencies at or below 5% in the populations below Devils Canyon. According to the Fisher-Wright model (Hedrick 2005), allele frequencies that start out at 5% in founding populations with an effective population size of 50 fish and no straying, would lose this allele on average in 31.5 generations. Lower frequency alleles would be lost even faster (i.e., 1% frequency lost in 9 generations). Although we detected slightly lower overall levels of allelic richness in the above (11.41) and within (11.53) Devils Canyon populations than in Portage Creek (12.18) and Indian River (12.17), differences in allelic richness between pairs of populations were not significant (Tables 7 and 8). In the absence of straying, one would expect that the allelic richness within and above Devils Canyon would be significantly lower than in populations below Devils Canyon given our estimates of N_e (Table 1).

Private Alleles

Low frequencies of private alleles within and above Devils Canyon indicate that at least 1 source population is outside the Middle Susitna River (Table 7). Unfortunately, this study only collected

µSAT data from samples in the Middle and Upper Susitna River, so identifying the source of these private alleles is not possible.

Population Structure in Context with full Susitna River

Barclay et al. (*In prep.*) found that the Cheechako Creek population was more genetically similar to Chulitna River populations than it was to the other 2 Middle Susitna River populations (Portage River and Indian Creek), which were less divergent and genetically similar to Lower River populations. However, this population structure between Cheechako and Middle Susitna River populations is weak, indicated by low F_{ST} values of 0.006 (Portage Creek) and 0.009 (Indian River).

In addition, Cheechako Creek is not well characterized in the baseline, represented by only 59 samples, most (56) of which were sampled in a single year (2014). As a result, temporal stability in allele frequencies cannot be tested. Instability of allele frequencies among years could affect the clustering of this population on the genetic tree.

PUTTING TOGETHER THE EVIDENCE

Based on the lines of evidence, we can conclude that spawning aggregates within and above Devils Canyon are neither self-perpetuating populations nor solely a collection of strays from Indian River and Portage Creek. What is less certain are the proportions and composition of the genetic migrants contributing to these spawning aggregates. Evidence supporting fairly high proportions of genetic migrants spawning within and above Devils Canyon comes from 1) the low N_e estimated for these areas, 2) observations of high allele richness within these collections, 3) low-frequency private alleles, and 4) low genetic distinction from Middle Susitna populations below Devils Canyon. Additional evidence for contributions from populations outside the Middle Susitna River includes 5) the presence of private alleles, and 6) the affinity of Cheechako Creek to Chulitna River populations. However, resolving the number and sources of genetic migrants that explain the observations cannot be done with the data available.

The presence of private alleles in the spawning aggregates within and above Devils Canyon along with the relationship of Cheechako Creek and Chulitna River populations points to genetic migration from non-Middle Susitna River population(s), specifically from the Chulitna River. The Chulitna River is a steep gradient tributary to the Susitna River and may be the most hydrologically similar to Devils Canyon. In addition, some of the headwater drainages for the Devils Canyon tributaries and the Chulitna River are very close and may produce similar olfactory cues which are used for homing in Pacific salmon (Cooper et. al 1976).

Further exploration of the evidence for the different hypotheses might be possible through simulations, but these analyses are beyond the scope of this study.

ADDITIONAL DATA TO TEST COMPETING HYPOTHESES OF POPULATION STRUCTURE

The original project design called for collecting samples from juvenile and adult Chinook salmon over multiple years (minimum of 3 to 4) from multiple locations both within and above Devils Canyon (AEA 2012). This distribution and number of samples was anticipated to enable testing among 3 primary hypotheses to explain population structure of Chinook salmon above and within Devils Canyon: 1) they represent self-sustaining population(s) (Hypothesis 1a; Figure 2); 2) they are individuals originating from other geographic spawning aggregates (e.g., Portage

Creek; Hypothesis 1b); or 3) they are some combination of a local population and a nearby stock(s) (Hypothesis 2).

To distinguish between hypothesis 1a and b, the original experimental design called for testing for interannual stability in allele frequencies within locations both within and above Devils Canyon. However, funding to collect samples for this project was only provided for 2 years (2013 and 2014), and AEA opportunistically collected samples above Devils Canyon in 2012. Unfortunately, in 2013, access to sampling within Devils Canyon and above Devils Canyon below the dam site was restricted. In the end, the project had access to 10 (2012), 6 (2013), and 14 (2014) adults above Devils Canyon, and 7 (2013) and 68 (2014) adults from within Devils Canyon. These numbers of adults were inadequate to robustly test for stability in allele frequencies across years or to examine variation among locations both within and above Devils Canyon. As a result, different analyses had to be performed using the available samples from both juvenile and adult Chinook salmon collected under the Su Hydro project.

The original experimental design would have allowed for the examination of population structure among spawning aggregates at much finer geographic scales both within and above Devils Canyon and the ability to parse out changes in allele frequencies due to annual and geographic variation. Because we did not have adequate sample sizes across fine-scale geographic areas within and above Devils Canyon, samples within these broader areas had to be combined for testing. This confounded test results because differences detected between fish collected within Devils Canyon and above Devils Canyon may be due to differences in the location or year of sampling.

Because we were unable to perform the originally proposed tests, we interpreted different lines of evidence to piece together the likely forces of genetic migration and drift that have occurred between spawning aggregates within and above Devils Canyon and other Susitna River populations. Additional years of juvenile and adult collections would allow for a re-examination of our conclusions.

ACKNOWLEDGEMENTS

This study required the efforts of a large number of dedicated people. Most importantly, we would like to acknowledge the work of the people in the ADF&G Gene Conservation Laboratory, including Christy Elmaleh, Zach Pechacek, Erica Chenoweth, and Zac Grauvogel who oversaw DNA extractions, screened microsatellites, and ensured quality control of laboratory data; Judy Berger for archiving samples; Eric Lardizabal for database support; and Matt Bowes, Jason Fox, and Nathan Shoutis for spearheading the 2013 and 2014 field collections. We also would like to thank the numerous people involved in the collection of samples analyzed in this study, including samplers with R2, LGL, HDR, and volunteers from other ADF&G projects. Field collections and laboratory and statistical analyses were funded by the Alaska Energy Authority for the Susitna–Watana Hydroelectric Project.

REFERENCES CITED

AEA (Alaska Energy Authority). 2012. Revised study plan: Susitna-Watana Hydroelectric Project. FERC Project No. 14241, December 2012. Prepared for the Federal Energy Regulatory Commission by the Alaska Energy Authority, Anchorage, Alaska. http://www.susitna-watanahydro.org/study-plan.

AEA. 2015. Salmon escapement study. Susitna-Watana Hydroelectric Project. FERC Project No. 14241. Study Plan Section 9.7, Study Completion Report.

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. Canadian Journal of Fisheries and Aquatic Sciences 38(12):1507–1514.

Baker, T. T., A. C. Wertheimer, R. D. Burkett, R. Dunlap, D. M. Eggers, E. I. Fritts, A. J. Gharrett, R. A. Holmes, and R. L. Wilmot. 1996. Status of Pacific salmon and steelhead escapements in southern Alaska. Fisheries 21(10):6–18.

Banks, M. A., M. S. Blouin, B. A. Baldwin, V. K. Rashbrook, H. A. Fitzgerald, S. M Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in chinook salmon. Journal of Heredity 90(2):281–288.

Barclay, A. W., and C. Habicht. *In prep*. Genetic population structure of Chinook salmon from Northern Cook Inlet with emphasis on Susitna River: A report to Alaska Energy Authority, Susitna-Watana Hydroelectric Project. Alaska Department of Fish and Game, Anchorage.

Cooper, J. C., A. T. Scholz, R. M. Horrall, A. D. Hasler, and D. M. Madison. 1976. Experimental confirmation fo the olfactory hypothesis with homing, artificially imprinted coho salmon (Onchorhynchus kisutch). Journal of the Fisheries research board of Canada 33:703–710.

Cornuet J. M., and G. Luikart. 1997. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014.

Dann, T. H., C. Habicht, J. R. Jasper, H. A. Hoyt, A. W. Barclay, W. D. Templin, T. T. Baker, F. W. West, and L. F. Fair. 2009. Genetic stock composition of the commercial harvest of sockeye salmon in Bristol Bay, Alaska, 2006–2008. Alaska Department of Fish and Game, Fishery Manuscript Series No. 09-06, Anchorage.

Greig C., D. P. Jacobson, and M. J. Banks. 2003. New tetranucleotide microsatellites for fine-scale discrimination among endangered chinook salmon (Oncorhynchus tshawytscha). Molecular Ecology Notes 3(3):376–379.

Hedrick, P. W. 2005. Genetics of Populations. 3rd edition. Boston, MA.

Jones, O. R., and J. Wang. 2010. COLONY: A program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551–555.

Kalinowski, S. T. 2004. Genetic polymorphism and mixed-stiock fisheries analysis. Canadian Journal of Fisheries and Aquatic Sciences 61:1075–1082.

Kalinowski, S. T. 2005. HP-RARE: A computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes 5:187–189. doi: 10.1111/j.1471-8286.2004.00845.x

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.

Kimura, M. and T. Ohta. 1975. Distribution of allele frequencies in a finite population under stepwise production of neutral alleles. Proceedings of the National Academy of Sciences 72:2761–2764.

NRC (National Research Council). 1996. Upstream: salmon and society in the Pacific Northwest. Committee on the Protection and Management of Pacific Northwest Salmonids. National Academy Press, Washington D. C. doi.org/10.17226/4976

Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7:1083–1090.

Pearcy, W. G. 1992. Ocean ecology of north Pacific salmonids. University of Washington Press, Seattle.

REFERENCES CITED (Continued)

Piry S., G. Luikart, and J. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502–503.

Rexroad, C. E., R. L. Coleman, A. M. Martin, W. K. Hershberger, and J. Killefer. 2001. Thirty-five polymorphic microsatellite markers for rainbow trout (Oncorhynchus mykiss). Animal Genetics 32(5):317–319.

Rice, W. R. 1989. Analyzing tables of statistical test. Evolution 43:223–225.

Ricker, W. E. 1958. Maximum sustained yields from fluctuating environments and mixed stocks. Journal of the Fisheries Research Board of Canada 15(5):991–1006.

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

Seeb, L. W., A. Antonovich, M. A. Banks, T. D. Beacham, M. R. Bellinger, S. M. Blankenship, M. R. Campbell, N. A. Decovich, J. C. Garza, C. M. Guthrie III, T. A. Lundrigan, P. Moran, S. R. Narum, J. J. Stephenson, K. J. Supernault, D. J Teel, W. D. Templin, J. K. Wenberg, S. F. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32(11):540–552.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. 3rd edition. Freeman, San Francisco.

Waples, R. S. 1990. Temporal changes of allele frequency in Pacific salmon: implications of mixed-stock fishery analysis. Canadian Journal of Fisheries and Aquatic Sciences 47(5):968–976.

Waples, R. S., D. W. Jensen, and M. McClure. 2010. Eco-evolutionary dynamics: fluctuations in population growth rate reduce effective population size in chinook salmon. Ecology 91(3): 902–914.

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

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

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology 18:2148–2164.

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

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

Williamson, K. S., J. F. Cordes, and B. May. 2002. Characterization of microsatellite loci in Chinook salmon (Oncorhynchus tshawytscha) and cross-species amplification in other salmonids. Molecular Ecolology Notes 2:17–19.

TABLES AND FIGURES

Table 1.–Mark–recapture, sonar, and aerial spawner survey census estimates (N) and effective populations size (N_e) by brood year (BY).

	BY2011		BY 2012		BY2013		BY2014	
	N[a]	N_e[b]	N[a]	N_e[b]	N[a]	N_e[b]	N[a]	N_e[b]
Juveniles								
Above Devils Canyon	-	NA	-	20(11–38)	-	31(19–52)	-	NA
Within Devils Canyon	-	8(4–23)	-	NA	-	14(8–30)	-	NA
Mark–Recapture								
Above Devils Canyon	NA	NA	NA	NA	48	27	15	9
Within Devils Canyon	NA	NA	NA	NA	111	63	15	9
Sonar								
Above Devils Canyon	NA	NA	NA	NA	NA	NA	24	14
Within Devils Canyon	NA	NA	NA	NA	NA	NA	NA	NA
Aerial spawner survey								
Above Devils Canyon	NA	NA	NA	NA	63	36	NA	NA
Within Devils Canyon	NA	NA	NA	NA	NA	NA	NA	NA

Note: NA indicates that there are no estimates available and dashes indicate when an estimate of N is not appropriate.

[a] AEA 2015.

[b] N_e estimates for juveniles were derived from a juvenile relatedness analysis and N_e estimates for mark-recapture, sonar, and aerial spawner surveys were derived by multiplying census estimates (N) by 0.57 (estimate for stream-type Chinook salmon; Waples et al. 2010).

Table 2.–Tissue collections of adult and juvenile Chinook salmon collected in the Upper and Middle Susitna River, including collection area, collection location, the years collected, and number of samples collected, genotyped, and used in the analysis.

Area	Location	Map No.	Collection Year(s)	Number of Samples		
				Collected	Genotyped	Used
			Adult Chinook Salmon			
Above Devils Canyon						
	Kosina Creek	2	2012	10	10	10
		2	2013	3	3	3
	Kosina Creek (radio tag)	2	2014	1	1	1
	Susitna River mainstem	3	2013	1	1	1
	Tsusena Creek (radio tag)	4	2013	1	1	1
	Fog Creek	6	2014	12	12	11
	Devil Creek	7	2014	1	1	1
	Devil Creek (radio tag)	7	2013	1	1	1
Within Devils Canyon						
	Chinook Creek	8	2014	7	7	7
	Chinook Creek (radio tag)	8	2013	1	1	1
	Cheechako Creek	9	2014	57	57	57
	Cheechako Creek (radio tag)	9	2013	6	6	6
		9	2014	4	4	4
Portage Creek						
		10	2011	116	95	92
Indian River						
		11	2013	81	81	78
		11	2014	20	20	20

-continued-

Table 2.–Page 2 of 2.

Area	Location	Map No.	Collection Year(s)	Number of Samples		
				Collected	Genotyped	Used
			Juvenile Chinook Salmon			
Above Devils Canyon						
	Oshetna River	1	2013	60	60	54
		1	2014	3	3	3
	Kosina Creek	2	2013	139	139	134
		2	2014	3	3	3
	Susitna River mainstem	3	2014	32	31	28
	Tsusena Creek	4	2014	1	1	1
	Unnamed Tributary 184	5	2014	1	1	1
	Devil Creek	7	2014	14	14	14
Within Devils Canyon						
	Chinook Creek	8	2014	62	61	55
	Cheechako Creek	9	2012	35	35	34
		9	2014	58	57	34

Note: Map numbers correspond to sampling locations on Figure 2.

Table 3.–Source, observed number of alleles, observed heterozygosity (H_o), and F_{ST} for 12 microsatellite loci used to analyze adult and juvenile Chinook salmon captured in the Upper and Middle Susitna River. These summary statistics are based upon the 2 juvenile and 4 adult area collections.

		Adult area collections			Juvenile area collections, Portage Cr., and Indian R.		
Marker Name	Source[a]	Alleles	H_o	F_{ST}[b]	Alleles	H_o	F_{ST}
Ogo2	A	6	0.72	0.00	6	0.55	0.03
Ogo4	A	10	0.75	0.01	10	0.76	0.05
Oki100	B	26	0.93	0.00	26	0.94	0.04
Omm1080	C	44	0.95	0.00	40	0.93	0.04
Ots201b	D	21	0.87	0.00	20	0.86	0.02
Ots208b	D	33	0.94	0.00	33	0.94	0.03
Ots211	D	21	0.92	0.01	21	0.80	0.08
Ots212	D	17	0.87	0.00	17	0.88	0.05
Ots213	D	22	0.91	0.01	22	0.93	0.03
Ots3M	E	6	0.59	0.00	6	0.57	0.04
Ots9	E	4	0.48	0.01	4	0.49	0.00
OtsG474	F	5	0.29	0.00	5	0.22	0.00
Average/Overall		17.92	0.77	0.01	17.50	0.74	0.04

[a] A = Olsen et al. 1989; B = DFO unpublished, Department of Fisheries and Oceans Canada, Nanaimo, British Columbia, Canada, contact K. Miller; C = Rexroad et al. 2001; D = Greig et al. 2003; E = Banks et al. 1999; F = Williamson et al. 2002.

[b] Weir and Cockerham 1984.

Table 4.–Pairwise F_{ST} between adult and juvenile Chinook salmon area collections from the Upper and Middle Susitna River.

	Adults above Devils Canyon	Juveniles above Devils Canyon	Adults within Devils Canyon	Juveniles within Devils Canyon	Portage Creek (adults)	Indian River (adults)
Adults above Devils Canyon	0.000					
Juveniles above Devils Canyon	0.026	0.000				
Adults within Devils Canyon	0.005	0.047	0.000			
Juveniles within Devils Canyon	0.039	0.074	0.031	0.000		
Portage Creek (adults)	0.008	0.042	0.009	0.041	0.000	
Indian River (adults)	0.009	0.045	0.006	0.041	0.000	0.000

Source: Pairwise F_{ST} (Weir and Cockerham 1984).

Table 5.–Hardy-Weinberg equilibrium test *p*-values by locus for juvenile area collections. *P*-values less than 0.05 indicate significant deviation from Hardy-Weinberg equilibrium.

Loci	Above Devils Canyon	Within Devils Canyon
Ogo2	0.000	0.016
Ogo4	0.000	0.740
Oki100	0.000	0.135
Omm1080	0.000	0.182
Ots201b	0.000	0.799
Ots208b	0.000	0.005
Ots211	0.000	0.443
Ots212	0.000	0.029
Ots213	0.000	0.133
Ots3M	0.000	0.263
Ots9	1.000	0.597
OtsG474	0.372	1.000
Overall Loci	0.000	0.008

Table 6.–Hardy-Weinberg equilibrium (HWE) test *p*-values by locus for adult area collections. *P*-values less than 0.05 indicate significant deviation from Hardy-Weinberg equilibrium. Note that 2 markers deviated from HWE for collections within Devils Canyon. Further investigation of these significant tests revealed deviation of only 1 or 2 observations, so we determined these deviations to be biologically nonsignificant.

Locus	Above Devils Canyon	Within Devils Canyon	Portage Creek	Indian River	Overall collections
Ogo2	0.744	0.065	0.983	0.097	0.171
Ogo4	0.153	0.049	0.335	0.218	0.064
Oki100	0.019	0.053	0.683	0.468	0.032
Omm1080	0.471	0.520	0.390	0.455	0.514
Ots201b	0.244	0.042	0.813	0.726	0.288
Ots208b	0.015	0.000	0.552	0.067	0.000
Ots211	0.714	0.000	0.905	0.437	0.001
Ots212	0.762	0.186	0.877	0.179	0.413
Ots213	0.020	0.022	0.932	0.394	0.009
Ots3M	0.686	0.604	0.703	0.177	0.686
Ots9	1.000	0.376	0.402	0.340	0.680
OtsG474	1.000	0.557	1.000	0.416	0.938
Overall Loci	0.071	0.000	0.996	0.148	0.000

Table 7.–Allelic richness (AR) and private allelic richness (PAR) based on a minimum sample size of 28 and the number of observed private alleles (PA) by locus for adult area collections.

	Above Devils Canyon			Within Devils Canyon			Portage Creek			Indian Creek			Overall Collections
Locus	AR	PAR	PA	AR	PAR	PA	AR	PAR	PA	AR	PAR	PA	AR
Ogo2	5.97	0.16	0	5.37	0.00	0	5.13	0.00	0	4.99	0.00	0	5.22
Ogo4	9.93	0.45	0	7.82	0.00	0	7.82	0.00	0	8.03	0.00	0	8.25
Oki100	17.86	2.32	1	17.34	0.19	0	18.74	0.99	1	18.77	0.88	0	18.65
Omm1080	20.00	0.94	0	22.66	3.57	4	22.66	3.54	5	21.30	2.29	3	22.33
Ots201b	10.86	0.60	0	11.51	0.77	1	12.88	0.76	0	14.11	1.91	3	12.75
Ots208b	19.72	1.10	0	21.88	0.29	0	25.11	1.70	0	24.48	1.23	0	23.88
Ots211	14.83	0.82	0	14.94	0.33	0	15.31	0.54	1	14.70	0.53	0	15.35
Ots212	9.97	0.07	0	11.20	0.32	0	12.90	1.26	1	12.73	1.09	1	12.42
Ots213	17.89	2.02	1	14.83	0.15	0	16.73	0.61	0	16.62	0.29	0	16.69
Ots3M	3.97	0.00	0	3.85	0.00	0	4.15	0.30	1	4.28	0.29	1	4.11
Ots9	2.97	0.97	1	2.61	0.30	0	2.30	0.08	0	2.29	0.08	0	2.43
OtsG474	3.00	0.00	0	4.31	0.55	0	3.69	0.09	0	3.90	0.09	0	3.94
Average/Total	11.41	0.79	3	11.53	0.54	5	12.28	0.82	9	12.18	0.72	8	12.17

Table 8.–Student's *t*-test results (*p*-values) testing the significance of allele richness and private allelic richness between pairs of populations (see Table 7).

Population	Above Devils Canyon	Within Devils Canyon	Portage Creek	Indian River
Allelic richness				
Above Devils Canyon	1.000	0.817	0.177	0.207
Within Devils Canyon		1.000	0.043	0.107
Portage Creek			1.000	0.585
Indian River				1.000
Private allelic richness				
Above Devils Canyon	1.000	0.498	0.909	0.832
Within Devils Canyon		1.000	0.093	0.358
Portage Creek			1.000	0.530
Indian River				1.000

Table 9.–Summary chi-square (χ^2) values, degrees of freedom (df), and *p*-values from exact tests (G test) for genotypic differentiation between population pairs across all 12 microsatellite loci.

Collections	χ^2	df	*p*-value
Above & within Devils Canyon	49.974	24	0.001
Above Devils Canyon & Portage Creek	63.581	24	0.000
Within Devils Canyon & Portage Creek	93.516	24	0.000
Above Devils Canyon & Indian River	82.726	24	0.000
Within Devils Canyon & Indian River	76.429	24	0.000
Portage Creek & Indian River	20.225	24	0.684

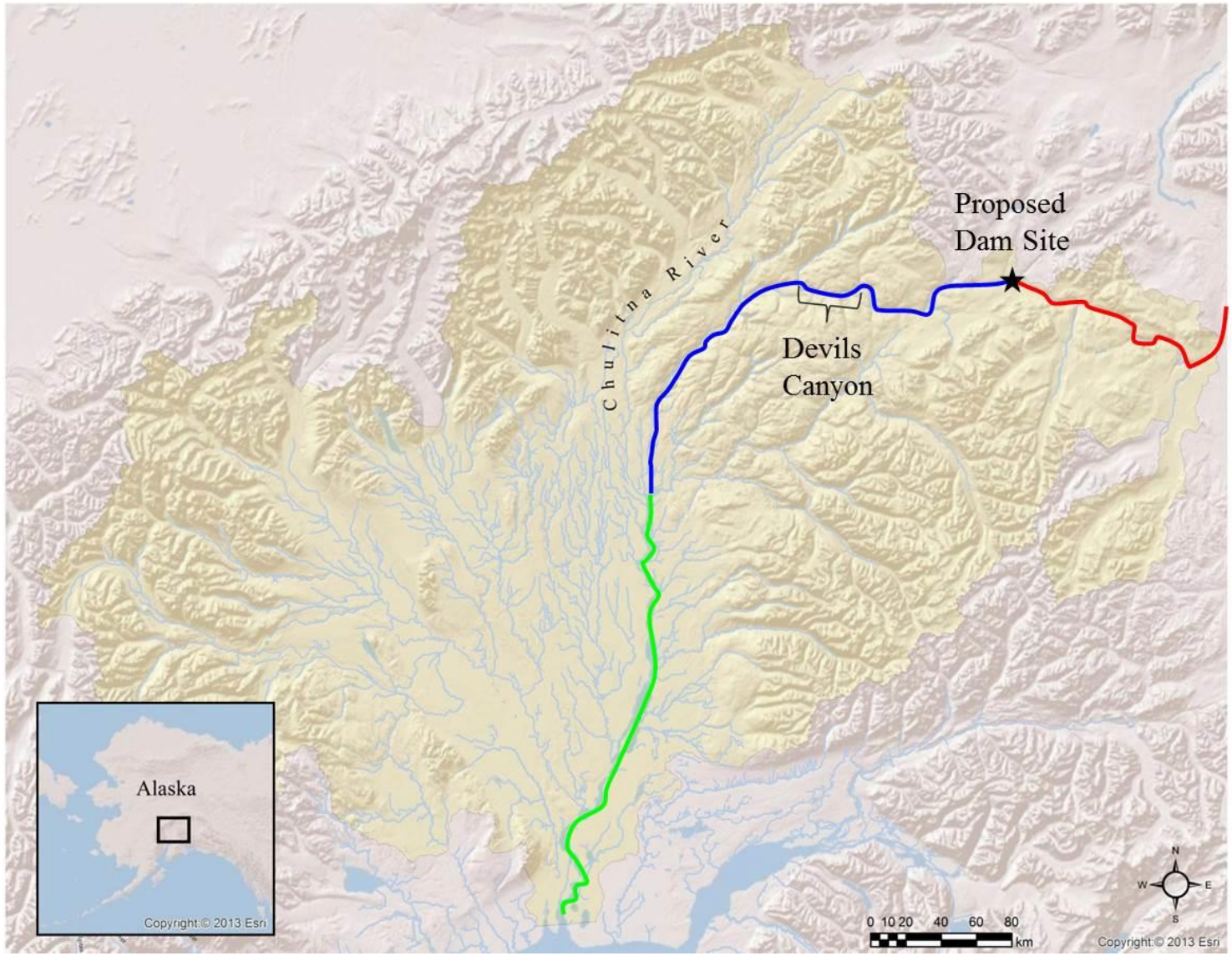

Figure 1.–Map of the Susitna River drainage (yellow) showing Upper River (red), Middle River (blue), and Lower River (green) segments.

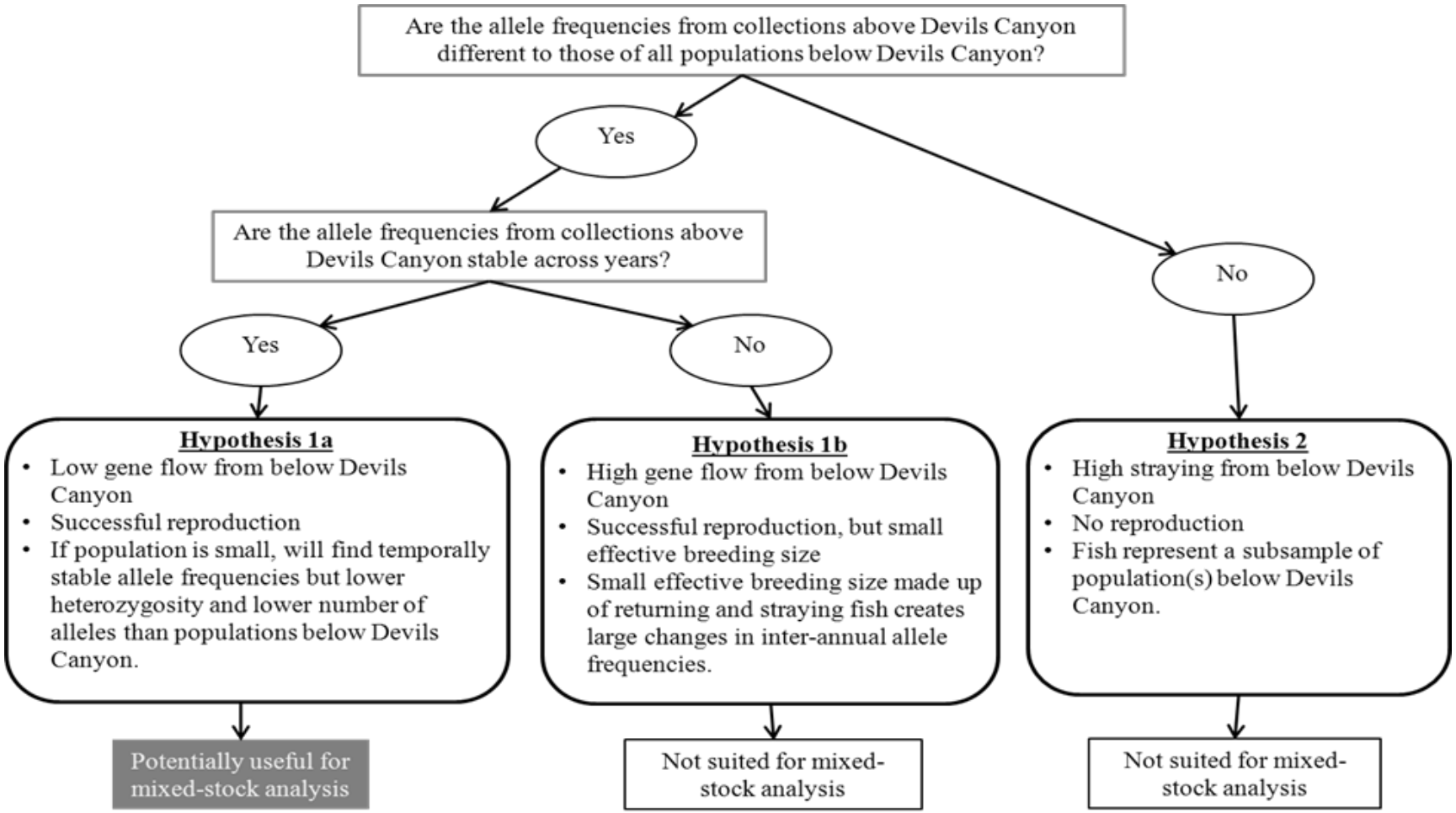

Figure 2.–A generalized flow chart to distinguish among hypotheses of population structure for Chinook salmon collected over spawning habitat above Devils Canyon in the Middle and Upper Susitna River.

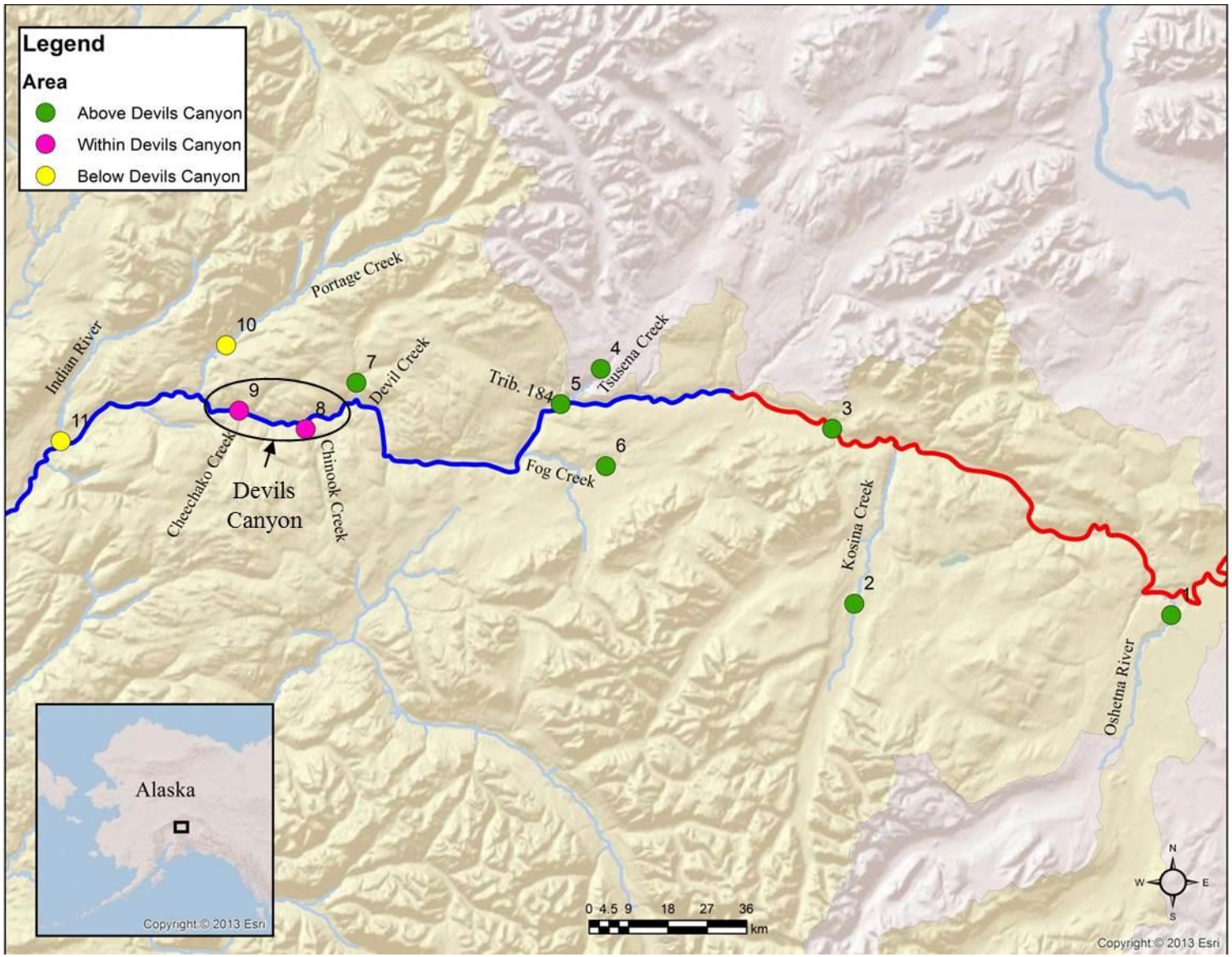

Figure 3.–Collection locations for Chinook salmon in the Upper (red) and Middle (blue) Susitna River, 2012–2014.

Note: Numbers correspond to map numbers on Table 2.

```
'C:\ZSL\Colony\test\test'

'C:\ZSL\Colony\test\test'

34      ! Number of offspring in the sample

12      ! Number of loci

1234    ! Seed for random number generator

1       ! 0/1=Not updating/updating allele frequency

2       ! 2/1=Dioecious/Monoecious species

0       ! 0/1=Inbreeding absent/present

0       ! 0/1=Diploid species/HaploDiploid species

0 0     ! 0/1=Polygamy/Monogamy for males & females

0       ! 0/1 = Clone inference = No/Yes

1       ! 0/1=Scale full sibship=No/Yes

1 3 3   ! 0/1/2/3/4=No/Weak/Medium/Strong sibship prior; 4=Optimal sibship prior for Ne

0       ! 0/1=Unknown/Known population allele frequency

1       ! Number of runs

2       ! 1/2/3/4 = Short/Medium/Long/VeryLong run

1       ! 0/1=Monitor method by Iterate#/Time in second

1       ! Monitor interval in Iterate# / in seconds

1       ! 0/1=DOS/Windows version

1       ! 0/1/2=Pair-Likelihood-Score(PLS)/Full-Likelihood(FL)/FL-PLS-combined(FPLS) method

1       ! 0/1/2/3=Low/Medium/High/VeryHigh precision


Ogo2v1,Ogo4v1,Oki100v1,Omm1080v1,Ots201bv1,Ots208bv1,Ots211v1,Ots212v2,Ots213v1,Ots3Mv1,Ots9v1,Ots
G474v1

0,0,0,0,0,0,0,0,0,0,0,0

0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000

0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100,0.0100
```

Figure 4.–Settings used for sibling relationship and effective population size analyses in the software COLONY.

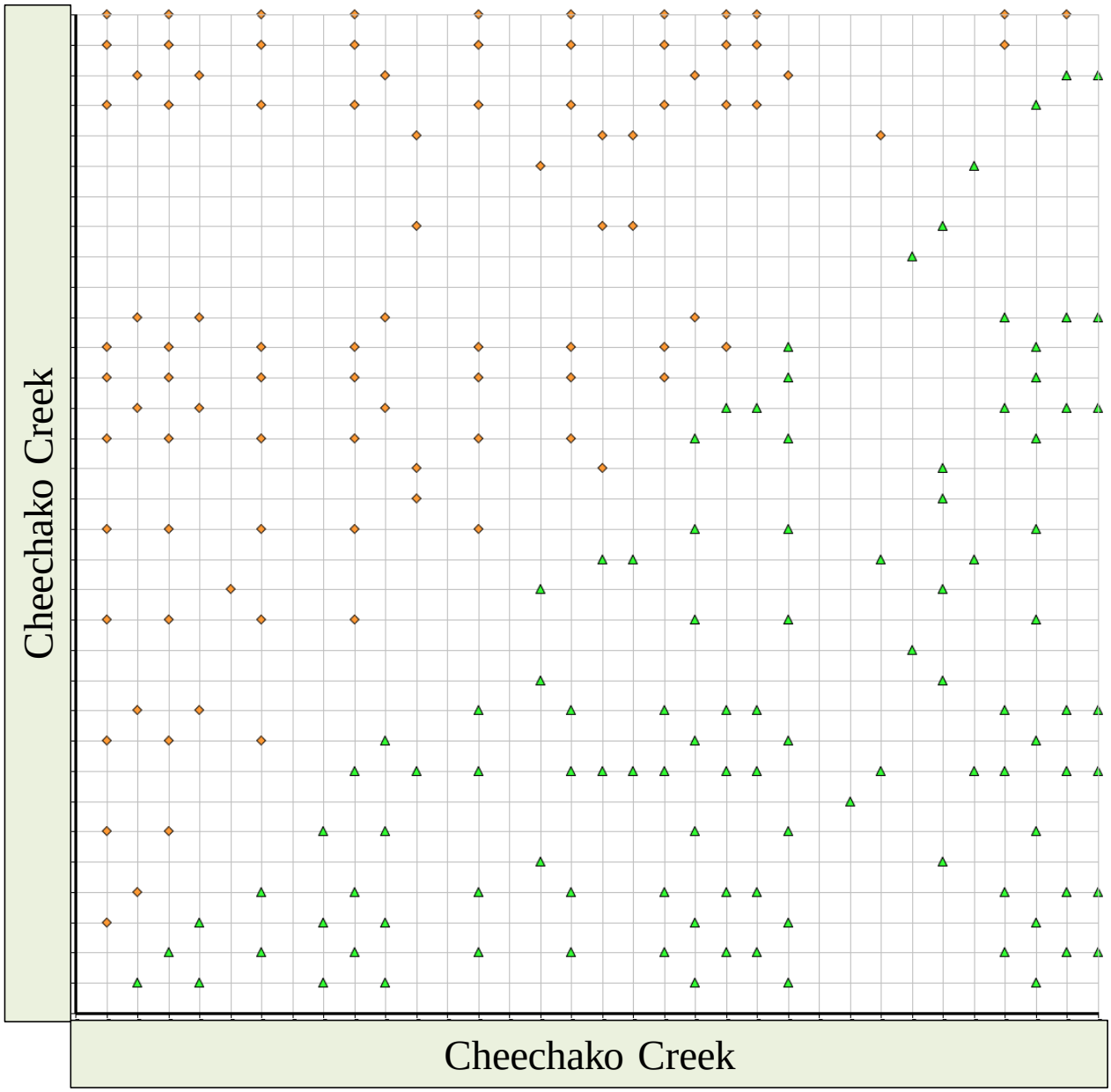

Figure 5.–Inferred pairwise sibling relationship results from COLONY for 34 juvenile Chinook salmon collected within Devils Canyon in 2012. Along both the X and Y axis is the same set of fish, in the same order, starting from the origin. Nodes on the grid indicate comparison of 2 individuals in the dataset where orange points indicate full siblings and green points indicate half siblings. The collection location is indicated on each axis.

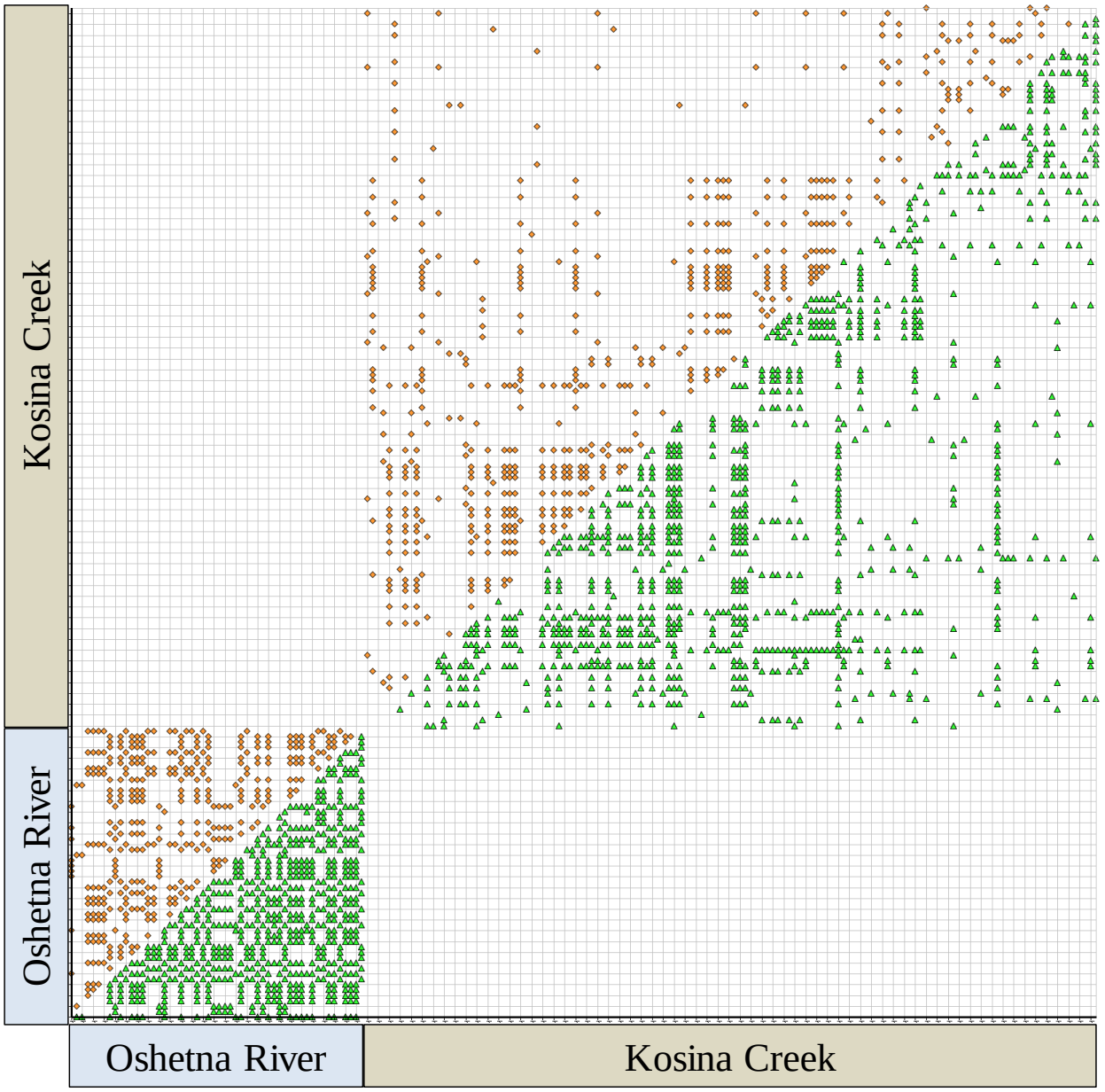

Figure 6.–Inferred pairwise sibling relationship results from COLONY for 188 juvenile Chinook salmon collected above Devils Canyon in 2013. Along both the X and Y axis is the same set of fish, in the same order, starting from the origin. Nodes on the grid indicate comparison of 2 individuals in the dataset where orange points indicate full siblings and green points indicate half siblings. The collection location is indicated on each axis.

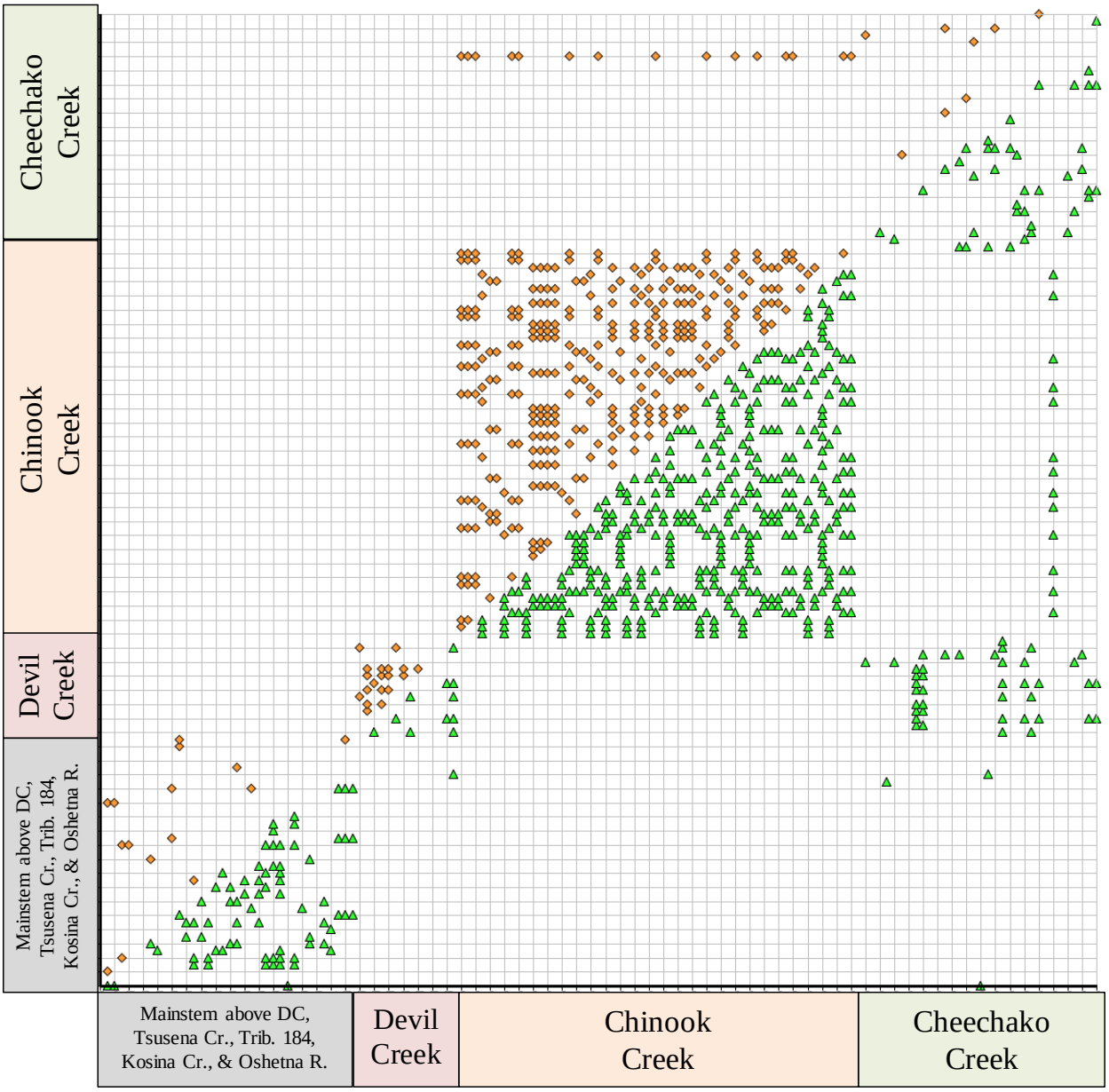

Figure 7.–Inferred pairwise sibling relationship results from COLONY for 139 juvenile Chinook salmon collected within and above Devils Canyon in 2014. Along both the X and Y axis is the same set of fish, in the same order, starting from the origin. Nodes on the grid indicate comparison of 2 individuals in the dataset where orange points indicate full siblings and green points indicate half siblings. The collection location is indicated on each axis.