Fishery Manuscript Series No. 12-02

Genetic Baseline for Upper Cook Inlet Chinook Salmon: 46 SNPs and 5,279 Fish

by

Andrew W. Barclay

Christopher Habicht

Richard A. Merizon

and

Richard J. Yanusz

May 2012

Alaska Department of Fish and Game **Divisions of Sport Fish and 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

FISHERY MANUSCRIPT SERIES NO. 12-02

GENETIC BASELINE FOR UPPER COOK INLET CHINOOK SALMON: 46 SNPS AND 5,279 FISH

by

Andrew W. Barclay, Christopher Habicht
Division of Commercial Fisheries, Gene Conservation Laboratory, Anchorage

and

Richard A. Merizon, Richard J. Yanusz
Division of Sport Fish, Palmer

Alaska Department of Fish and Game
Division of Sport Fish, Research and Technical Services
333 Raspberry Road, Anchorage, Alaska, 99518-1565

May 2012

This report was prepared under award #NA07NMF4380288 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, administered by the Alaska Department of Fish and Game. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

The Fishery Manuscript Series was established in 1987 by the Division of Sport Fish for the publication of technically-oriented results of several years' work undertaken on a project to address common objectives, provide an overview of work undertaken through multiple projects to address specific research or management goal(s), or new and/or highly technical methods, and became a joint divisional series in 2004 with the Division of Commercial Fisheries. Fishery Manuscripts are intended for fishery and other technical professionals. Fishery Manuscripts are available through the Alaska State Library and on the Internet: http://www.adfg.alaska.gov/sf/publications/ This publication has undergone editorial and peer review.

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

Christopher Habicht,
Alaska Department of Fish and Game, Division of Commercial Fisheries, Gene Conservation Laboratory
333 Raspberry Road, Anchorage, AK 99518, USA

Richard A. Merizon,
Alaska Department of Fish and Game, Division of Sport Fish,
1800 Glenn Highway, Suite 4, Palmer, AK 99645, USA

and

Richard J. Yanusz,
Alaska Department of Fish and Game, Division of Sport Fish,
1800 Glenn Highway, Suite 4, Palmer, AK 99645, USA

This document should be cited as:
Barclay, A. W., C. Habicht, R. A. Merizon, and R. J. Yanusz. 2012. Genetic baseline for Upper Cook Inlet Chinook salmon: 46 SNPs and 5,279 fish. Alaska Department of Fish and Game, Fishery Manuscript Series No. 12-02, Anchorage.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

Chinook salmon support important commercial, sport, personal use, subsistence, and educational fisheries in Cook Inlet, Alaska with annual harvests that can exceed 100,000 fish. Many populations contribute to these harvests, which often occur in areas where stocks intermingle, so the exploitation and productivity of individual stocks are not well known. This lack of knowledge compromises both protection of stocks from overharvest and utilization of stocks with fish in excess of spawning needs. Genetic analysis, using single nucleotide polymorphism (SNP) technology, holds promise in discriminating among discrete fish stocks in fishery samples if adequate stock structure exists. Here, we examine a baseline of 30 populations using 46 SNP markers to determine population structure, which provides insight into potential identifiable units (reporting groups) for mixed stock analysis (MSA). Greater genetic diversity was found in southern stocks than in northern stocks, indicating a greater potential for stock discrimination among southern stocks than in northern stocks. An ongoing study to increase the representation and the number of populations and the number of markers in the baseline will provide a more comprehensive analysis for the potential reporting groups for MSA.

Key words: Chinook salmon, Cook Inlet, *Oncorhynchus tshawytscha*, single nucleotide polymorphism (SNP), mixed stock analysis (MSA).

INTRODUCTION

BACKGROUND

Populations of Chinook salmon (*Oncorhynchus tshawytscha*) support important commercial, sport, personal use, subsistence, and educational fisheries in the Upper Cook Inlet (UCI; drainages north of Anchor Point) management area (Figure 1). Total harvests of Chinook salmon in all fisheries in UCI averaged 78,900 fish and ranged from 40,800 to 100,200 fish during the years 2000–2009 (Begich and Pawluk 2010; Bosch 2010; Oslund and Ivey 2010; Shields 2010; Szarzi et al. 2010). Most harvests occur during the homeward migration from the open ocean or in the lower reaches of river drainages, and this migration behavior results in mixed stock harvests. Without stock-specific harvest information, the exploitation and productivity of any single stock cannot be estimated. This data gap compromises the ability of the Alaska Department of Fish and Game (ADF&G) to manage for sustained yield (the policy by which ADF&G manages salmon stocks: Alaska Administrative Code 5 AAC 39.222), because understanding the return-per-spawner relationship (through the development of brood tables) for each stock provides the basis for the escapement goals and subsequent management strategy. Appropriate escapement goals allow for an adequate escapement of fish into natural streams to produce sustained yield without foregoing harvest of the surplus. Selecting a management strategy to meet an escapement goal requires knowledge of a stock's harvest by location and time.

Genetic analyses have proven successful in discriminating among discrete Pacific salmon stocks in mixed stock fishery samples (e.g. Crane et al. 2000, Miller et al. 2001, Seeb et al. 2000; 2004; Smith et al. 2005a; Templin et al. 2005; Beacham et al. 2006a; 2008; 2009; Flannery et al. 2010; Habicht et al. 2010). Recent advances in single nucleotide polymorphism (SNP) technology now allow rapid, efficient, and cost-effective analyses to discriminate stocks over a wide geographic area, making SNPs the marker of choice (e.g. Smith et al. 2007, Narum et al. 2008,Habicht et al. 2010, Barclay et al. 2010, Dann et al. 2011, Seeb et al. 2011, Templin et al. 2011).

Population structure provides the basis for successful mixed stock analysis (MSA). The population structure of Chinook salmon on a broad geographic scale is well understood, but the population structure within Cook Inlet is less well understood. In the north Pacific, Chinook salmon are separated into 2 large groups (Eastern and Western) delineated at Cape Fairweather

(Martin et al. 2010, Templin et al. 2011). These broad-scale groups are thought to reflect a secondary contact of 2 groups that were isolated in the Pleistocene. Finer-scale structure has also been detected among populations, especially among those in the Eastern group, the Yukon and Copper river drainages, and Russia (Utter et al. 1989; Teel et al. 2000; Gharrett et al. 1987 Guthrie III and Wilmot 2004; Templin et al. 2005; Beacham et al. 2006b; Ackerman et al. 2011; Templin et al. 2011). Previous studies examining population structure of Cook Inlet Chinook salmon populations were limited to Kenai and Kasilof rivers (Adams et al. 1994) and broad-scale analyses with a few representative populations (Crane et al. 1996; Teel et al. 1999; Templin et al. 2011). The most comprehensive information on population structure of Chinook salmon in Cook Inlet appears in Templin et al. (2011), in which stock structure was investigated on a large scale, and in a study by Begich et al. (2010), which focused on populations in the Kenai River drainage. The present study was designed to investigate the population structure of Chinook salmon within all of UCI to provide insights into potential MSA applications in the area.

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.

Bootstrapping. A method of resampling data with replacement to assess the variation of parameters of interest.

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.

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 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.

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

Mixed stock analysis (*MSA*). A method using allele frequencies from baseline populations and genotypes from mixture samples to estimate stock compositions of mixtures.

Polymerase Chain Reaction (*PCR*). A method to amplify DNA sequences, which can be used to generate millions of copies of the DNA.

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).

Single nucleotide polymorphism (*SNP*). DNA nucleotide variation (A, T, C, or G) at a single nucleotide site. SNPs can differ among individuals or within an individual between homologous nucleotide sites on paired chromosomes.

Stock. A locally interbreeding group of salmon (population) that is distinguished by a distinct combination of genetic, phenotypic, life history, and habitat characteristics or an aggregation of 2 or more interbreeding groups (populations) that occur within the same geographic area and are managed as a unit (from 5 AAC 39.222(f)).

OBJECTIVES

1) Collect Chinook salmon tissue samples for genetic analysis from spawning locations throughout UCI to fill gaps in the existing SNP baseline.
2) Analyze selected tissues for 46 SNP markers.
3) Examine population structure and genotypic diversity among UCI Chinook salmon populations.

METHODS

TISSUE SAMPLING

Tissue samples suitable for genetic analyses (hereafter "genetic samples") were collected, and were subsequently frozen (heart, muscle, liver, and eye; samples collected prior to 2003) or preserved in 95% ethanol (axillary fin). Frozen tissues were placed into individual vials, and ethanol-preserved samples were placed collectively into 125–500ml containers, with 1 or more containers for each collection site for each year.

Baseline genetic samples were collected from spawning populations of Chinook salmon by ADF&G personnel using gillnets, beach seines, or hook-and-line gear (Table 1; Figure 1). Target sample size for each baseline population was 95 individuals across all years to achieve acceptable precision to estimate allele frequency (Allendorf and Phelps 1981;Waples 1990a).

LABORATORY ANALYSIS

Assaying Genotypes

We extracted genomic DNA from tissue samples using a DNeasy 96 Tissue Kit by QIAGEN (Valencia, CA)[1]. We screened 46 SNP markers (Table 2) using Fluidigm 96.96 Dynamic Arrays (http://www.fluidigm.com). The Fluidigm 96.96 Dynamic Array contains a matrix of integrated

[1] Product names used in this publication are included for completeness but do not constitute product endorsement.

channels and valves housed in an input frame. On one side of the frame are 96 inlets to accept the sample DNA from individual fish and on the other are 96 inlets to accept the assays for a unique SNP marker. Once in the wells, the components are mixed under pressure applied to the chip using the IFC Controller HX (Fluidigm). The 96 samples and 96 assays are thereby systematically combined into 9,216 parallel reactions. Each reaction consists of a mixture of 4µl of assay mix (1× DA Assay Loading Buffer (Fluidigm), 10× TaqMan SNP Genotyping Assay (Applied Biosystems), and 2.5× ROX (Invitrogen)) and 5µl of sample mix (1× TaqMan Universal Buffer (Applied Biosystems), 0.05× AmpliTaq Gold DNA Polymerase (Applied Biosystems), 1× GT Sample Loading Reagent (Fluidigm), and 60–400 ng/µl DNA) combined in a 7.2 nL chamber. To increase loading efficiency and reduce liquid handling, 96 well plates of assay mix and reaction mix (working stocks) were made to load up to 40 and 100 Dynamic Arrays, respectively. Thermal cycling was performed on an Eppendorf IFC Thermal Cycler as follows: 70°C for 30 min for "Hot-Mix" step, initial denaturation of 10 min at 96°C followed by 40 cycles of 96°C for 15 s and 60°C for 1 min. The Dynamic Arrays were read on a Fluidigm EP1 System after amplification and scored using Fluidigm SNP Genotyping Analysis software.

Assays that failed to amplify on the Fluidigm system were reanalyzed on the Applied Biosystems platform. Each reaction on this platform was performed in 384-well reaction plates in a 5µL volume consisting of 5–40 ng/µl of template DNA, 1× TaqMan Universal PCR Master Mix (Applied Biosystems), and 1× TaqMan SNP Genotyping Assay (Applied Biosystems). Thermal cycling was performed on a Dual 384-Well GeneAmp PCR System 9700 (Applied Biosystems) as follows: an initial denaturation of 10 min at 95°C followed by 50 cycles of 92°C for 1s and annealing/extension temperature for 1 min. The plates were scanned on an Applied Biosystems Prism 7900HT Sequence Detection System after amplification and scored using Applied Biosystems Sequence Detection Software (SDS) version 2.2.

Genotypes produced on both platforms were imported and archived in the Gene Conservation Laboratory Oracle database, LOKI.

Laboratory Failure Rates and Quality Control

The overall failure rate was calculated by dividing the number of failed single-locus genotypes by the number of assayed single-locus genotypes. An individual genotype was considered a failure when a locus for a fish could not be satisfactorily scored.

Quality control (QC) measures were instituted to identify laboratory errors and to determine the reproducibility of genotypes. In this process, 8 of every 96 fish (1 row per 96-well plate) were reanalyzed for all markers by staff not involved with the original analysis.

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.

Assuming that the inconsistencies among analyses (original vs. QC genotyping) were due equally to errors in original genotyping and errors during the QC genotyping and that these analyses are unbiased, error rates in the original genotyping were estimated as ½ the rate of inconsistencies. Because baseline collections were genotyped during several separate laboratory analyses, we report quality control results for a single recent laboratory baseline analysis as a representative of the QC on the entire dataset.

STATISTICAL ANALYSIS

Data Retrieval and Quality Control

We retrieved genotypes from LOKI and imported them into *R* (*R* Development Core Team 2011) with the *RODBC* package (Ripley 2010). All subsequent analyses were performed in *R*, unless otherwise noted.

Prior to statistical analysis, we performed 3 analyses to confirm the quality of the data. First, we identified SNP markers that were invariant in all individuals or that had very few individuals with the alternate allele in only 1 collection. We excluded these markers from further statistical analyses. Second, 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 and reduce the accuracies of mixed stock analyses.

The final QC analysis identified individuals with duplicate genotypes and removed 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 95% of screened loci. 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.

Baseline Development

Hardy-Weinberg expectations

For each locus within each collection, we tested for conformance to Hardy-Weinberg expectations (HWE) using Monte Carlo simulation with 10,000 iterations in the *Adegenet* package (Jombart 2008). We combined probabilities for each collection across loci and for each locus across collections using Fisher's method (Sokal and Rohlf 1995) and removed collections and loci that violated HWE from subsequent analyses after correcting for multiple tests with Bonferroni's method ($\alpha = 0.05$ / # of collections).

Pooling collections into populations

When appropriate, we pooled some collections to obtain better estimates of allele frequencies following a step-wise protocol. First, we pooled collections from the same geographic location, sampled at similar calendar dates but in different years, as suggested by Waples (1990b). We then tested for differences in allele frequencies between pairs of geographically proximate collections that were collected at similar calendar dates and that might represent the same population. We defined collections as being "geographically proximate" if they were within the same river. We used Fisher's exact test (Sokal and Rohlf 1995) of allele frequency homogeneity and based our decisions on a summary across loci using Fisher's method. When these tests indicated no difference between collections ($P > 0.01$), we pooled them. When all individual collections within a pooled collection were geographically proximate to other collections, we followed the same protocol until we found significant differences between the pairs of collections being tested. After this pooling protocol, we considered these final collections as populations. Finally, we tested populations for conformance to HWE following the same

protocol described above to ensure that our pooling was appropriate, and that tests for linkage disequilibrium would not result in falsely positive results due to departure from HWE.

Linkage disequilibrium

We tested for linkage disequilibrium between each pair of nuclear markers in each population to ensure that subsequent analyses would be based on independent markers. We used the program *Genepop* version 4.0.11 (Rousset 2008) with 100 batches of 5,000 iterations for these tests. We summarized the frequency of significant linkage disequilibrium between pairs of SNPs ($P < 0.05$). We considered pairs to be linked if they exhibited linkage in more than half of all populations.

Analysis of Genetic Structure

Temporal variation

We examined the temporal variation of allele frequencies with a hierarchical, 3-level analysis of variance (ANOVA). We treated the temporal samples as sub-populations based on the method described in Weir (1996). This method allowed the quantification of the sources of total allelic variation and permitted the calculation of the among-years component of variance and the assessment of its magnitude relative to the among-population component of variance. This analysis was conducted using the software package *GDA* (Lewis and Zaykin 2001).

Hierarchical log-likelihood tests

We examined genetic diversity within Cook Inlet with a 3-level hierarchical log-likelihood ratio (G) analysis. Populations were grouped hierarchically into 8 fine-scale groups based on drainage and geographic features: 1) "West" (populations from spawning streams that drain into western UCI south of the Susitna River), 2) "Susitna" (Susitna River populations), 3) "Yentna" (Yentna River populations, 4) "Knik" (populations from spawning streams that drain into Knik Arm), 5) "Turnagain" (populations from spawning streams that drain into Turnagain Arm), 6) "Kenai" (Kenai River populations), 7) "Kasilof" (Kasilof River populations), and 8) "SKenaiPen" (Kenai Peninsula spawning streams that drain into UCI south of the Kasilof River) (Table 1; Figure 1). These were further grouped into 2 broad-scale regions: 1) "Northern" (West, Susitna, Yentna, Knik, and Turnagain groups) and 2) "Southern" (Kenai, Kasilof, and SKenaiPen groups). We tested for homogeneity of allele frequencies within groups, among groups within regions, and between regions. To compare levels of heterogenity between regions and groups, scaled G-statistics (G′) were calculated by dividing G by degrees of freedom.

Visualization of genetic distances

We took 2 approaches to visualizing genetic distances among collections. Both approaches were based on pairwise F_{ST} estimates from the final set of independent markers with the package *hierfstat* (Goudet 2006). The first approach was to construct 1,000 bootstrapped neighbor-joining (NJ) trees by resampling loci with replacement to assess the stability of tree nodes. We plotted the consensus tree with the *APE* package (Paradis et al. 2004). While these trees provided insight into the variability of the genetic structure of these collections, pairwise distances visualized in 3 dimensions were more intuitive. In a second approach, we plotted pairwise F_{ST} in a multidimensional scaling (MDS) plot using the package *rgl* (Adler and Murdoch 2010).

RESULTS

TISSUE SAMPLING

A total of 6,758 genetic samples were collected from spawning populations of Chinook salmon throughout UCI (Table 1; Figure 1). In 1992 and 1995, samples were collected from locations in the Susitna River drainage (3), Matanuska River (1), and Kasilof River (1) (Crane et al. 1996; Teel et al. 1999). Between 2003 and 2007, 27 collections were made primarily from locations on the Kenai Peninsula (Begich et al. 2010). Between 2008 and 2011, an additional 71 collections were made focusing on unrepresented and underrepresented areas of UCI. The 104 collections were taken at 55 locations throughout UCI drainages; individuals from 30 of these locations were collected in multiple years. Target sample sizes of 95 fish were met at 29 locations.

LABORATORY ANALYSIS

Assaying Genotypes

A total of 5,279 fish collected over spawning areas were selected for analysis and assayed for 46 SNP markers (Table 1).

Laboratory Failure Rates and Quality Control

For all 63 collections in the baseline, the overall failure rate for genotypes at the 46 SNP markers was 2%. The recent baseline project included 1,229 individuals (~24% of the current baseline). The discrepancy rate for this project was 0.18%; therefore the overall error rate was 0.09%.

STATISTICAL ANALYSIS

Data Retrieval and Quality Control

For all analyzed collections, a total of 4 SNP markers were found to be invariant among all individuals, and 2 SNP markers were invariant for all but 2 individuals (Table 2). These markers were removed from further analyses. Based upon the 80% rule, 3.01% of individuals were removed from the baseline collections. Based on the 95% of loci criterion for detecting duplicate individuals, 0.56% of individuals were removed from baseline collections as duplicate individuals. No duplicate individuals were detected in 48 of the 63 baseline collections (76%).

Baseline Development

Hardy-Weinberg expectations

Over all nuclear markers and collections, 36 of 2,835 tests deviated significantly from HWE ($P < 0.01$) without adjusting for multiple tests. These were spread over 25 markers, and no markers were out of HWE in more than 4 of the 63 collections. No collections departed HWE at more than 4 of the 45 markers. After adjusting for multiple tests, all collections conformed to HWE.

Pooling collections into populations

A total of 30 populations were identified after pooling collections taken at the same geographic location over multiple years and geographically proximate collections (pooled collections and collections taken at different sites are referred to as “populations”; Table 1). Over all variant markers and populations, 30 of 1,200 tests did not conform to HWE ($P < 0.01$) without adjusting for multiple tests. These were spread over 17 markers, and no markers were out of HWE in more

than 6 of the 30 populations. After adjusting for multiple tests, 2 SNP markers (*Ots_SERPC1-209* and *Ots_il-1racp-166*) did not conform to HWE and were removed from further analyses. No population was out of HWE at more than 4 of 40 markers. After adjusting for multiple tests, all populations conformed to HWE.

Linkage disequilibrium

In the tests for linkage disequilibrium, no SNP pairs showed significant linkage ($P < 0.05$) in more than 50% of populations.

Analysis of Genetic Structure

Temporal variation

Twenty populations were included in the analysis of temporal variation of allele frequencies. Allele frequencies for all populations appeared to be temporally stable. Within populations, 17 pairs of collections were 1–4 years apart, and 3 were 13–14 years apart (Table 1). The 3-level ANOVA indicated that the ratio of variation among temporal collections to the variation among populations was 5.2%.

Hierarchical log-likelihood test

In the analysis of genetic heterogeneity, grouping populations into 7 fine-scale groups and 2 broad-scale regions, significant variation was found within each group, among groups, and between regions (Table 3). Within the Northern region, Susitna ($G' = 9.85$) had greatest genetic heterogeneity among populations, followed by the Knik ($G' = 3.11$), Yentna ($G' = 2.72$), and West ($G' = 2.20$) population groups. A test could not be conducted on the Turnagain group, because it included only 1 population. Within the Southern region, the Kasilof ($G' = 22.32$) group had the greatest genetic heterogeneity followed by the Kenai ($G' = 13.92$) and SKenaiPen ($G' = 1.71$) groups, which had the lowest genetic heterogeneity of the 7 groups. Greater among-group heterogeneity was found in the Southern region.

Visualization of genetic distances

The first approach to visualize genetic relationships among baseline populations is shown in an NJ tree (Figure 2). In general, populations clustered with other populations within the same group; however, several populations clustered with populations outside of their group. Moose Creek, in the Knik group, is more genetically similar to Susitna populations than to other populations in the Knik group. Deshka River clusters with populations from the West group (Chuitna River and Straight Creek). Chickaloon River, located on the Kenai Peninsula, is more genetically similar to Knik populations (Ship Creek and Little Susitna River) than to other Kenai Peninsula populations. Kasilof River populations group more closely to other Kenai Peninsula populations than to each other; Kasilof River mainstem groups with Kenai River mainstem and Crooked Creek groups with Slikok Creek and SKenaiPen populations. Several populations appear to be more genetically distinct (on longer genetic branches): Chulitna River, Middle Fork; Deshka River; Russian River; and Kasilof River mainstem. Well supported nodes (>50% of bootstrap trees) occurred almost entirely in the Southern region (7 out of 8 significant nodes).

The second approach to visualizing genetic relationships among baseline populations is shown in MDS plots (Figures 3 and 4). In an MDS plot of the entire baseline, populations clustered into Northern and Southern regions (Figure 3). However, a Knik population (Little Susitna River) was intermediate between the regional clusters. As with the tree, the MDS also shows the

Chulitna River, Middle Fork; Deshka and Russian rivers as genetically distinct from other populations (population numbers (#) 4, 9, 19; Figures 3 and 4). Sunflower Creek (#11) also stands out as an outlier in Figure 3 but not so much in Figure 4. Fine-scale MDS plots of each region showed similar genetic relationships among populations to those in the NJ tree (Figure 4). In the Northern region, Yentna River populations (Talachulitna River and Sunflower Creek; #10, 11) do not group closely with Susitna populations (#5-9) and Talachulitna River shows genetic similarity to West population, Coal Creek (#3). Crooked Creek (#27; Kasilof group) clusters closely with Slikok Creek (#25; Kenai group). Although the Kasilof River mainstem (#26) and Middle Kenai River mainstem (#24) populations clustered with high bootstrap values in the NJ tree (Figure 2), they appear genetically distinct from one another in the MDS (Figure 4). The long branch length for the Kalilof River mainstem population in the NJ tree is congruent with the relationship in the MDS.

DISCUSSION

COMPARISONS TO PREVIOUS FINDINGS

This is the most comprehensive set of baseline samples collected to test specifically for fine-scale population structure among Chinook salmon returning to streams in UCI. Previous studies included collections from a limited number of drainages within UCI. In one of the earliest studies, Adams et al. (1994) used mitochondrial DNA and allozyme markers to discriminate between Kenai and Kasilof rivers early- and late-run Chinook salmon based on 400 samples representing 4 populations. In the first broad-scale study of genetic structure containing Cook Inlet populations, Crane et al. (1996) found significant heterogeneity among populations from the Susitna, Kenai, and Kasilof rivers using allozyme loci based on 496 samples representing 6 populations. Teel et al. (1999) used the same samples and populations from Crane et al. (1996) in an updated coastwide allozyme baseline, which was later used to analyze mixtures of fish caught in southeast Alaska fisheries (Crane et al. 2000). Another study focused on the Kenai River drainage and found differences adequate for MSA within the drainage among 4 groups of populations (Lower Kenai River tributaries, Kenai river mainstem, Killey River, and Quartz Creek; Begich et al. 2010). The Begich et al. (2010) study was based on 977 samples representing 9 populations. The most recently published study focused on broad-scale genetic structure of Chinook salmon populations around the Pacific Rim for high seas MSA and found differences adequate for MSA within the Kenai River and between the Kenai River, Kasilof River, Lower Kenai Peninsula rivers, and the Susitna River drainages and was based on 2,699 samples representing 15 populations (Templin et al. 2011). This current baseline represents a doubling of both the number of samples and the number of populations analyzed (5,279 samples, 30 populations).

Genetic relationships among populations within the current study agree with previous findings. Genetic relationships between early- (tributary) and late-run (mainstem) populations in the Kenai and Kasilof rivers are similar to those described in Adams et al. (1994); Crooked Creek clusters with Kenai River tributary populations and Kenai and Kasilof rivers mainstem populations cluster together. Genetic relationships among the Kenai River populations are similar to those described with microsatellites (Begich et al. 2010); upper and mainstem Kenai River populations form a cluster, lower Kenai River populations form another cluster, and Crescent Creek is divergent from either. Genetic relationships among Kasilof, Kenai, and SKenaiPen populations were also consistent with those described in Templin et al. (2011); Kasilof and SKenaiPen

populations clustered with Kenai populations below a node with good bootstrap support. The Kasilof River mainstem population clustered with the middle Kenai River mainstem population and the lower Kenai River tributary populations form a cluster with the Kasilof River tributary population (Crooked Creek) and SKenaiPen populations. Finally, the results confirm the divergence between the Susitna and Yentna populations and the Kenai Peninsula populations that was first reported in Templin et al. (2011).

NEW FINDINGS

Relationships among populations that were not previously examined provide additional insight into population structure within UCI. This is the first study to include samples from the West, Knik, and Turnagain groups. These additional populations all clustered with the Susitna and Yentna populations, making up the Northern region (Figure 3). However, we found little segregation among these groups (Figures 2 and 4) with some affinity between West and Yentna populations. For example, the Deshka River in the Susitna River Drainage, clustered with the Chuitna River in the West (Figure 2); however, the populations appear to be genetically distinct from one another (Figure 4). Overall, the Susitna group has the most divergent populations within the Northern region. Within this region, the Chulitna River was the most divergent, followed by the Deshka River and Prairie Creek.

Relationships among some of the populations in this study were not expected given their geographic proximity. For example, the Moose Creek population located in the Matanuska River, clusters with Susitna River populations instead of clustering with Knik populations (Figures 2 and 4). In the 1920s Moose Creek was realigned to allow for construction of a railroad, which resulted in a waterfall that blocked fish passage. Recently, Chickaloon Village Traditional Council, NOAA Restoration Center, and United States Fish and Wildlife Service were able to recreate the historical channel and fish are now using the restored channel (http://www.fakr.noaa.gov/habitat/restoration/moosecreek06.pdf). During the interim, the Chinook salmon population was likely reduced or extirpated. A combination of a small effective population size and low straying from the nearby Susitna River may have resulted in a population that is genetically similar to other Susitna River populations. Another unexpected relationship is between Kenai River populations and the Chickaloon River on the Kenai Peninsula. Although the headwaters of Chickaloon River are within a mile of the headwaters to Juneau Creek (Kenai River drainage), this population is more genetically similar to populations in the Northern region.

This study also contained new populations from groups previously studied, and these populations add texture to the results of the previous analyses. In the Kenai River, this study added Upper Kenai river mainstem and Russian River. The new mainstem population clustered with Juneau Creek and the middle mainstem collection (Figure 4). Genetic similarities between the Juneau Creek and Upper Kenai River mainstem populations may be due to proximity and similarity of the two spawning areas. Juneau Creek has a barrier waterfall about 3 kilometers from its mouth on the Kenai River. As a result, the spawning habitats within Juneau Creek and the Upper Kenai River mainstem are contiguous and geomorphically similar to each other allowing for little genetic isolation. The Russian River population, on the other hand, was highly divergent from all other Kenai River populations. The Russian River spawning habitat is above a difficult-to-navigate waterfall (Engel 1972) and is geomorphically divergent from mainstem spawning sites (which are lower order streams). The genetic divergence of the Russian River population

parallels the pattern observed for sockeye salmon in this drainage and is thought to be due to a combination of a more recent founding of this population and smaller current effective population size because of migration barriers (Allendorf and Seeb 2000).

The most notable finding of this study is the differences in the amounts of genetic diversity within the Northern and Southern regions. These differences in the amount genetic diversity within regions are visualized in the overall MDS plot, where, except for a few outlier populations, populations in the Northern region form a tight cluster (Figure 3). Results from the hierarchical log-likelihood tests support these findings, where within region heterogeneity in the Northern region ($G' = 6.37$) is less than half of that in the Southern region ($G' = 14.57$; Table 3). Although the Southern region represents a much smaller geographic area, populations in the Southern region are spaced farther apart indicating greater genetic diversity than the Northern region. Within the Southern region, the Kenai River drainage contains the greatest genetic diversity. This greater genetic diversity can be seen in the Southern region MDS plot, where Kenai River populations are spread out from one end of the plot to the other (Figure 4).

POTENTIAL FOR MSA

The potential for MSA among regions within UCI looks promising based on these genetic relationships and on previous baseline MSA performance. In proof tests, using a similar set of SNPs, Templin et al. (2011) demonstrated support for the following reporting groups: Kenai River early (Killey River, Benjamin Creek, Funny River and Slikok Creek), Kenai River late (Kenai River mainstem and Juneau Creek), Kasilof River (Kasilof River mainstem and Crooked Creek), Lower Kenai Peninsula rivers (Ninilchik River and Anchor River), and the Susitna River drainages (Deshka River, Willow Creek, Prairie Creek, and Talachulitna River). Microsatellite data reported in Begich et al. (2010) also indicated that Quartz Creek is divergent from other Kenai River populations, and that this population met the criteria for reporting group in simulations. However, Begich et al. (2010) did not have a sufficient number of individuals from the Russian River to include in their analysis or to perform mixture simulations.

Testing performance of reporting groups for MSA will provide more definitive conclusions regarding the use of this baseline to answer management-related questions. These tests will be performed after additional collections are added and markers screened under a related project and over the next year or two. In addition to the analyzed samples, collection efforts in the current project yielded an additional 1,479 samples from potentially 20 additional populations that are suitable for screening and incorporating into the baseline (Table 1). In 2012, additional effort will be made to fill out potential populations that are represented by fewer than 50 fish and to sample additional spawning sites. Finally, an additional set of SNPs has been developed under the Southeast Sustainable Salmon Fund project number 44515 “High Resolution SNPs for Chinook” and is being used to screen a subset of collections from this study. Results from these tests may yield loci that may be under selection and therefore may be more likely to differentiate among reporting groups that do not show adequate structure with neutral loci. Extracted DNA is available for screening for some such additional loci, as they become available.

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: Lisa Fox, Serena Rogers, Heather Liller, Heather Hoyt, Paul Kuriscak, Tara Harrington, and Zac Pechacek, who oversaw DNA extractions, screened SNPs, and ensured quality control of laboratory data; Judy Berger for archiving samples; Eric Lardizabal for database support, Chase Jalbert, Cassey McConnell, and Jed Konsor for spearheading field collections in 2011. We would like to thank Don Reeves, Kiche and Kaasan Braendel, Darly Lescanec, Leslie and Gabrielle Vail, Chase Jalbert, Clint McBride, Nick Cassara, Matt Warnke, and Ingrid Jensen for their sampling efforts during 2008–2010. We also would like to thank people involved in the collection of samples analyzed in this study; collected over a 17-year period prior to the inception of this project. Finally, we would like to thank Stewart Grant for providing editorial review that increased the clarity of the report. Field collections and laboratory and statistical analyses were funded by State of Alaska and Alaska Sustainable Salmon Fund project numbers 44517 "West Cook Inlet Chinook Baseline" and 45864 "Northern Cook Inlet Chinook GSI". This report was prepared under award #NA07NMF4380288 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, administered by the Alaska Department of Fish and Game.

REFERENCES CITED

Ackerman, M. W., C. Habicht, and L. W. Seeb. 2011. SNPs under diversifying selection provide increased accuracy and precision in mixed stock analyses of sockeye salmon from Copper River, Alaska. Transactions of the American Fisheries Society. 140: 865-881.

Adams, N. S., W. J. Spearman, C. V. Burger, K. P. Currens, C. B. Schreck, and H. W. Li. 1994. Variation in mitochondrial DNA and allozymes discriminates early and late forms of Chinook salmon (*Oncorhynchus tshawytscha*) in the Kenai and Kasilof Rivers, Alaska. Canadian Journal of Fisheries and Aquatic Sciences 51 (Supplement 1): 172-181.

Adler, D., and D. Murdoch. 2010. 3D visualization device system (OpenGL). R package version 0.91. http://CRAN.R-project.org/package=rgl (Accessed November 11, 2011).

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

Allendorf, F. W., and L. W. Seeb. 2000. Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mitochondrial DNA markers. Evolution 54(2): 640–651.

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.

Barclay, A. W., C. Habicht, T. Tobias, and T. M. Willette. 2010. Genetic stock identification of Upper Cook Inlet sockeye salmon harvest, 2009. Alaska Department of Fish and Game, Fishery Data Series No. 10–93, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FDS10-93.pdf

Beacham, T. D., J. R. Candy, K. L. Jonsen, J. Supernault, M. Wetklo, L. Deng, K. M. Miller, R. E. Withler, and N. Varnavskaya. 2006a. Estimation of stock composition and individual identification of Chinook salmon across the Pacific Rim by use of microsatellite variation. Transactions of the American Fisheries Society 135(4): 861-888.

REFERENCES CITED (Continued)

Beacham, T. D., J. R. Candy, C. Wallace, S. Urawa, S. Sato, N. V. Varnavskaya, K. D. Le, and M. Wetklo. 2009. Microsatellite stock identification of chum salmon on a Pacific Rim basis. North American Journal of Fisheries Management 29(6): 1757-1776.

Beacham, T. D., K. L. Jonsen, J. Supernault, M. Wetklo, L. Deng, and N. Varnavskaya. 2006b. Pacific Rim population structure of Chinook salmon as determined from microsatellite analysis. Transactions of the American Fisheries Society 135(6): 1604–1621.

Beacham, T. D., M. Wetklo, C. Wallace, J. B. Olsen, B. G. Flannery, J. K. Wenburg, W. D. Templin, A. Antonovich, and L. W. Seeb. 2008. The application of microsatellites for stock identification of Yukon River Chinook salmon. North American Journal of Fisheries Management 28(1): 283–295.

Begich, R. N., and J. A. Pawluk. 2010. 2008-2010 Recreational fisheries overview and historical information for North Kenai Peninsula: fisheries under consideration by the Alaska Board of Fisheries, February 2011. Alaska Department of Fish and Game, Fishery Management Report No. 10-51, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FMR10-51.pdf

Begich, R. N., W. D. Templin, A. W. Barclay, and L. W. Seeb. 2010. Development of microsatellite genetic markers for Kenai River Chinook salmon. Alaska Department of Fish and Game, Fishery Data Series No. 10-38, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FDS10-38.pdf

Bosch, D. 2010. Area management report for the recreational fisheries of Anchorage, 2009 and 2010. Alaska Department of Fish and Game, Fishery Management Report No. 10-53, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/Fmr10-53.pdf

Crane, P. A., W. D. Templin, D. M. Eggers, and L. W. Seeb. 2000. Genetic stock identification of Southeast Alaska chinook salmon fishery catches; a report of the Alaska Department of Fish and Game pursuant to National Oceanic and Atmospheric Administration Award No. NA87FPO408. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J00-01, Anchorage.

Crane, P. A., W. D. Templin, and L. W. Seeb. 1996. Genetic stock identification of Alaska Chinook salmon: a report of the Alaska Department of Fish and Game pursuant to National Oceanic and Atmospheric Administration Award No. NA46FD0356. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J99-17, Juneau.

Dann, T. H., C. Habicht, H. A. Hoyt, T. T. Baker, and F. W. West. 2011. Genetic stock composition of the commercial harvest of sockeye salmon in Bristol Bay, Alaska, 2009. Alaska Department of Fish and Game, Fishery Data Series No. 11-21, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FDS11-21.pdf

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. http://www.adfg.alaska.gov/FedAidPDFs/FMS09-06.pdf

Engel, L. J. 1972. Studies on the Russian River red salmon sport fishery. Alaska Department of Fish and Game. Federal Aid in Fish Restoration, Annual Performance Report, 1971-1972, Project F-9-4, 13 (G-II-G), Juneau http://www.adfg.alaska.gov/FedAidpdfs/FREDF-9-4(13)G-II-G.pdf

Flannery, B. G., T. D. Beacham, J. R. Cand, R. R. Holder, G. F. Maschmann, E. J. Kretschmer, and J. K. Wenburg. 2010. Mixed-stock analysis of Yukon River chum salmon: application and validation in a complex fishery. North American Journal of Fisheries Management 30(5): 1324-1338.

Gharrett, A. J., S. M. Shirley, and G. R. Tromble. 1987. Genetic relationships among populations of Alaskan Chinook salmon (*Oncorhynchus tshawytscha*). . Canadian Journal of Fisheries and Aquatic Sciences 44(4):765–774.

Goudet, J. 2006. Estimation and tests of hierarchical F-statistics. R package version 0.04-4. http://www.r-project.org, http://www.unil.ch/popgen/softwares/hierfstat.htm (Accessed December 6, 2011).

REFERENCES CITED (Continued)

Guthrie III, C. M., and R. L. Wilmot. 2004. Genetic structure of wild Chinook salmon populations of Southeast Alaska and northern British Columbia. Environmental Biology of Fishes 69:81–93.

Habicht, C., L. W. Seeb, K. W. Myers, E. V. Farley, and J. E. Seeb. 2010. Summer–Fall distribution of stocks of immature sockeye salmon in the Bering Sea as revealed by single-nucleotide polymorphisms. Transactions of the American Fisheries Society 139(4): 1171-1191.

Jombart, T. 2008. Adegenet: a *R* package for the multivariate analysis of genetic markers. Bioinformatics 24(11): 1403-1405. doi: 10.1093/bioinformatics/btn129

Lewis, P. O., and D. Zaykin. 2001. Genetic data analysis: computer program for the analysis of allelic data. Version 1.0. URL http://lewis.eeb.uconn.edu/lewishome/software.html (Accessed March 10, 2009; site currently discontinued).

Martin, K. E., C. A. Steele, J. P. Brunelli, and G. H. Thorgaard. 2010. Mitochondrial variation and biogeographic history of Chinook salmon. Transactions of the American Fisheries Society 139(3):792–802.

Miller, K. M., K. H. Kaukinen, T. D. Beacham, and R. E. Withler. 2001. Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. Genetica 111:237–257.

Narum, S. R., M. Banks, T. Beacham, M. Campbell, J. DeKoning, A. Elz, C. Guthrie, C. Kozfkay, K. Miller, P. Moran, R. Phillips, L. Seeb, C. Smith, K. Warheit, S. Young, and J. C. Garza. 2008. Differentiating salmon populations at broad and fine geographic scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology 17(15):3464-3477.

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.

Oslund, S., and S. Ivey. 2010. Recreational fisheries of Northern Cook Inlet, 2009-2010: Report to the Alaska Board of Fisheries, February 2011. Alaska Department of Fish and Game, Fishery Management Report No. 10-50, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FMR10-50.pdf

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in *R* language. Bioinformatics 20(2): 289–290.

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

R Development Core Team. 2011. *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/ (Accessed October 6, 2011).

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.

Ripley, B. 2010. RODBC: ODBC Database Access. R package version 1.3-2. http://CRAN.R-project.org/package=RODBC (Accessed December 15, 2011).

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., P. A. Crane, C. M. Kondzela, R. L. Wilmot, S. Urawa, N. V. Varnavskaya, and J. E. Seeb. 2004. Migration of Pacific Rim chum salmon on the high seas: insights from genetic data. Environmental Biology of Fishes 69: 21–36.

Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R. Z. Davis, L. K. Brannian, and J. E. Seeb. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the Exxon Valdez oil spill. Transactions of the American Fisheries Society 129(6):1223-1249.

Seeb, L. W., J. E. Seeb, C. Habicht, Edward V. Farley Jr., and F. M. Utter. 2011. Single-nucleotide polymorphic genotypes reveal patterns of early juvenile migration of sockeye salmon in the Eastern Bering Sea. Transactions of the American Fisheries Society 140:734-748.

REFERENCES CITED (Continued)

Shields, P. 2010. Upper Cook Inlet commercial fisheries annual management report, 2010. Alaska Department of Fish and Game, Fishery Management Report No. 10-54, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/FMR10-54.pdf

Smith, C. T., A. Antonovich, W. D. Templin, C. M. Elfstrom, S. R. Narum, and L. W. Seeb. 2007. Impacts of marker class bias relative to locus-specific variability of population inferences in Chinook salmon: a comparison of SNPs to STRs and allozymes. Transactions of the American Fisheries Society 136:1674-1687.

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

Smith, C. T., W. D. Templin, J. E. Seeb, and L. W. Seeb. 2005b. Single nucleotide polymorphisms provide rapid and accurate estimates of the proportions of U.S. and Canadian Chinook salmon caught in Yukon River fisheries. North American Journal of Fisheries Management 25:944-953.

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

Szarzi, N. J., C. M. Kerkvliet, B. J. Failor, and M. D. Booz. 2010. Recreational fisheries in the Lower Cook Inlet Management Area, 2008-2010, with updates for 2007. Alaska Department of Fish and Game, Fishery Management Report No. 10-38, Anchorage. http://www.adfg.alaska.gov/FedAidpdfs/Fmr10-38.pdf

Teel, D. J., P. A. Crane, C. M. G. III, A. R. Marshall, D. M. V. Doornik, W. D. Templin, N. V. Varnavskaya, and L. W. Seeb. 1999. Comprehensive allozyme database discriminates chinook salmon around the Pacific Rim. (NPAFC document 440). Alaska Department of Fish and Game, Division of Commercial Fisheries, 333 Raspberry Road, Anchorage, Alaska USA 99518.

Teel, D. J., G. B. Milner, G. A. Winans, and W. S. Grant. 2000. Genetic population structure and origin of life history types in Chinook salmon in British Columbia, Canada. Transactions of the American Fisheries Society 129(1):194–209.

Templin, W. D., J. E. Seeb, J. R. Jasper, A. W. Barclay, and L. W. Seeb. 2011. Genetic differentiation of Alaska Chinook salmon: the missing link for migratory studies. Molecular Ecology Resources 11(S1): 226-246.

Templin, W. D., R. L. Wilmot, C. M. G. III, and L. W. Seeb. 2005. United States and Canadian Chinook salmon populations in the Yukon River can be segregated based on genetic characteristics. Alaska Fishery Research Bulletin 11(1):44-60. http://www.adfg.alaska.gov/FedAidpdfs/AFRB.11.1.044-060.pdf

Utter, F., G. Milner, G. Stahl, and D. Teel. 1989. Genetic population structure of Chinook salmon, *Oncorhynchus tschawytscha*, in the Pacific Northwest. Fishery Bulletin 87(2) 239–264.

Waples, R. S. 1990a. Conservation genetics of Pacific salmon. III. Estimating effective population size. Journal of Heredity 81(4):277-289.

Waples, R. S. 1990b. 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.

Weir, B. 1996. Genetic data analysis (2nd edition). Sinauer Associates, Inc, Sunderland, MA.

TABLES AND FIGURES

Table 1.–Tissue collections of Chinook salmon throughout Upper Cook Inlet, including the year sampled, number of samples collected (N), the number of individuals analyzed from each collection included in the baseline, and the assigned group for tests of homogeneity. Map numbers correspond to sampling sites on Figure 1; unique population numbers represent all the analyzed collections that contribute to a single population.

Map No.	Pop. No.	Group	Location	Year Collected	N	Analyzed	Source[a]
1	---	West	Crescent River	2010	3	---	a
2	1		Straight Creek	2010	105	95	b
3	2		Chuitna River	2008	20	20	a
3	2			2009	122	122	a
4	3		Coal Creek	2009	42	42	a
4	3			2010	35	35	a
4	---			2011	43	---	a
5	---		Theodore River	2010	34	---	a
5	---			2011	55	---	b
6	---		Lewis River	2011	48	---	b
7	---	Susitna	East Fork Chulitna River	2009	5	---	c
7	---			2010	2	---	c
7	---			2011	6	---	c
8	4		Middle Fork Chulitna River	2009	72	72	a
8	4			2010	97	97	a
9	---		Portage Creek	2009	15	---	a
9	---			2010	10	---	a
9	---			2011	116	---	a
10	5		Stephan Lake weir	2008	19	19	a
11	5		Prairie Creek	1995	52	52	c
11	5			2008	98	98	a
12	6		Chunilna Creek	2009	50	50	a
13	7		Montana Creek	2008	33	33	a
13	7			2009	155	155	a
13	7			2010	30	30	a
14	8		Deception Creek	2009	122	100	c
15	8		Willow Creek	2005	74	74	c
16	9		Moose Creek	1995	51	51	c
17	9		Deshka River weir	2005	200	200	c
18	---		Sucker Creek	2011	91	---	a
19	---		Wolverine Creek	2011	1	---	a
20	10	Yentna	Talachulitna River	1995	58	58	c
20	10			2008	74	72	a
20	10			2010	48	48	a

-continued-

Table 1.–Part 2 of 3.

Map No.	Pop. No.	Group	Location	Year Collected	N	Analyzed	Source[a]
21	---	Yentna	Lake Creek	2008	1	---	a
22	11		Sunflower Creek	2009	53	53	a
22	---			2011	74	---	a
23	---		Peters Creek	2009	27	---	a
23	---			2010	6	---	a
23	---			2011	37	---	a
24	12	Knik	Little Susitna River	2009	3	3	a
24	12			2010	122	122	a
25	13		Moose Creek	1995	20	20	c
25	13			2008	33	33	c
25	13			2009	22	22	c
26	---		South Fork Eagle River	2009	1	---	c
26	---			2011	4	---	b
27	---		Meadow Creek	2009	6	---	c
28	14		Ship Creek	2009	311	311	c
29	---	Turnagain	Campbell Creek	2010	3	---	b
29	---			2011	30	---	a
30	---		Rabbit Creek	2011	8	---	a
31	---		Bird Creek	2009	2	---	c
31	---			2011	35	---	a
32	---		Carmen River	2003	5	---	c
32	---			2011	19	---	a
33	---		Granite Creek	2011	1	---	a
34	---	Turnagain	Resurrection Creek	2010	24	---	b
34	---			2011	61	---	a
35	15		Chickaloon River	2008	2	2	c
35	---			2009	1	---	c
35	15			2010	66	65	b
35	---			2011	58	---	b
36	---	Kenai	Grant Creek	2011	23	---	b
37	16		Quartz Creek	2006	35	34	c
37	16			2008	34	34	c
37	16			2009	41	41	c
38	16		Dave's Creek	2007	8	8	c
38	16			2008	5	5	c

-continued-

Table 1.–Part 3 of 3.

Map No.	Pop. No.	Group	Location	Year Collected	N	Analyzed	Source[a]
39	17	Kenai	Crescent Creek	2006	165	165	c
40	18		Juneau Creek	2005	32	32	c
40	18			2006	100	91	c
40	18			2007	24	24	c
41	19		Russian River	2005	24	24	c
41	19			2006	16	16	c
41	19			2007	84	83	c
41	19			2008	91	91	c
42	20		Upper Kenai River mainstem	2009	200	200	c
43	21		Benjamin Creek	2005	56	56	c
43	21			2006	150	150	c
44	22		Killey River	2005	68	68	c
44	22			2006	190	190	c
45	23		Funny River	2005	37	37	c
45	23			2006	183	183	c
46	24		Middle Kenai River mainstem	2003	80	80	c
46	24			2004	39	39	c
46	24			2006	183	183	c
47	25		Slikok Creek	2004	48	48	c
47	25			2005	100	95	c
47	25			2008	58	57	c
48	---		Lower Kenai River mainstem	2011	90	---	c
49	26	Kasilof	Lower Kasilof River mainstem	2005	144	49	c
49	26		Middle Kasilof River mainstem	2005	273	273	c
50	27		Crooked Creek	1992	95	95	c
50	27			2005	212	212	c
51	28	SKenaiPen	Ninilchik River weir	2006	190	162	c
51	---			2009	93	---	c
51	---			2010	50	---	c
52	29		Deep Creek	2009	100	100	c
52	---			2010	100	---	c
53	---		Stariski Creek	2011	56	---	b
54	30		Anchor River weir	2006	200	200	c
54	---			2009	10	---	c
54	---			2010	50	---	c

[a] Collection sources: a) AKSSF project 45864, b) AKSSF project 44517, c) other projects.

Table 2.–Source, observed heterozygosity (H_o), F_{IS}, and F_{ST} for 46 single nucleotide polymorphisms (SNPs) used to analyze the population genetic structure of Upper Cook Inlet Chinook salmon. These summary statistics are based upon the 30 populations within Upper Cook Inlet.

Assay Name	Source [a]	H_o	F_{IS}	F_{ST}
Ots_AsnRS-60	b	0.382	0.010	0.042
Ots_C3N3 [b,c]	e	---	---	---
Ots_E2-275	b	0.302	0.003	0.027
Ots_E9BAC [c]	c	0.000	---	---
Ots_ETIF1A	c	0.494	-0.013	0.023
Ots_FARSLA-220	d	0.341	-0.009	0.026
Ots_FGF6A	a	0.416	-0.008	0.041
Ots_GH2	e	0.282	-0.005	0.023
Ots_GPDH-338	b	0.078	-0.019	0.032
Ots_GPH-318	d	0.127	0.003	0.043
Ots_GST-207	d	0.061	0.015	0.035
Ots_GST-375 [c]	d	0.000	---	---
Ots_GTH2B-550	a	0.460	-0.014	0.016
Ots_hnRNPL-533	d	0.318	-0.009	0.024
Ots_HSP90B-100	d	0.199	0.030	0.025
Ots_HSP90B-385	d	0.021	-0.008	0.055
Ots_IGF-I.1-76	b	0.465	-0.019	0.096
Ots_Ikaros-250	b	0.123	0.000	0.028
Ots_il-1racp-166 [d]	b	0.451	-0.052	0.037
Ots_ins-115	b	0.035	-0.003	0.011
Ots_LEI-292	d	0.053	0.042	0.036
Ots_LWSop-638	b	0.063	-0.012	0.018
Ots_MHC1	e	0.424	0.015	0.033
Ots_MHC2	e	0.032	-0.007	0.012
Ots_NOD1	a	0.457	0.005	0.069
Ots_P450	e	0.309	-0.006	0.022
Ots_P53	e	0.439	0.010	0.020
Ots_PGK-54	a	0.025	0.011	0.019
Ots_Prl2	e	0.473	0.008	0.039
Ots_PSMB1-197 [e]	d	0.000	-0.002	0.002
Ots_RAG3	a	0.304	-0.003	0.051
Ots_RFC2-558 [c]	b	0.000	---	---

-continued-

Table 2.–Part 2 of 2.

Assay Name	Source [a]	H_o	F_{IS}	F_{ST}
Ots_S7-1	a	0.188	0.008	0.027
Ots_SClkF2R2-135	b	0.393	0.003	0.072
Ots_SERPC1-209 [d]	d	0.189	0.079	0.049
Ots_SL	e	0.466	-0.029	0.076
Ots_SWS1op-182	b	0.453	-0.027	0.011
Ots_TAPBP	c	0.296	-0.003	0.033
Ots_Tnsf	e	0.185	-0.005	0.021
Ots_u202-161	b	0.074	-0.012	0.036
Ots_u211-85	b	0.244	-0.029	0.035
Ots_U212-158	b	0.037	0.005	0.040
Ots_u6-75	b	0.113	-0.008	0.038
Ots_unk526	a	0.233	0.005	0.067
Ots_ZNF330-181 [e]	b	0.000	0.000	0.000
Ots_Zp3b-215	b	0.059	0.034	0.075
Average/Overall		0.252	-0.004	0.040

[a] Marker sources: a) Northwest Fisheries Science Center-NOAA (Unpublished); b) Smith et al. 2005a; c) Washington State University Vancouver (Unpublished); d) Smith et al. 2007; e) Smith et al. 2005b.

[b] Mitochondrial SNP.

[c] These were removed because they were invariant.

[d] These were removed because they did not conform to Hardy-Weinberg expectations.

[e] These were removed because they were invariant in all but 2 individuals.

Table 3.–Hierarchical log-likelihood ratio (G) analysis of population structure based on allele frequencies at 38 SNP loci. The probability of the statistic (P), assuming the null hypothesis is true, is provided for inferring significance. The scaled G statistic (G′) is provided for comparing levels of heterogeneity.

Population groupings	DF	G	G′	P
Upper Cook Inlet	1102	14168	12.86	<0.001
Between Regions	38	3028	79.68	<0.001
Within Regions	1064	11141	10.47	<0.001
Northern	532	3387	6.37	<0.001
Among Groups	152	1009	6.64	<0.001
Within Groups	380	2378	6.26	<0.001
West	76	167	2.20	<0.001
Susitna	190	1871	9.85	<0.001
Yentna	38	103	2.72	<0.001
Knik	76	236	3.11	<0.001
Turnagain	---	---	---	---
Southern	532	7754	14.57	<0.001
Among Groups	76	2014	26.50	<0.001
Within Groups	456	5740	12.59	<0.001
Kenai	342	4761	13.92	<0.001
Kasilof	38	848	22.32	<0.001
SKenaiPen	76	130	1.71	<0.001

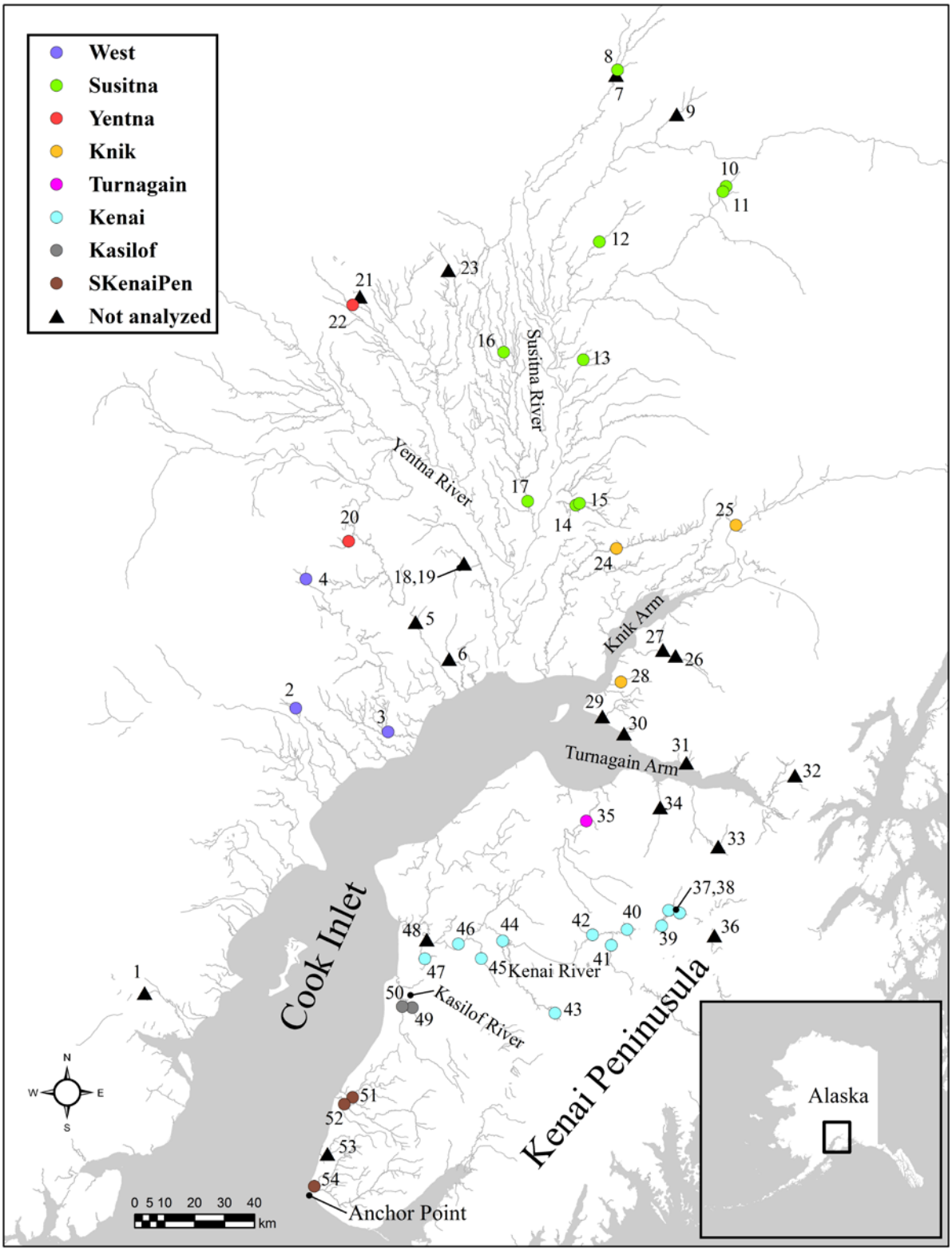

Figure 1.–Sampling locations for Chinook salmon originating for Upper Cook Inlet, Alaska, 1992–2011. Numbers correspond to map numbers on Table 1. Locations included in the baseline analysis (circles) are colored by group and black triangles indicate locations not included in the analysis.

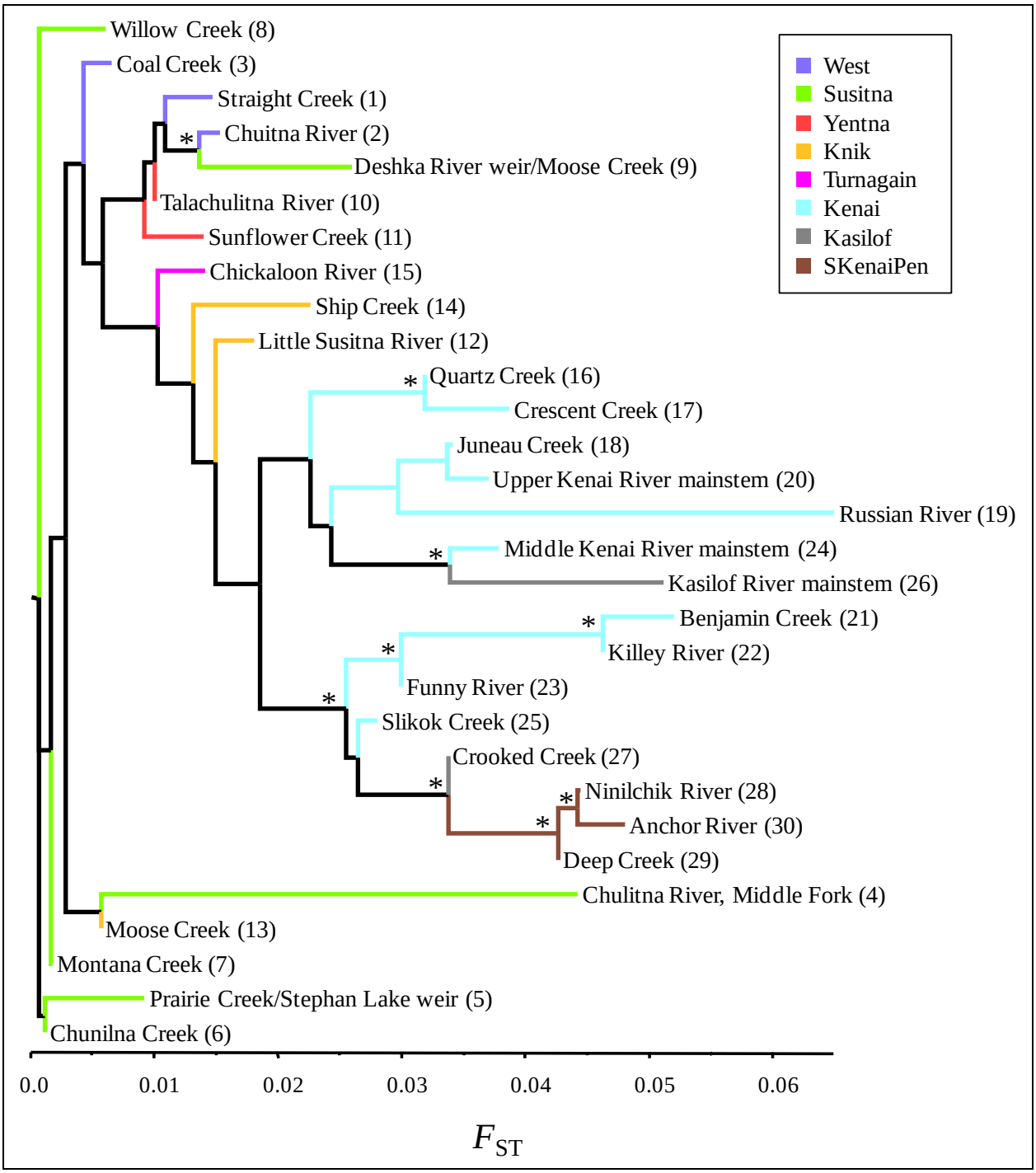

Figure 2.–Consensus NJ tree based on F_{ST} between Chinook salmon populations sampled from spawning areas in drainages of Upper Cook Inlet, Alaska (see Table 1 for collection details).

Note: Colors denote groups as in Figures 1, 3, and 4. Numbers in parentheses correspond to unique population numbers on Table 1. Bootstrap consensus nodes occurring in >50% of trees are marked with an asterisk.

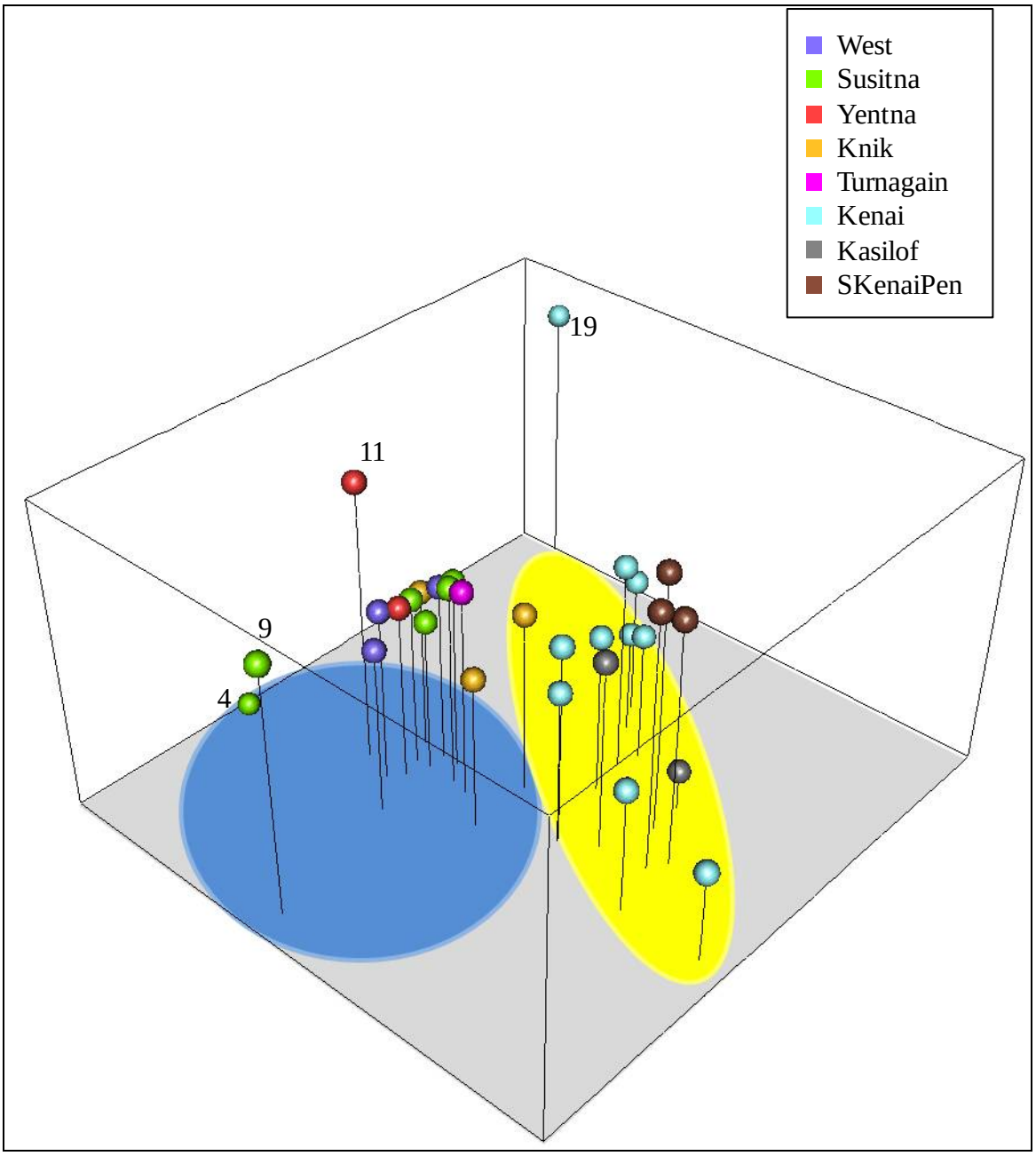

Figure 3.–Multidimensional scaling (MDS) plot based on F_{ST} between Chinook salmon populations sampled from spawning areas in drainages of Upper Cook Inlet, Alaska (see Table 1 for collection details). Group colors correspond to those in Figures 1, 2, and 4 and numbers on outlier populations correspond to unique population numbers on Table 1. Northern and Southern regions are indicated with blue and yellow ovals, respectively.

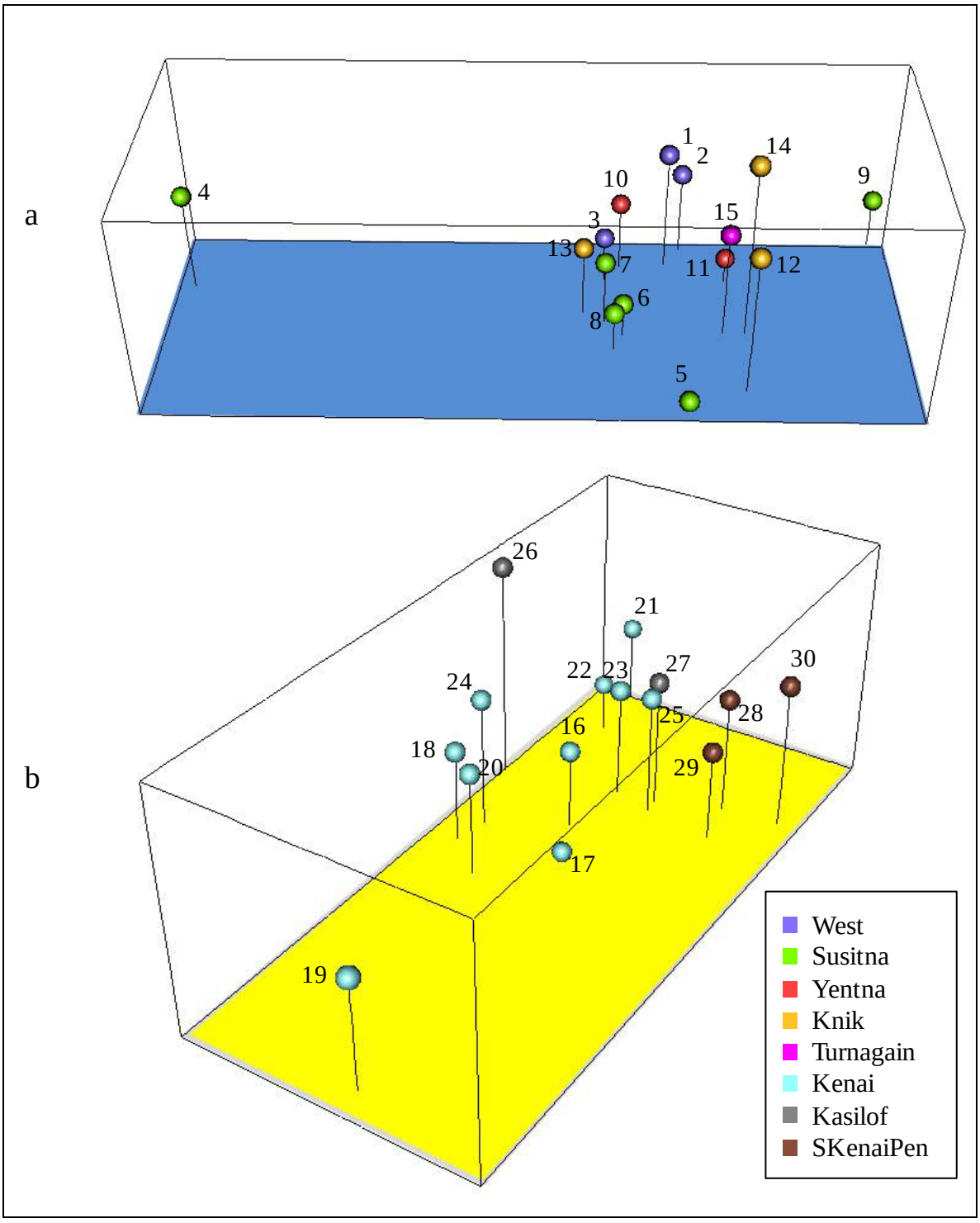

Figure 4.–Multidimensional scaling (MDS) plots based on F_{ST} between Chinook salmon from the a) Northern region (blue) and b) Southern region (yellow) of Cook Inlet, Alaska. Group colors correspond to those in Figures 1–3 and numbers correspond to unique population numbers on Table 1.