U.S. Fish and Wildlife Service

Run Timing, Migratory Patterns, and Harvest Information of Chinook Salmon Stocks within the Yukon River

Alaska Fisheries Technical Report Number 92

Conservation Genetics Laboratory
October 2006

The Alaska Region Fisheries Program of the U.S. Fish and Wildlife Service conducts fisheries monitoring and population assessment studies throughout many areas of Alaska. Dedicated professional staff located in Anchorage, Juneau, Fairbanks, Kenai, and King Salmon Fish and Wildlife Offices and the Anchorage Conservation Genetics Laboratory serve as the core of the Program's fisheries management study efforts. Administrative and technical support is provided by staff in the Anchorage Regional Office. Our program works closely with the Alaska Department of Fish and Game and other partners to conserve and restore Alaska's fish populations and aquatic habitats. Additional information about the Fisheries Program and work conducted by our field offices can be obtained at:

http://alaska.fws.gov/fisheries/index.htm

The Alaska Region Fisheries Program reports its study findings through two regional publication series. The **Alaska Fisheries Data Series** was established to provide timely dissemination of data to local managers and for inclusion in agency databases. The **Alaska Fisheries Technical Reports** publishes scientific findings from single and multi-year studies that have undergone more extensive peer review and statistical testing. Additionally, some study results are published in a variety of professional fisheries journals.

Alaska Fisheries Technical Report Number 92. October 2006
U.S. Fish and Wildlife Service

Run Timing, Migratory Patterns, and Harvest Information of Chinook Salmon Stocks within the Yukon River

Blair Flannery[1], Terry Beacham[2], Michael Wetklo[2], Christian Smith[3], Bill Templin[3], Anton Antonovich[3], Lisa Seeb[3], Steve Miller[1], Ora Schlei[1], and John Wenburg[1]

Abstract

Yukon River Chinook salmon (*Oncorhynchus tshawytscha*) stocks were assayed for diversity at 34 microsatellite loci to examine stock structure and evaluate mixed-stock analysis (MSA) potential. Laboratories from three agencies, U.S. Fish and Wildlife Service, Department of Fisheries and Oceans Canada, and the Alaska Department of Fish and Game, collaborated in the survey of microsatellite diversity. Each laboratory surveyed a subset of the microsatellite loci. Allele frequencies were subsequently pooled to form a single, joint baseline. Yukon River Chinook salmon stocks are geographically structured, with moderate divergence (F_{ST} = 0.044) within and among geographic regions. Using the 10 most powerful loci, accuracy to country of origin is > 99%, ranges from 94%–99% for eight regional groupings, and 86%–99% for individual stocks (for the 19 individual stocks, 16 > 90%). The standardized Pacific Salmon Commission Chinook Technical Committee 13-locus baseline produced comparable results. Microsatellite analysis can be used to accurately and precisely allocate Chinook salmon in mixtures to region and, in many cases, drainage and tributary of origin, providing managers with a powerful tool for assessing and regulating fisheries.

Introduction

Chinook salmon (*Oncorhynchus tshawytscha)* spawn throughout the Yukon River drainage and support important subsistence and commercial fisheries. In the Yukon River, fisheries that harvest Chinook salmon are managed to meet established escapement and transboundary passage goals. Providing for fisheries and escapement is made difficult by the compressed return time and mixed-stock nature of the run ("stock" in this report refers to a particular river or stream where a collection was made). However, mixed-stock analysis (MSA), a method of estimating the proportion of individuals from different source stocks contributing to a mixture, can help management meet these obligations by supplying information to prevent excessive exploitation and protect genetic diversity, both essential for sustained salmon productivity (NRC 1996). Furthermore, the additional stock return data that MSA provides may assist

Authors: [1] USFWS, Conservation Genetics Laboratory, 1011 E. Tudor Rd., Anchorage, Alaska 99503, U.S.A.; [2] DFO Canada, Pacific Biological Station, Molecular Genetics Laboratory, Nanaimo, B.C. V9T 6N7, Canada; [3] ADFG, Division of Commercial Fisheries, Gene Conservation Laboratory, 333 Raspberry Rd., Anchorage, Alaska 99518, U.S.A.

run reconstruction and allow a more rigorous assessment of escapement goals (Bergstrom et al. 2001).

Yukon River Chinook salmon MSA has been conducted using scale pattern analysis (SPA), radio telemetry, and allozyme markers, but all have drawbacks in comparison to DNA markers. Apportioning harvests with SPA is limited to three geographic regions: lower (U.S.), middle (U.S.), and upper (Canada) Yukon River. Average classification accuracies between 65% and 80% for such large-scale reporting groups indicate that SPA is not a very powerful tool, but it is useful for distinguishing between U.S. and Canada Chinook salmon in a cost-effective, if not timely, manner (JTC 1997). Scale pattern analysis is environmentally influenced, requires new baselines annually, and is only applicable as a post-season tool. The level of effort, low accuracy of estimates, and limited resolution of this technique often restricts its use. Mark-recapture and radio telemetry resolve contributions of stocks to mixtures (Eiler et al. 2004), but this is impractical for continual application on large systems of wild fish because of the logistics, expense, difficulty in tagging and tracking large numbers of individuals, and low tag return (Ihssen et al. 1981). Allozyme diversity reveals a geographically hierarchical stock structure with strong divergence among six regions (Gharrett et al. 1987; Beacham et al. 1989; Wilmot et al. 1992; Templin et al. 2005). Mean accuracies of MSA simulations to those six regions are 94%–97% and to country of origin are 98%–99%. However, the allozyme baseline does not include any U.S. stocks above the Tanana River, so simulation accuracies may be overly optimistic. Furthermore, allozyme samples are logistically difficult to collect and store properly, and there are no available loci left to assay, which precludes increasing power to distinguish among stocks through additional loci. Allozyme MSA estimates of commercial and test fishery catches from 1987–1991 are concordant with SPA estimates for that period (Wilmot et al. 1992). The Joint Technical Committee (JTC), therefore, decided to discontinue allozyme MSA for Chinook salmon and focus research on chum salmon as no other stock identification tool was available for that species (JTC 1997).

Following severe run shortages from 1998–2002, with the lowest return in state history occurring in 2000, focus returned to Chinook salmon. The question of whether finer level stock structure exists than what SPA and allozymes reveal prompted research into using DNA-based genetic markers for MSA. It is hoped that greater resolution will improve estimates of run timing, migratory patterns, and harvest composition and allow for better assessment of factors influencing stock productivity.

Microsatellites are a class of highly variable DNA-based markers with a wealth of available loci. They have strong potential for accurate and precise MSA. Diversity at microsatellite loci discriminates among salmonid stocks within watersheds (Small et al. 1998; Beacham and Wood 1999; Beacham et al. 2001) and is a powerful tool for stock discrimination of Chinook salmon (Banks et al. 2000; Beacham et al. 2003). Here we report on a three-year project funded by the Office of Subsistence Management Fisheries Information Service to develop a microsatellite baseline for Yukon River Chinook salmon as a cooperative effort among the U.S. Fish and Wildlife Service (USFWS), the Alaska Department of Fish and Game (ADFG), and the Department of Fisheries and Oceans Canada (DFOC). The main objectives are to collect, exchange, and analyze 200 samples from each of 10 U.S. and 10 Canadian stocks at a

minimum of 30 microsatellite loci to assess stock structure and the potential utility of using a microsatellite baseline for MSA of Yukon River Chinook salmon.

Methods

Collection of Samples and Laboratory Analysis

Baseline samples from adults returning to spawn were collected from 19 Yukon River Chinook salmon stocks between 1987 and 2003 (Table 1, Figure 1). The vast size, habitat diversity, and remoteness of the Yukon River drainage along with the limited knowledge of Chinook salmon ecology in the system precluded the collection of the target number of samples in many systems (see Appendix 1). Mixed fishery samples were collected from Pilot Station test fisheries in 2002–2003. All samples, either tissue or DNA, were distributed among the labs. Genomic DNA was extracted from either liver, scales, operculum punches, or fin clips. Extractions were conducted with either a chelex resin protocol outlined by Small et al. (1998) for the DFOC extractions or a QIAGEN 96-well Dneasy® procedure for the ADFG and USFWS extractions. The USFWS DNA was quantified using a 96-well Packard FluoroCount® Microplate Fluorometer and diluted to 30ng/µl for use in polymerase chain reaction (PCR) DNA amplifications.

For the survey of baseline stocks, the DFOC surveyed diversity at 13 microsatellite loci: *Ots100*, *Ots101*, *Ots102*, *Ots104*, *Ots107* (Nelson and Beacham 1999), *Ssa197* (O'Reilly et al. 1996), *Ogo2*, *Ogo4* (Olsen et al. 1998), *Oke4* (Buchholz et al. 2001), *Omy325* (O'Connell et al. 1997), *Oki100* (K. M. Miller, unpublished data), and *Ots2*, *Ots9* (Banks et al. 1999). The ADFG surveyed diversity at 10 microsatellite loci: *Ots100*, *Ots107* (Nelson and Beacham 1999), *Ots1*, *Ots2* (Banks et al. 1999), *Oke4* (Buccholz et al. 2001), *One7*, *One9* (Scribner et al. 1996), *One102* (Olsen et al. 2000), *Ots212* (Greig et al. 2003), and *µ73* (Estoup et al. 1993). The USFWS surveyed diversity at 11 microsatellite loci: *Oke2*, *Oke4* (Buccholz et al. 2001), *Ots3.1* (Banks et al. 1999), *Oki10*, *Oki11* (Smith et al. 1998), and *Ots311*, *OtsG474*, *OtsG68*, *OtsG432*, *OtsG3*, *OtsG253b* (Williamson et al. 2002).

In general, PCR DNA amplifications were conducted using MJResearch thermalcylers (BioRad, Hercules, CA) in 10µl volumes consisting of 0.06 units of *Taq* polymerase, 1µl of 30ng DNA, 1.5–2.5mM $MgCl_2$, 1mM 10x buffer, 0.8mM dNTPs, 0.006–0.065µM of labeled forward primer (depending on the locus), 0.4µM unlabeled forward primer, 0.4µM unlabeled reverse primer, deionized H_2O, and 1M betaine (majority of loci). The thermal cycling profile involved one cycle of 2 minutes at 92°C, followed by 30 cycles of 15 seconds at 92°C, 15 seconds at 52–60°C (depending on the locus), and 30 seconds at 72°C, with a final extension for 10 minutes at 72°C. Specific PCR conditions for a particular locus could vary from this general outline. For the DFOC, PCR fragments were size fractionated in denaturing polyacrylamide gels using an ABI 377 automated DNA sequencer, and genotypes were scored by Genotyper 2.5 software (Applied Biosystems, Foster City, CA) using an internal lane sizing standard. For the ADFG, microsatellites were size fractionated in an ABI 3730 capillary DNA sequencer, and genotypes were scored by GeneMapper software 3.0 (Applied Biosystems, Foster City, CA) using an internal lane sizing standard. For the USFWS, microsatellites were size fractionated in denaturing polyacrylamide gels using Li-Cor IR^2 scanners, and genotypes

Table 1. Stock, sample collection years, number of adult fish sampled per year, and total number of fish sampled for 10 Canadian and 9 U.S. stocks of Yukon River Chinook salmon.

Stock	Year Sampled	*N*	Total *N*
Canadian Stocks			
Whitehorse	1985, 1987, 1997	39, 89, 114	242
Takhini	1997, 2002, 2003	63, 67, 38	168
Big Salmon	1987, 1997	76, 40	116
Tatchun	1987, 1996, 1997, 2002, 2003	27, 200, 58, 36, 48	369
Little Salmon	1987, 1997	20, 80	100
Blind	1997, 2003	1, 138	139
Mayo	1992, 1997, 2003	135, 32, 38	205
Stewart	1997	99	99
Pelly	1996, 1997	39, 113	152
Chandindu	1998, 2001, 2003	123, 158, 85	366
U.S. stocks			
Chandalar	2002, 2003	4, 113	117
Beaver	1997	96	96
Chena	2001	200	200
Salcha	2003	55	55
Tozitna	2002, 2003	200, 250	450
Henshaw	2001	150	150
Gisasa	2001	368	368
Anvik	2002, 2003	75, 38	113
Andreafsky	2002, 2003	28, 209	237

were scored by Saga™ GT 3.1 software (Li-Cor, Lincoln, NE) using positive controls, consisting of known genotypes, and sizing standards loaded throughout the gel.

Standardization

Shortly after the initiation of this project, a large consortium of U.S. and Canadian genetic laboratories was funded by the Pacific Salmon Commission's Chinook Technical Committee (CTC) to develop a microsatellite baseline for Chinook salmon from Alaska to California. This process yielded a final set of 13 microsatellite loci that have recently been standardized for use in MSA. In light of this process, ADFG proposed a resolution, which was accepted, at a Yukon River Panel meeting that adopting these loci for Yukon River Chinook salmon MSA was the best course of action. Eight of the 34 loci analyzed in this project are also included in the 13 CTC loci. The JTC funded DFOC and ADFG to analyze the CTC loci on the Canadian and U.S. stocks collected in this study. To ensure consistent scoring across individuals, all individuals were reanalyzed for the full suite of 13 loci.

Data Analysis

A joint baseline was built by integrating allele frequencies for loci that were successfully surveyed in all 19 stocks by the three labs. All analyses documented in this report were based on

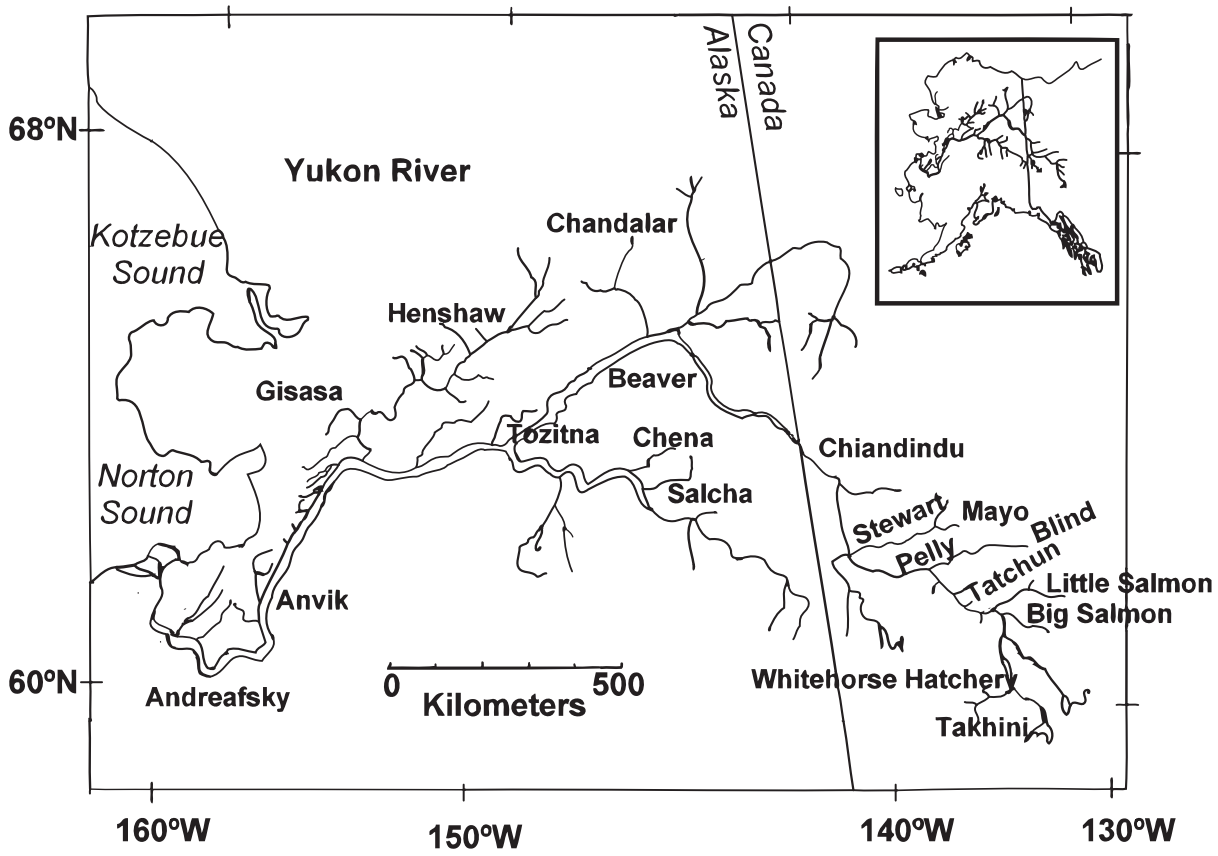

Figure 1. Sampling locations for the 19 stocks of Chinook salmon in the Yukon River drainage.

one of four baselines (the joint baseline or one of the three lab-specific baselines). In analyses where multiple tests of the same hypothesis were performed, significance levels were adjusted to prevent type I error by dividing the alpha by the number of tests (Cooper 1968). Regional groups for hierarchical analyses such as MSA, gene diversity analysis, and log-likelihood ratio tests were formed initially from neighbor-joining analysis results.

For each stock, allele frequencies at each locus were estimated (available upon request). Each locus within each stock was tested for departures from Hardy-Weinberg equilibrium (HWE) using a randomization test implemented in the computer program FSTAT 2.9.3 (Goudet 1995), and the results assessed following the binomial probability expectation method of Apostal et al. (1996). Estimates of polymorphism at the 95% criterion, allele number, heterozygosity, and F_{ST} (Weir and Cockerham 1984) were calculated for each locus over all stocks with FSTAT 2.9.3.

Stock Structure

Chord distances (Cavalli-Sforza and Edwards 1967; Wright 1978) calculated from allele frequencies were used to estimate genetic distances among the stocks. A neighbor-joining dendrogram (Saitou and Nei 1987) was constructed from the chord distances using MEGA version 2.1 (Kumar et al. 2001). Linear regression and lowess smoothing of the genetic and geographic distance matrices, calculated between all stock pairs, were used to assess the fit

of the data to the isolation by distance model (Wright 1943). The matrices correlation coefficient was assigned significance by randomizing one matrix 10,000 times with the program Mantel 1.0 (Jeffrey Bromaghin, U.S. Fish and Wildlife Service, 1011 East Tudor Road, Anchorage, Alaska, 99503).

Stocks within and among regions were tested for genetic divergence through log-likelihood ratio analysis of allelic frequencies (*G*-test, Sokal and Rohlf 1995). To preserve the *G*-test's approximation of the χ^2 probability distribution, alleles with expected overall counts of less than four were pooled, until counts equaled or exceeded four, with alleles of similar electrophoretic mobility (Kondzela et al. 1994; Sokal and Rohlf 1995). The ratio of the normalized *G*-test statistics (*G*/df) for within and among regions was compared to evaluate the distribution of genetic diversity (Smouse and Ward 1978).

An unbiased gene diversity analysis, with corresponding measures of gene differentiation (G_{ST}-statistics), was used to partition total gene diversity (H_T) into hierarchical components of within stocks (H_S), among stocks within regions (D_{SR}), and among regions (D_{RT}) (Chakraborty and Leimar 1987; Nei and Chesser 1983). Heterogeneity of allelic frequencies, measured by the *G*-test, indicated whether the G_{ST}-statistics differed significantly from zero (Chakraborty and Leimar 1987). Gene flow was estimated from the formula: $N_e m = ((1\text{-}G_{ST})\text{-}1)/4)/(g/(g\text{-}1))^2$ (Zhivotovsky et al. 1994).

Concordance of Data Sets

The three lab-specific baselines were tested for concordance to determine if the same genetic structure was observed and whether pooling the data into one joint baseline was justified (Scribner et al. 1998). A chi-square test was used to determine if G_{ST}-statistics were significantly different ($P < 0.05$) (Allendorf and Seeb 2000). Additionally, the chord distance matrices were tested for correlation (Smouse et al 1986). Both the ADFG and DFOC lab-specific chord distance matrices were regressed on the USFWS chord distance matrix. The resulting residual deviations ($D_{i.k}$ and $D_{j.k}$) from the two regressions were used to compute the partial correlation coefficient ($r_{ij.k}$) of the ADFG and DFOC matrices, given the USFWS matrix. The significance of $r_{ij.k}$ was determined by 10,000 randomizations using Mantel 1.0.

Mixed-Stock Analysis Simulations

An analysis was initially conducted to determine whether allele binning (Bromaghin and Crane 2005) or Bayesian allele frequency estimation (Rannala and Mountain 1997) as implemented in SPAM 3.7 (Debevec et al. 2000) to control sampling zero problems provided the best results in 100% individual stock MSA simulations. A sampling zero is the occurrence of fish in the mixed sample from a specific stock having an allele not observed in the baseline samples from that stock.

The evaluation of the baseline for stock composition estimation included analysis of simulated fishery samples. All simulations were performed using SPAM 3.7 (Debevec et al. 2000). In these simulations, mixture samples with 100% contribution from a single stock (or region) of interest were generated from the baseline allele frequencies. Then, the stock composition for each mixture was estimated by the maximum likelihood algorithm. Parametric bootstrap resampling of both the baseline and mixture was performed to derive mean allocation esti-

mates and to evaluate precision. Size of simulated mixtures was set to 400 fish and the number of bootstrap iterations to 1000. Either the Bayesian correction (Rannala and Mountain 1997) to baseline allele frequencies or allele binning (Bromaghin and Crane 2005) was used in the analysis to control sampling zeros. All allele frequencies were assumed to be in Hardy-Weinberg equilibrium, and all loci were assumed to be independent.

To evaluate the ability of the baseline to estimate stock composition, individual loci and multilocus subsets were subjected to 100% MSA simulations for individual stocks and regions using equal proportions of the stocks from the region in the simulated mixtures. The following loci subsets were analyzed: 10 best, 10 worst, 10 random, and 13 from the standardized CTC baseline. The best and worst loci for MSA were determined by analyzing the data with the computer program WHICHLOCI (Banks et al. 2003). The program WHICHLOCI ranks loci based on their ability to correctly assign individuals to stock of origin and determines the combination of loci that will provide a specified assignment success rate. Although WHICHLOCI ranks loci based on individual assignment, which is different than MSA, preliminary findings from a study developing a method to select loci combinations for MSA indicate that results will be similar to WHICHLOCI (Jeffrey Bromaghin pers. comm.). To allow a direct comparison with previously published allozyme and single nucleotide polymorphism (SNP) data (Templin et al. 2005; Smith et al. 2005a), 100% simulations were also done on six regions used in those studies. Lastly, realistic regional simulations of mixtures that could be expected in a lower river fishery were performed using the best 10 loci.

Evaluation of Accuracy and Precision

Scatter plots were created to analyze the relationships between combinations of accuracy and precision against sample size and number of alleles. Specifically, the mean accuracies of the 100% individual stock simulations were compared to sample size. In addition, the mean accuracy and standard deviation for each locus from the 100% individual stock simulations were compared to the number of alleles observed at that locus. Using the DFOC baseline, accuracy and standard deviation versus number of alleles were compared for a stock with the highest and for a stock with the lowest accuracy in 100% MSA simulations by sequentially adding loci in the MSA simulation analysis. Lastly, using the ADFG baseline, histograms were created of the accuracy and standard deviation in 100% individual stock simulations of one highly polymorphic locus (*Ots100*) against all 10 loci (including *Ots100*).

Mixed-Stock Analysis of Pilot Station Test Fishery Samples

Analysis of test fishery samples collected at Pilot Station was performed using the ADFG baseline and SPAM 3.7. Maximum likelihood estimates of mixture composition were obtained using the observed baseline frequencies with either the Bayesian or allele binning correction. Variability of these estimates was quantified by the bootstrap method. During the bootstrap iterations, the fishery mixture was resampled non-parametrically whereas the baseline was resampled parametrically. Apportionments were made to country of origin for the purposes of the Yukon River Salmon Agreement and to the three regions for which comparable assignments were made using the allozyme baseline (lower U.S., middle U.S., and Canada). Stock composition estimates were compared to those generated using the allozyme and SNP baselines.

Results

Sample Collections

The Yukon River occupies an area of more than 330,000 square miles, comprises eight major river tributaries, and forms 20 eco-regions (Brabets et al. 2000). The diverse hydrology includes clear, rocky, fast moving tributaries that drain mountain basins and murky, slow, meandering tributaries that drain tundra. The combination of remoteness and different habitat types demanded flexibility and adaptability in designing and implementing sampling regimens. At the conclusion of the year 2003 field season, 19 Canada/U.S collections were considered adequate for use in the baseline. Additional samples were collected in the year 2004 but were not incorporated into this report due to time constraints. For details on U.S. collections, see Appendix 1.

Microsatellite Analysis

Data were successfully collected from 30 of the 34 microsatellite loci for all 19 stocks. Limited DNA quantities, variable DNA extraction chemistries, and the age of some samples precluded full data collection for some locus/stock combinations, most notably: *Otsg253b* for Tatchun, Pelly, Big Salmon, and Little Salmon; and *Otsg68*, *Otsg432*, and *Otsg474* for Tatchun.

For other stocks, there are more minor but often significant differences in the success rates among loci. The majority of the Canadian samples used in this study were from archived DNA, provided by DFOC, many of which were from DNA extractions up to six years old. The DFOC lab was able to initially run these samples soon after they were extracted and have optimized their reaction conditions to the chelex chemistry. However, the USFWS and ADFG labs had lower levels of success than DFOC when amplifying various loci for the Canadian samples because of the age of these samples and lab protocols that were optimized for different extraction methods. The ADFG lab first encountered difficulties with the Canadian samples and provided information on the more recalcitrant samples to the USFWS lab. The ADFG lab attempted to analyze some of these samples multiple times often with limited success but eventually managed to improve yield by various optimization strategies. For several of these difficult samples, the USFWS lab chose not to incur the added expense of attempting to PCR the samples multiple times. The USFWS lab had already incurred unanticipated costs of over $100,000 for the collections in this project and, as such, decided that the added expense of rerunning some of these recalcitrant samples outweighed the potential gain in information. It should be noted that DFOC did recently go back to these samples in order to run the CTC loci for the Canadian stocks and did have better success rates in general than the U.S. labs.

With the exception of *One9*, all loci were polymorphic at the 95% criterion. The number of observed alleles at each locus ranged from 2 to 61; lower heterozygosity was observed at those loci with fewer alleles (e.g. *One7*; Table 2). Expected heterozygosity varied among the loci and ranged from 0.12 to 0.94 (Table 2).

Per locus F_{ST} values ranged from 0.015 to 0.118 with a mean value of 0.044, indicative of moderate differentiation among stocks. Significant ($P < 0.05$) departures from HWE were

Table 2. Number of alleles, expected heterozygosity (H_E), observed heterozygosity (H_O), and F_{ST} (standard deviation in parentheses) among Yukon River Chinook salmon stocks for microsatellite loci.

Locus	Alleles	H_E	H_O	F_{ST}
DFOC				
Ots100	39	0.90	0.88	0.017 (0.002)
Ots101	44	0.90	0.91	0.027 (0.004)
Ots102	52	0.92	0.71	0.030 (0.005)
Ots104	34	0.92	0.90	0.019 (0.003)
Ots107	35	0.88	0.86	0.036 (0.004)
Ogo2	12	0.57	0.58	0.065 (0.014)
Ogo4	15	0.69	0.69	0.053 (0.011)
Oke4	6	0.65	0.62	0.104 (0.037)
Oki100	33	0.93	0.91	0.024 (0.003)
Omy325	15	0.73	0.72	0.038 (0.007)
Ots2	7	0.26	0.24	0.015 (0.002)
Ots9	6	0.45	0.45	0.048 (0.013)
Ssa197	36	0.92	0.91	0.032 (0.005)
ADFG				
Ots100	61	0.92	0.91	0.029 (0.004)
Ots107	33	0.88	0.84	0.034 (0.004)
Oke4	6	0.72	0.70	0.096 (0.034)
One7	2	0.18	0.18	0.077 (0.029)
One9	3	0.07	0.06	0.017 (0.005)
Ots1	9	0.64	0.63	0.068 (0.015)
Ots2	16	0.25	0.24	0.019 (0.005)
One102	3	0.45	0.41	0.054 (0.013)
Ots212	23	0.67	0.70	0.044 (0.011)
μ73	6	0.30	0.27	0.065 (0.014)
USFWS				
Oke2	22	0.83	0.84	0.059 (0.007)
Oke4	4	0.67	0.60	0.096 (0.034)
Ots3.1	11	0.61	0.63	0.040 (0.012)
Oki10	44	0.94	0.92	0.020 (0.003)
Oki11	5	0.47	0.36	0.028 (0.008)
Ots311	41	0.93	0.87	0.013 (0.002)
OtsG3	4	0.63	0.61	0.058 (0.013)
OtsG68	32	0.92	0.81	0.024 (0.005)
OtsG253	41	0.92	0.91	0.020 (0.005)
OtsG432	24	0.69	0.67	0.118 (0.015)
OtsG474	6	0.12	0.12	0.037 (0.013)

noted at 17 of the 34 loci, but only 6 loci were out of HWE in more than 1 or 2 stocks. *Ssa197* and *Oki100* showed an excess of homozygotes in three stocks, but this was consistent with binomial expectations. *Ots102* showed an excess of homozygotes in 18 stocks, likely the result of a null allele or differential allelic amplification (Wattier et al. 1998). *Oki11*, *OtsG68*, and *Ots311* showed an excess of homozygotes in six or seven stocks. However, 12 of the 19 departures occurred in Canadian stocks, which consisted of older chelex extractions that often proved difficult to amplify; hence, genotyping errors may be responsible for the disequilibrium (Lewis 2002). Overall 646 tests were performed whereof approximately 9% showed significant departures from HWE, which generally conformed to the type I error rate expectation. With the exception of *Ots102* and possibly *Oki11*, *Otsg68*, and *Ots311*, no other loci or stocks were judged to be out of HWE. These loci were dropped from all further analyses; in addition, data from only one lab was retained in further stock structure analyses from those loci that were assayed by more than one lab (*Oke4*, *Ots107*, *Ots100*, *Ots2*). Because the labs had varying levels of success at collecting data from the overlapping loci, the versions that produced the most accurate mean estimates in 100% individual stock MSA simulations were retained. Stock by locus sample sizes for the 22 loci that were further analyzed can be found in Appendix 2.

Stock Structure

The distribution of genetic diversity had a geographic pattern as indicated by the neighbor-joining analysis (Figure 2). The dendrogram revealed eight apparent regional groups: lower river, middle river, border U.S., border Canada, Stewart drainage, Pelly drainage, Canada mainstem, and upper Canada. At a November 2005 meeting, JTC members representing Alaska defined three geographic management units for Alaska that incorporated some genetic data but were based primarily on the nature of the fishery in that area (i.e. subsistence or commercial); therefore, to aid interpretation of the genetic data with respect to those geographic management units, the neighbor-joining groups were realigned accordingly. This involved removing Henshaw Creek from the middle river group and placing it in the border U.S. group. The three geographic management units for Alaska and the five neighbor-joining groups for Canada were used in the following hierarchical and MSA analyses (Table 3).

The correlation between genetic distance and geographic distance was significant ($r = 0.71$; $P < 0.00001$; Figure 3). This suggested migration-drift equilibrium, and that the isolation by distance model explained the distribution of genetic diversity for Yukon River Chinook salmon (Slatkin 1993). However, closer inspection revealed that the fit of the data to the model differed depending on spatial scale (Figure 3). A lowess curve, which smoothes data by locally fitting data to the regression model and generally provides a clearer picture of the x-y relationship, revealed that gene flow had a greater influence among stocks within 1400 km of each other whereas genetic drift predominated among stocks separated by greater distances.

Stocks were genetically divergent ($P < 0.0001$) within and among regions as measured by log-likelihood ratio analysis (Table 4). Divergence among regions was approximately six times greater than within regions ($F_{476, 748} = 5.99$, $P < 0.0001$). The Stewart drainage and upper river stocks were the least and most divergent groups, respectively. While maintaining the type I error rate at $\alpha = 0.05$, the number of individual loci that were divergent within the hierarchy follows: 22 loci among regions, 14 loci within the lower river, 3 loci within the

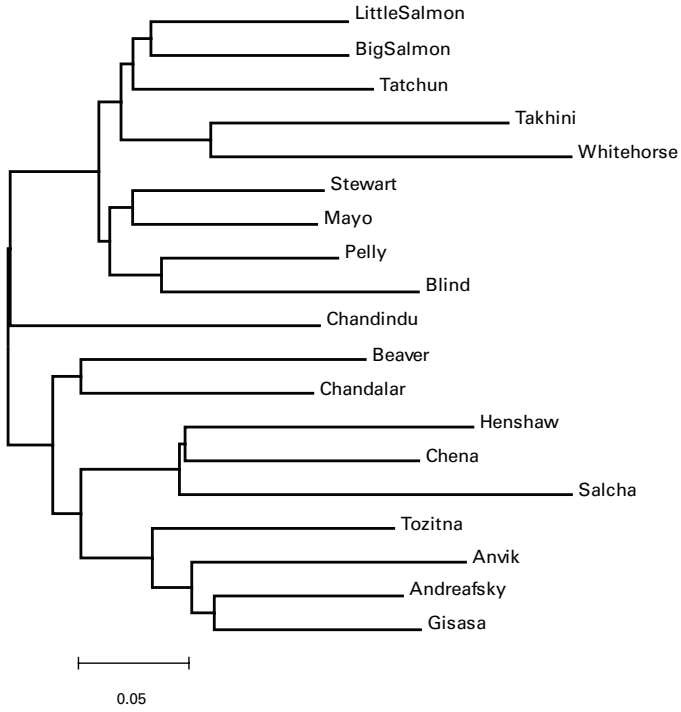

Figure 2. Neighbor-joining dendrogram of chord distances.

Table 3. Hierarchy of stocks used in log-likelihood ratio, gene diversity and mixed-stock analyses.

Region	Stock
Lower River	Andreafsky River
	Anvik River
	Gisasa River
	Tozitna River
Tanana Drainage	Salcha River
	Chena River
Border U.S.	Beaver Creek
	Chandalar River
	Henshaw River
Border Canada	Chandindu River
Stewart Drainage	Stewart River
	Mayo River
Pelly Drainage	Pelly River
	Blind River
Canada Mainstem	Tatchun River
	Little Salmon River
	Big Salmon River
Upper Canada	Takhini River
	Whitehorse Hatchery

Tanana drainage, 13 loci within the border U.S. region, 0 loci within the Stewart drainage, 4 loci within the Pelly drainage, 7 loci within the Canada mainstem region, and 19 loci within the upper river. The high number of loci showing divergence among stocks in the lower river, border U.S., and upper river groups suggests that further substructure exists, and that realignment may be appropriate in the future as baseline collections are added.

Partitioning the genetic diversity revealed that the majority (96.37%) occurred within stocks while among stocks accounted for 3.63% (Table 5). Regional divergence was responsible for 2.91% whereas within region divergence explained 0.72% of the total among stocks diversity. All values were significantly ($P < 0.0001$) different from zero. The overall estimate of gene flow was 6.62 migrants per generation and ranged from 6.38 to 34.16 depending on the hierarchical level. In cases where $4 < N_e m < 1$, gene flow estimation using G_{ST} statistics is robust to violations of drift-migration equilibrium (Hutchinson and Templeton 1999), an assumption which may not have been satisfied for Yukon River Chinook salmon.

Concordance of Data Sets

No significant differences were found among estimates of gene diversity from the three lab-specific baselines (Table 6). Furthermore, the partial correlation coefficient between

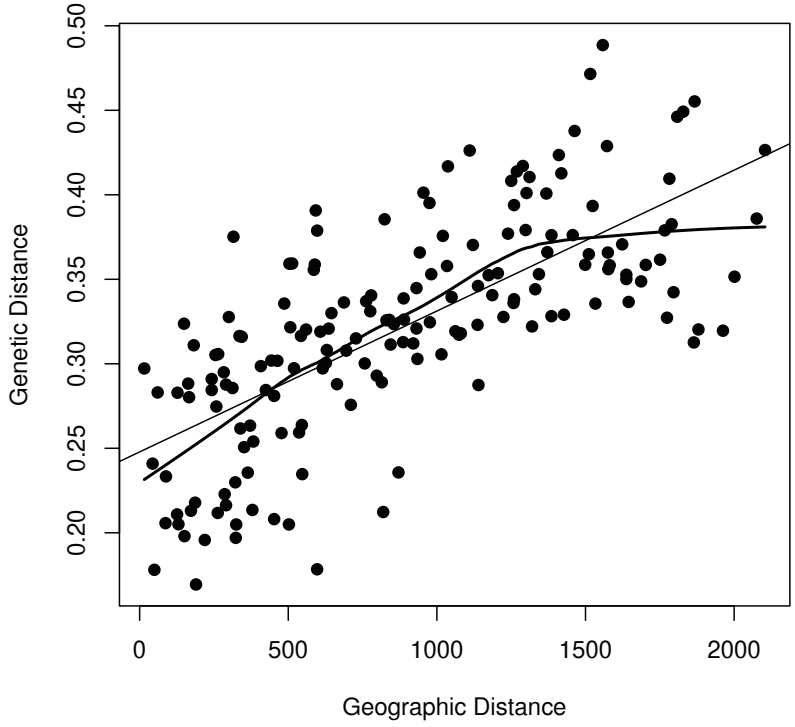

Figure 3. Regression analysis of genetic distance (chord) on geographic distance (km) separating stocks. Both linear ($y = 0.0001x + 0.248$, $r = 0.71$, $P < 0.00001$) and lowess trend lines are displayed.

the residuals from the regression of the ADFG and DFOC matrices on the USFWS matrix was significant ($r_{ij.k} = 0.46$, $P < 0.00001$) and positively correlated. These analyses provided evidence that the three labs observed the same genetic structure, and that combining the data was appropriate.

Mixed-Stock Analysis Simulations

Bayesian allele frequency estimation (Rannala and Mountain 1997) to prevent sampling zeros provided the best 100% individual stock simulation results as compared to the allele binning method; thus, the Bayesian method was used in all MSA analyses.

Analysis of simulated mixtures composed of individual stocks using the best 10 loci resulted in estimates above 90% accuracy for all but the Salcha, Stewart, and Little Salmon stocks (Table 7). The best 10 loci produced regional estimates that ranged in accuracy from 94%–99% and estimates for country of origin that were 99% accurate (Table 8). The worst 10 loci failed to provide estimates that were ≥ 90% accurate for six of the eight regions, yet the ran-

Table 4. Hierarchical tests of homogeneity using log-likelihood ratio analysis (Sokal and Rohlf 1995) of allele frequencies at 22 microsatellite loci among stocks within a region, and among regions. No test indicates that a single stock was collected for that region.

Source of Variation	Total df	Total *G*-test
Among Regions within Total	476	14344.5*
Among Populations within Regions	748	3763.2*
Lower	204	950.2*
Tanana	68	285.7*
Border U.S.	136	739.9*
Border Canada	No test	
Stewart	68	103.7*
Pelly	68	220.2*
Mainstem Canada	136	465.9*
Upper	68	997.6*
Total	1224	18107.7*

*$P < 0.0001$

Table 5. Hierarchical gene diversity analysis of Yukon River Chinook salmon using 22 microsatellite loci.

Source	Absolute Gene Diversity	Relative Gene Diversity	$N_e m$
Average within stocks	$H_S = 0.6524$	$H_S/H_T = 0.9636$	
Average among stocks within regions	$D_{SR} = 0.0049$	$G_{SR} = 0.0072$*	34.16
Average among regions	$D_{RT} = 0.0197$	$G_{RT} = 0.0291$*	6.38
Total gene diversity	$H_T = 0.6771$	$G_{ST} = 0.0363$*	6.62

*$P < 0.0001$ Inferred from hierarchical tests of homogeneity

dom 10 loci provided estimates that exceeded 90% accuracy for seven of the eight regions. The 13 CTC loci provided comparable estimates to that of the best 10 loci (Table 8).

Country of origin and regional apportionment accuracies of simulated fish for previously published allozyme (Templin et al. 2005; 22 polymorphic loci) and SNP (Smith et al. 2005a; 9 polymorphic loci) baselines were compared with apportionment accuracies for the loci subsets (best, random, worst) from the present study and with those for the 13 loci in the CTC

Table 6. Chi-square test to determine if coefficients of gene differentiation among the three lab-specific baselines are significantly different.

	USFWS	DFOC	ADFG
G_{ST}	0.0376	0.0296	0.0476
	df	Chi-square	
FWS DFOC	18	22.86*	
FWS ADFG	18	22.79*	
DFOC ADFG	18	28.95*	

*$P > 0.05$

baseline. Apportionment accuracies of simulated fish to country of origin were similar for all baselines (98–99% with allozymes, 98%–99% with SNPs, 98%–99% with microsatellites); however, apportionment accuracies to the six regions were higher using the present best 10 and 13 CTC microsatellites (Table 9) than they were using the published allozyme (94%–97%) or SNP (85%–94%) baselines.

Analysis of Simulated Multi-Stock Mixtures

Three realistic mixtures that could be expected in a lower river fishery were simulated using the best 10 loci, which provided the best MSA estimates in 100% simulations. Estimated stock compositions of a simulated mixture containing fish from both the U.S. and Canada were usually within 1% of the actual value for a specific region of origin (Table 10, mixture 1). Mixtures containing only simulated Canadian-origin Chinook salmon, derived from 10 Canadian stocks, were estimated to specific region within 1% of the actual value, and virtually all of the samples were estimated to be of Canadian origin (Table 10, mixture 2). Mixtures containing only simulated U.S.-origin Chinook salmon, derived from nine stocks in Alaska, were usually within 2% of the actual value (Table 10, mixture 3). The above 100% and multi-stock simulations suggest that accurate estimates of stock composition should be available when a microsatellite baseline is applied to any mixture sample drawn from the Yukon River drainage.

Evaluation of Accuracy and Precision

An analysis of simulated single-stock mixtures revealed that the number of fish surveyed in a baseline stock had a marked effect upon the accuracy of estimated stock compositions. A substantial increase in accuracy of estimated stock compositions was observed for sample sizes up to approximately 150–200 individuals, only marginal increases in accuracy were obtained by increasing sample size beyond this point (Figure 4).

As noted previously, the number of alleles observed at the microsatellite loci surveyed ranged from 2 to 61 alleles (Table 2). The range in the number of alleles observed among the loci allowed for a comparison of the effect of allele number on the relative power of the locus to estimate stock composition of representative single-stock samples from stocks located throughout the Yukon River drainage (Table 7). The number of alleles observed at a locus was significantly related to the power of the locus in providing accurate estimates of stock composition of single-stock mixtures ($r^2 = 0.84$, $P < 0.01$) (Figure 5). Mean estimated stock compositions of single-stock mixtures for loci with < 20 alleles, with 20–40 alleles, and with > 40 alleles were 46%, 80%, and 84%, respectively. In general, loci with more alleles provided greater resolution of single-stock mixtures than did loci with fewer alleles. This was supported by multilocus simulations using the best, worst and random 10 loci. The best 10

Table 7. Mean estimated percentage stock compositions of 100% single stock mixtures (correct = 100%) for 19 representative stocks of Chinook salmon calculated for individual loci and for a multilocus group that comprised the best 10 loci as determined by WHICHLOCI. The loci that are a part of the best 10 group have bold lettering. Simulations were conducted using a 19-stock baseline, 400 fish in the mixture sample and 1000 resamplings of the mixture sample and baseline samples. *N* represents the mean sample size for the best 10 loci.

	No. of Alleles	36	44	35	33	39	34	15	12	6	6	3	2	3	9
	Locus	***Ssa197***	***Ots101***	***Ots107***	***Oki100***	*Ots100*	***Ots104***	*Ogo4*	*Ogo2*	*Oke4*	***Oke4***	*One102*	*One7*	*One9*	*Ots1*
Stock	Andreafsky	85.1	82.1	80.9	71.4	78.0	84.6	74.6	50.7	27.3	55.4	17.3	40.7	11.7	54.9
	Anvik	82.7	80.6	69.6	72.7	73.5	62.0	67.5	60.3	9.5	28.9	19.7	27.6	82.5	53.9
	Gisasa	81.7	81.5	79.2	88.2	74.0	81.9	88.1	49.0	78.1	65.7	17.0	79.0	16.0	44.3
	Tozitna	92.7	94.6	89.9	91.4	76.0	90.8	85.8	81.5	36.7	48.2	18.5	24.9	18.2	89.2
	Beaver	84.7	81.7	76.0	92.0	84.4	85.7	80.5	53.4	42.2	64.1	10.5	13.7	12.3	41.8
	Chandalar	80.7	90.9	74.4	83.8	81.4	75.3	76.0	86.4	24.9	39.1	13.8	13.7	10.8	57.9
	Chena	89.2	87.8	89.4	85.3	90.6	81.9	88.6	75.2	78.4	76.2	39.8	13.8	51.7	47.9
	Henshaw	92.3	96.8	92.9	95.4	93.8	95.0	83.6	77.7	53.2	72.6	10.8	16.5	12.9	63.8
	Salcha	75.7	80.8	82.1	74.1	85.1	81.6	78.4	55.6	69.2	74.6	19.0	15.8	25.5	78.0
	Blind	89.8	85.0	92.8	78.5	89.3	87.0	81.5	50.6	29.6	67.7	17.5	10.9	7.9	63.7
	Mayo	79.4	75.9	74.2	72.4	86.6	74.9	68.9	49.1	53.1	53.4	9.2	15.5	7.1	22.0
	Pelly	90.9	81.4	77.1	75.4	76.8	70.7	51.0	72.9	52.5	45.2	18.5	15.8	8.3	53.9
	Stewart	62.7	72.9	80.6	72.8	54.2	65.3	48.9	39.9	63.8	55.1	14.8	11.1	16.4	45.7
	Big_Salmon	83.0	68.4	76.6	75.9	79.2	63.0	82.3	44.6	12.3	80.1	10.5	9.7	7.2	40.2
	L._Salmon	65.6	59.3	62.7	58.4	75.1	71.4	59.7	44.6	8.4	20.1	8.8	8.6	10.1	53.0
	Tatchun	87.7	86.5	92.0	89.4	88.5	90.5	80.0	70.0	73.7	74.7	66.8	9.8	12.7	71.1
	Chandindu	95.4	94.6	93.1	96.6	96.3	94.3	95.4	89.6	94.2	94.6	12.6	19.2	19.2	78.3
	Takhini	95.5	92.2	91.6	94.6	89.2	86.5	86.8	53.2	71.9	72.2	12.4	13.8	15.5	55.3
	Whitehorse	96.6	97.7	97.5	94.9	90.8	94.9	94.1	81.2	96.0	94.4	34.7	13.2	11.5	91.6
	Average	84.8	83.7	82.8	82.3	82.3	80.9	77.5	62.4	51.3	62.2	19.6	19.6	18.8	58.2

Continued on next page.

Table 7. *Continued.*

	No. of Alleles	61	33	16	23	6	44	22	11	4	4	6	7	15		
	Locus	***Ots100***	*Ots107*	*Ots2*	*Ots212*	*µ73*	***Oki10***	***Oke2***	*Ots3.1*	*Oke4*	*OtsG3*	*Ots9*	*Ots2*	***Omy325***	Best 10	*N*
Stock	Andreafsky	84.0	77.8	68.6	75.0	51.0	77.6	80.9	80.0	37.5	32.8	71.2	60.2	48.9	97.0	189
	Anvik	76.5	79.2	19.5	77.5	85.9	71.1	51.0	53.4	29.9	27.7	19.4	72.1	66.0	88.6	67
	Gisasa	87.4	81.3	62.4	71.0	81.2	90.9	76.1	67.2	86.2	41.5	53.8	48.9	62.3	96.5	172
	Tozitna	91.2	88.4	38.5	95.1	74.8	93.7	85.5	36.8	23.5	45.6	54.8	30.1	80.3	98.9	355
	Beaver	92.6	74.9	9.6	83.7	19.9	89.5	86.0	74.8	29.0	34.0	13.8	7.2	59.9	95.8	89
	Chandalar	86.8	78.6	45.1	30.3	29.2	86.8	68.1	58.9	25.5	60.1	40.2	6.7	55.5	95.7	95
	Chena	91.4	91.6	39.1	68.8	25.4	94.5	93.3	85.9	90.7	21.6	56.6	58.5	47.0	98.1	155
	Henshaw	96.2	92.1	45.0	83.3	79.1	95.4	83.6	66.5	68.2	65.9	71.6	19.3	84.6	97.5	139
	Salcha	91.2	81.4	18.8	34.2	68.2	85.9	56.0	46.8	67.8	64.2	13.5	31.2	t	90.6	38
	Blind	92.1	89.4	13.8	72.9	60.2	90.2	88.8	51.9	69.2	24.3	21.7	9.1	80.8	95.1	111
	Mayo	78.6	60.0	34.2	43.2	26.2	82.7	71.4	27.4	50.3	12.9	28.1	29.8	53.9	94.4	145
	Pelly	80.1	76.7	53.4	43.9	24.7	76.3	73.2	71.6	41.8	17.9	48.9	35.2	66.1	95.2	129
	Stewart	73.5	77.9	17.7	58.4	56.1	63.3	63.1	54.9	69.4	19.2	13.0	11.9	60.3	87.4	90
	Big_Salmon	78.2	75.4	52.1	52.3	33.2	60.4	73.2	40.5	65.3	32.2	14.4	15.4	70.4	91.7	99
	L._Salmon	68.1	55.1	6.3	49.5	17.9	74.9	62.6	14.6	15.2	20.0	14.2	23.1	42.3	86.3	74
	Tatchun	89.7	92.6	63.3	84.7	72.8	75.9	70.0	51.9	35.7	18.5	33.1	44.2	86.5	97.8	255
	Chandindu	96.3	93.0	27.9	91.7	60.3	92.6	91.3	74.8	94.0	17.3	71.8	19.8	92.8	98.8	276
	Takhini	95.8	91.8	17.5	79.5	27.4	94.7	93.3	90.5	74.2	92.6	89.7	10.6	93.2	98.4	149
	Whitehorse	96.7	96.0	55.3	93.0	91.0	97.3	94.4	79.2	93.6	72.6	19.6	49.1	94.2	98.5	196
	Average	86.6	81.7	36.2	67.8	51.8	83.9	76.9	59.3	56.2	37.9	39.4	30.7	69.3	94.9	

Table 8. Mean estimated percentage stock compositions of 100% regional and country of origin mixtures with the 10 best and worst microsatellite loci as determined by WHICHLOCI, 10 randomly chosen loci, and 13 CTC loci. Simulations were conducted using a 19-stock baseline representing eight regions, 400 fish in the mixture sample, and 1000 resamplings of the mixture sample and baseline samples. Regional and country of origin mixtures represented by equal proportions of fish from stocks in the region or country.

	Best 10	**Worst 10**	**Random 10**	**CTC 13**
Region	Est. (S.D.)	Est. (S.D.)	Est. (S.D.)	Est. (S.D.)
Lower	99.5 (0.4)	97.3 (1.5)	99.1 (0.6)	98.6 (0.7)
Tanana	95.3 (1.0)	85.0 (5.5)	93.8 (1.4)	95.6 (1.4)
Border U.S.	96.5 (1.3)	88.8 (4.4)	93.2 (1.9)	95.4 (1.3)
Border Canada	98.9 (0.5)	88.7 (5.1)	98.0 (0.8)	99.3 (0.4)
Stewart	93.9 (1.5)	74.3 (8.8)	89.2 (2.9)	94.1 (1.8)
Pelly	96.3 (1.2)	84.2 (5.1)	93.0 (2.3)	94.7 (1.5)
Canada Mainstem	96.3 (1.2)	75.3 (6.7)	91.7 (2.3)	94.1 (1.9)
Upper Canada	98.4 (0.6)	94.1 (2.5)	97.4 (2.3)	98.3 (0.7)
U.S.	99.1 (0.5)	97.6 (1.7)	98.9 (0.7)	99.1 (0.5)
Canada	99.4 (0.5)	98.2 (1.5)	99.1 (0.7)	99.5 (0.4)

Table 9. Mean estimated percentage stock compositions of 100% regional and country of origin mixtures with the 10 best and worst microsatellite loci as determined by WHICHLOCI, 10 randomly chosen loci, and 13 CTC loci. Simulations were conducted using a 19-stock baseline representing six regions, 400 fish in the mixture sample, and 1000 resamplings of the mixture sample and baseline samples. Regional and country of origin mixtures were represented by equal proportions of fish from stocks in the region or country.

	Best 10	**Worst 10**	**Random 10**	**CTC 13**
Region	Est. (S.D.)	Est. (S.D.)	Est. (S.D.)	Est. (S.D.)
Lower	99.5 (0.4)	97.3 (1.5)	99.1 (0.6)	98.6 (0.7)
Middle	96.8 (1.1)	93.7 (2.4)	95.7 (1.2)	97.2 (1.1)
Border Canada	98.9 (0.5)	88.7 (5.1)	98.0 (0.8)	99.3 (0.4)
Pelly/Stewart	96.8 (1.1)	83.6 (6.4)	93.8 (2.0)	96.5 (1.3)
Canada Mainstem	96.3 (1.2)	75.3 (6.7)	91.7 (2.3)	94.1 (1.9)
Upper Canada	98.4 (0.6)	94.1 (2.5)	97.4 (1.3)	98.3 (0.7)
U.S.	99.1 (0.5)	97.6 (1.7)	98.9 (0.7)	99.1 (0.5)
Canada	99.4 (0.5)	98.2 (1.5)	99.1 (0.7)	99.5 (0.4)

loci averaged 34 alleles per locus and produced estimates for the regions that ranged in accuracy from 94%–99%. The worst 10 loci had on average 7 alleles per locus and estimates that ranged in accuracy from 74%–94%, and the random 10 loci had on average 20 alleles per locus and estimates that ranged in accuracy from 89%–98%.

The number of microsatellite alleles used in the stock composition analysis directly influenced the average accuracy and precision obtained in resolving single-stock mixtures. For stocks difficult to identify, such as the Little Salmon River, increasing the number of microsatellite alleles to the maximum 282 alleles available in the DFOC loci set resulted in the maximum accuracy of estimates for this stock (Figure 6). For distinct stocks, such as Chandindu River (Table 7), 94% accuracy in estimated stock compositions was achieved by employing approximately six alleles. For the average stock, increasing the number of alleles employed in stock composition analysis consistently increased the accuracy of the estimates. For estimates up to 80% accuracy, each additional allele used in the estimation increased accuracy by about 1%, so that 80% accuracy for the average stock was achieved by employing approximately 80 alleles in the analysis. Increasing the accuracy of estimated stock compositions to 90% for the average stock required approximately 180 microsatellite alleles. Diminishing returns in accuracy per allele added were observed once approximately 200 alleles were employed in the analysis.

Table 10. Estimated percentage stock compositions of simulated mixed-stock mixtures of Yukon River Chinook salmon as may be encountered in mixed-fisheries. Simulations were conducted using a 19-stock baseline, 400 fish in the mixture sample, and 1000 resamplings of the mixture sample and baseline samples. Regional and country of origin mixtures were represented by equal proportions of fish from stocks in the region or country. The best 10 loci were used in the analysis.

	Expected	Estimated	S.D.
Mixture 1			
Upper	40	39.2	2.4
Pelly Drainage	10	9.7	1.4
Canada	50	50.3	2.4
Lower	20	20.1	2.0
Tanana Drainage	20	18.9	2.1
Border U.S.	10	9.7	1.3
U.S.	50	49.7	2.4
Mixture 2			
Upper	20	19.9	2.1
Canada Mainstem	25	25.3	2.5
Pelly Drainage	25	24.8	1.7
Stewart Drainage	20	19.5	2.4
Border Canada	10	9.9	1.6
Canada	100	99.4	0.5
Mixture 3			
Border U.S.	20	19.2	2.2
Tanana Drainage	30	28.5	2.5
Lower	50	51.4	2.6
U.S.	100	99.1	0.5

Relationships between accuracy and precision of estimated stock compositions were investigated further with the ADFG set of loci. The results from using a single locus with high numbers of alleles (*Ots100* with 61 alleles) for stock composition estimation were compared with the results obtained from using the entire ADFG set of 10 loci with 164 alleles. Accuracy of estimated stock compositions for each of the 19 stocks was very similar between *Ots100* and the entire ADFG set of loci (Figure 7a). However, increasing the number of loci used in the estimation procedure reduced the variance associated with the estimates (Figure 7b).

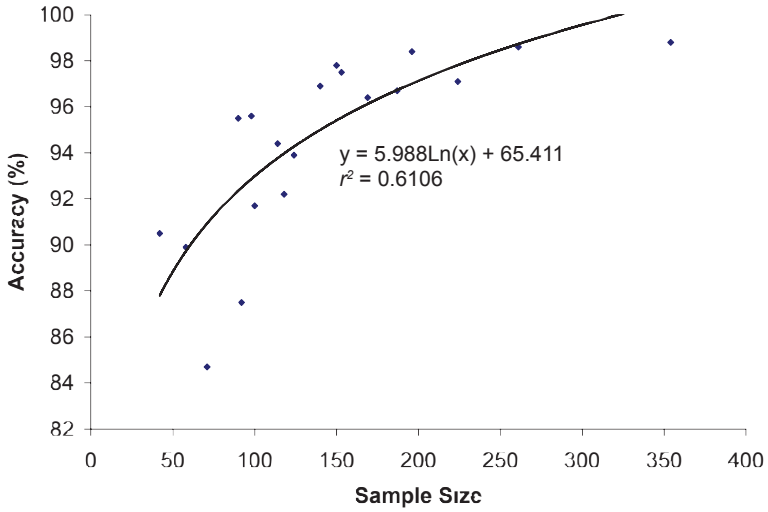

Figure 4. Relationship between stock sample size and accuracy for 19 stocks of Yukon River Chinook salmon. Simulations were conducted using a 19-stock baseline, 22 microsatellite loci, 400 fish in the mixture sample, and 1000 resamplings in the mixture sample and baseline samples.

Mixed-Stock Analysis of Pilot Station Test Fishery Samples

Apportionment of Pilot Station samples using the ADFG microsatellite baseline indicated a decreasing Canadian contribution over the collection period (Table 11). Comparing the apportionments of microsatellites, allozymes and SNPs revealed some disagreements. For the first period, the estimates were similar for the lower Yukon River, but only allozyme and microsatellite estimates were similar for the middle and upper Yukon River and for country of origin. For the second period, all estimates for the upper Yukon River and country of origin were similar, but they were different for the lower and middle Yukon River; the microsatellite estimate was intermediate to those from allozymes and SNPs. For the third period, all estimates were similar.

Discussion

Stock Structure

Microsatellite data reveals spatial structuring among Yukon River Chinook salmon stocks, which generally corresponds well with allozyme (Templin et al. 2005) and SNP (Smith et al. 2005a) data. Values of G_{ST} are similar for Yukon River Chinook salmon among allozymes (0.051; Templin et al. 2005), SNPs (0.069; Smith et al. 2005a) and microsatellites (0.036), but the different mutational properties of these markers precludes a direct comparison of this statistic (Olsen et al. 2004). Contrasting the normalized log-likelihood ratios (*G*-test/df) over all stocks reveals that microsatellites harbor approximately twice the level of heterogeneity than do allozymes ($F_{1224, 539} = 1.83$, $P < 0.00001$) and SNPs ($F_{1224, 220} = 2.13$, $P < 0.00001$).

An exception to the general marker correspondence exists. Microsatellite and allozyme (Templin et al. 2005) data show similar patterns of correlation between genetic and geo-

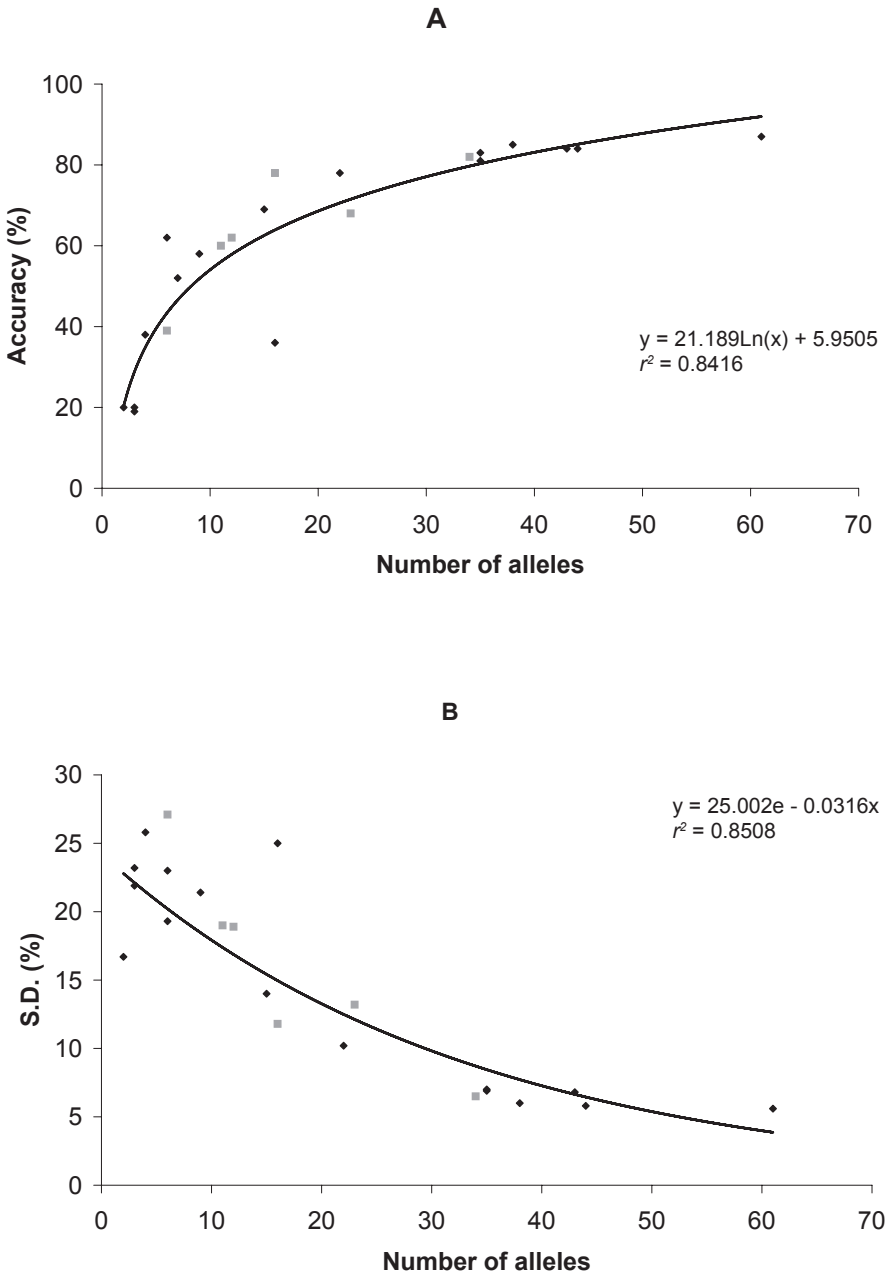

Figure 5. A) Average accuracy and B) average standard deviation for single stock, single locus simulations versus the number of alleles observed at the locus. Accuracy for the locus was determined as the mean value over all 19 stocks. Gray squares are loci in common with the 13-locus baseline defined by the Genetic Analysis of Pacific Salmon consortium under the Chinook Salmon Technical Committee of the Pacific Salmon Commission.

Table 11. Stock composition estimates of 2003 Pilot Station test fishery Chinook salmon samples using the ADFG microsatellite baseline and SPAM 3.7. Upper Yukon and Canada are the same reporting group. Allozyme and SNP estimates are from Templin et al. (2005) and Smith et al. (2005a). CI = 90% confidence intervals.

		Period 1 Estimate	CI	Period 2 Estimate	CI	Period 3 Estimate	CI
Lower Yukon	Micosatellites	12	9–19	14	10–22	46	37 –55
	Allozymes	6	0–15	35	15–44	47	34 –63
	SNPs	1	0–10	5	0–12	49	37–61
Middle Yukon	Micosatellites	30	22–36	33	23–39	17	10–24
	Allozymes	35	24–44	14	3–29	16	5–25
	SNPs	56	42–68	47	31–58	14	8–28
Upper Yukon	Micosatellites	57	50–64	50	41–56	36	28–43
	Allozymes	59	48–70	51	41–67	37	24–50
	SNPs	43	28–55	49	39–62	36	24–48
U.S.	Micosatellites	42	35–49	47	39–55	63	55–70
	Allozymes	41	28–50	49	33–59	63	49–75
	SNPs	57	45–72	51	38–62	64	52–76
Canada	Micosatellites	57	50–64	50	41–56	36	28–43
	Allozymes	59	48–70	51	41–67	37	24–50
	SNPs	43	28–55	49	39–62	36	24–48

graphic distance, suggesting that stocks are influenced less by gene flow than by genetic drift when separated by more than 1000–1400 km, and that isolation by distance and drift/migration equilibrium may not be representative of the entire drainage. However, SNP data (Smith et al. 2005a) do not show this trend, a lowess regression analysis reveals a strong correlation between genetic and geographic distances throughout the Yukon River drainage for Chinook salmon. Such a discrepancy could be a sampling/statistical artifact or reflect differences in mutation rates and/or selection gradients among the markers and highlights potential risks associated with inferring population structure from a small number of loci or loci whose allele frequencies are potentially driven by selection (Smith et al. 2005b).

In comparison to Yukon River chum salmon (Wilmot et al. 1992; Scribner et al. 1998; Flannery 2004) the resolution among Chinook salmon stocks is quite high and likely the result of either small stock sizes and/or increased spawning fidelity. Straying and stock size indirectly relate to gene flow and genetic drift, two of the main factors affecting contemporary stock structure. While little is known about straying rates of Yukon River Chinook or chum salmon, estimates for both from other locations overlap but are generally lower for Chinook salmon, ranging from 1.4%–27% for Chinook salmon and 2.5%–42% for chum salmon (Meyers et al. 1998; Quinn 1993; Tallman and Healey 1994). Stock sizes appear to be much smaller for Chinook salmon in the Yukon River with no more than 10,000 fish for any individual stock and most at 1,000 fish or less while many chum salmon stocks exceed 50,000 fish (JTC 2001). The relatively high levels of divergence for Chinook salmon are likely more a result

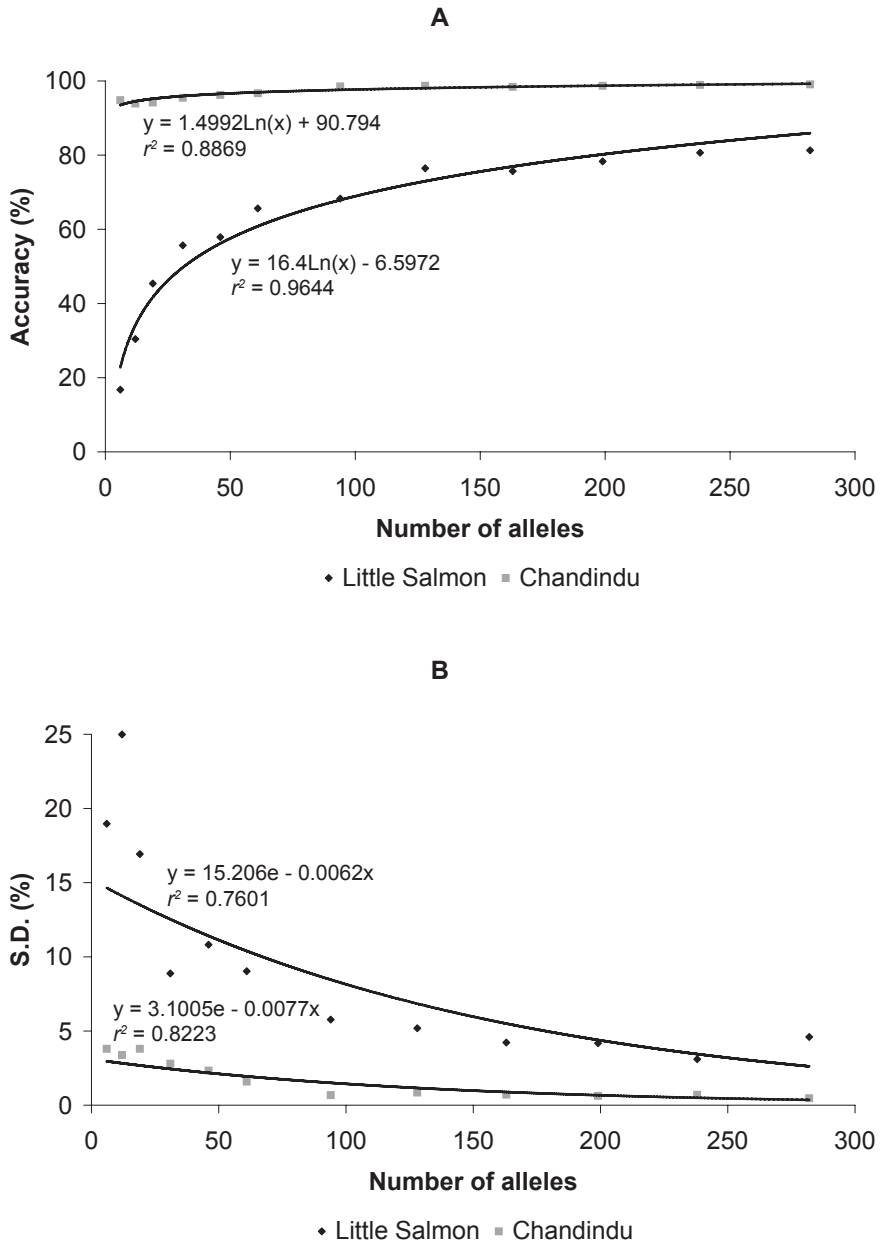

Figure 6. Relationship between the number of microsatellite alleles used in estimating stock compositions and the A) accuracy and B) standard deviation obtained for single-stock mixtures of Chandindu River Chinook salmon (stock identified most accurately) and Little Salmon River Chinook salmon (stock identified least accurately). Simulations were conducted using a 19-stock baseline, 400 fish in the mixture sample and 1000 resamplings in the mixture sample and baseline samples. Allele counts were determined by adding the DFOC loci sequentially to the analysis, one locus at a time, beginning with the loci with the fewest number of alleles.

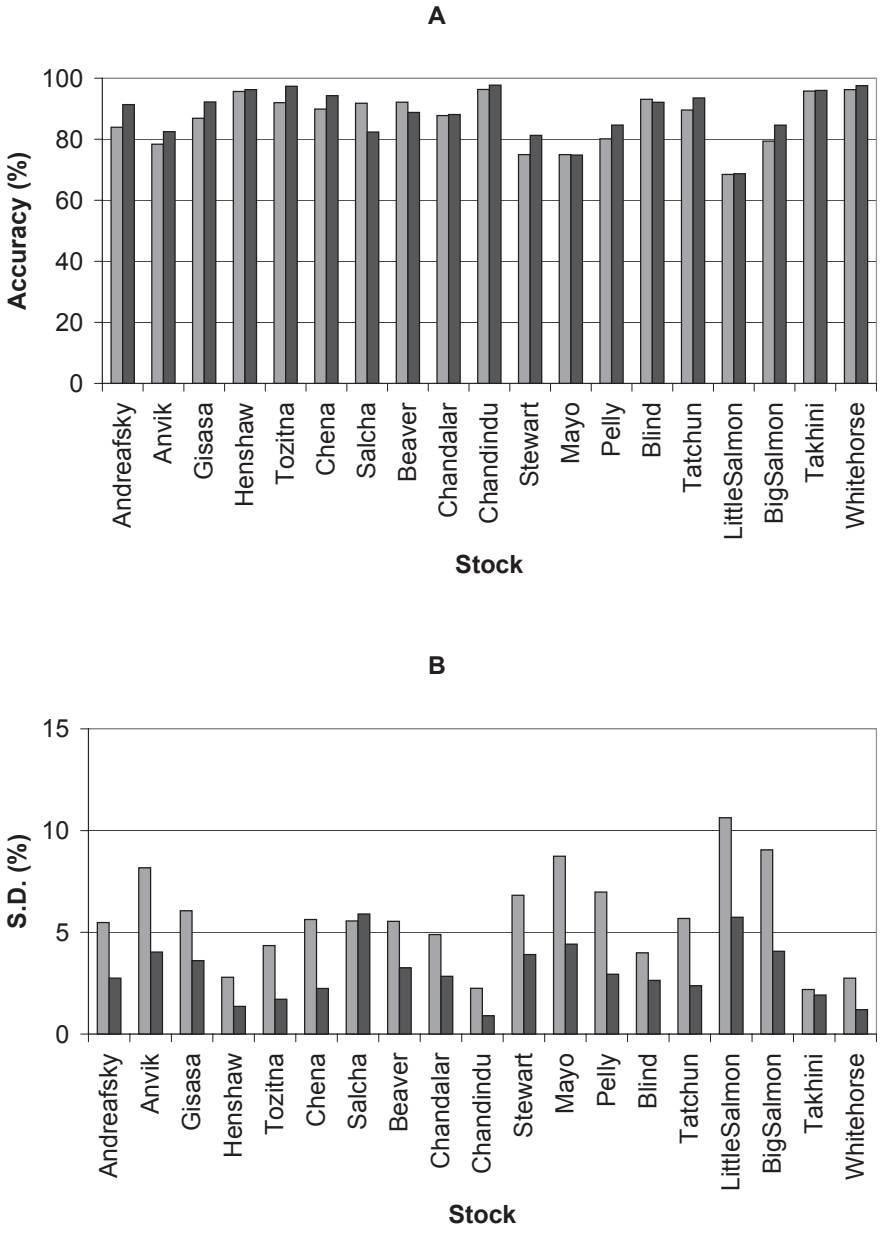

Figure 7. A) Accuracy of only *Ots100* (gray bars) and the entire ADFG suite of 10 loci (including *Ots100*, black bars) for 19 Chinook salmon stocks in the Yukon River drainage. Simulations were conducted using a 19-stock baseline, 400 fish in the mixture sample and 1000 resamplings in the mixture sample and baseline samples. B) Standard deviations of individual stock estimates for only *Ots100* and the entire ADFG suite of 10 loci for 19 stocks of Yukon River Chinook salmon.

of disparate census stock sizes and presumably effective stock sizes (N_e) and less a result of variations in straying rates between the species. Chinook salmon stocks probably have smaller N_e than do chum salmon stocks and, thus, are subject to higher rates of genetic drift, which tends to make stocks more dissimilar.

Historic events can also influence stock structure. Yukon River Chinook salmon probably have experienced severe bottlenecks resulting from Pleistocene glaciations greatly reducing water levels. Salmon survive these periods in glacial refuges with expansion following glacial maxima. If residual genetic similarity lingers from historic fragmentation or range expansion events, then contemporary estimates of gene flow can be biased. Distinguishing historic from contemporary factors can be accomplished by taking into account an "alleles' existence through evolutionary time and geographical space" (Templeton 1998). Unfortunately, allele frequency data is unordered as it does not keep track of haplotypes and, thus, contains no intraspecific genealogical information. Alternatively, mitochondrial DNA (mtDNA) restriction site data provide such information, and a nested phylogeographical analysis of the Yukon River chum salmon mtDNA haplotype genealogy detects only contemporary recurrent but restricted gene flow and no historical events (Flannery 2004). Such results give some confidence in the gene flow estimates for Yukon River Chinook salmon assuming the two species experienced similar conditions during Pleistocene glaciations.

Mixed-Stock Analysis

The divergence we observe with microsatellites produces highly accurate and precise MSA simulation estimates for country of origin, regions, most of the major drainages and individual stocks. Managing by country of origin and major geographic region is easily achievable with the present baseline. With a more comprehensive baseline, individual drainage and even specific stock management will likely be possible as 16 of the 19 stocks already achieve $\geq$ 90% MSA simulation accuracy, a level that is typically assumed to indicate a baseline can accurately apportion fishery harvests (Seeb and Crane 1999). Stocks with estimates below 90% accuracy have lower sample sizes; thus, sample size rather than a dearth of genetic divergence may be the limiting factor. To assess this question, we recommend continuing to augment the current baseline to bring sample sizes above 200 for each stock and to add all major contributing stocks not yet sampled (see also Appendix 1).

Additionally, regarding the subset of loci to apply to Yukon River Chinook salmon fisheries, while the three lab-specific baselines are highly correlated and reveal the same structure, it is known that the amount of divergence varies among loci (Ewens 1983), which the MSA comparison of the best and worst 10 loci indicates. The results from this study show a trend wherein larger sample sizes of both alleles per locus and individuals per stock improve MSA estimates. Although comparing loci with disparate sample sizes of individuals is inherently biased, the intent of the analysis is to determine the power of the baseline for MSA and to select the most powerful loci, not necessarily to specifically compare locus performance. These results suggest that future microsatellite work (e.g., increasing sample sizes, adding new stocks) should focus on the more variable loci because this will provide the greatest increase in information. The comparable simulation accuracy estimates for the 13 CTC loci and the best 10 loci from this study indicate that either would be appropriate; however, the benefits of having a Yukon River Chinook salmon baseline that is a subset of a much larger geographic

baseline, which can be used to address questions of a broader nature if needed, clearly dictate the use of the 13 CTC loci. In addition to information content per locus, locus selection will likely be driven by the specific questions being asked and the costs of analyses.

Mixture analysis of 2003 Pilot Station test fishery samples reveals that stock run timing is generally geographically dependent with earlier run timing for fish having a longer migratory distance. Although there is overlap among all stock groups, the data suggest that upper Yukon River Chinook salmon stocks run first followed by middle river stocks and lastly lower river stocks; similar findings are reported by previous studies using MSA and tagging data (Wilmot et al. 1992; Eiler et al. 2004). Comparing the stock composition estimates to those of Templin et al. (2005) and Smith et al. (2005a) reveals that all markers produce similar results for period three, but disagreements occur for periods one and two. Microsatellites and allozymes gave similar results for period one, but SNP estimates are only similar for the lower Yukon River. During period two, there is agreement among markers for country of origin, but for the lower and middle Yukon River the markers are in disagreement, with microsatellite estimates roughly intermediate to those of allozymes and SNPs. Incomplete baseline representation and varying levels of baseline heterogeneity could account for these discrepancies as both can bias estimates (Pella and Milner 1987). Tagging data shows that large stocks exist between the Tanana River and Canada (Eiler et al. 2004). The agreement of estimates for country of origin for period two may suggest that border U.S. stocks are moving through, resulting in regional bias but accurate country of origin estimates for allozymes. Absent border U.S. stocks also cause discrepancies between allozyme and radio-tag country of origin stock composition estimates; adjusting the reporting groups to above and below the Tanana River ameliorates the situation (Templin et al. 2005).

Potential Fishery Management Implications

In the Yukon River, uncertainty of in-season run abundance requires conservative management of Chinook salmon commercial fisheries. Current in-season stock assessment tools do not provide consistently reliable estimates of harvestable surpluses (JTC 2001). Not having precise in-season stock abundance estimates results in forgoing potential harvests, as most regions exceed escapement goals. Lost commercial earnings affect subsistence fishers because most fish commercially to support their subsistence efforts.

Mixed stock analysis using this microsatellite baseline can accurately and precisely allocate Chinook salmon in mixtures to units useful for management, such as region or major drainage, providing managers with a powerful tool for assessing and regulating fisheries. With a comprehensive baseline, which all involved agencies are striving to develop, it is likely that MSA using microsatellites will be able to accurately and precisely allocate Chinook salmon to all the major drainages and tributaries in the Yukon River. Further, information on migration patterns and run timing of regional genetic groups will provide valuable information useful to managers throughout the drainage.

Project Objectives Assessment

1) Attempt to collect 200 samples from each of 20 stocks, 10 from the U.S. and 10 from Canada.

While it will continue to grow and almost continuously be augmented with new samples and markers, the acquisition and compilation of this extensive baseline for use by various agencies and countries is a primary success of this project. These collections will likely form the foundation for Chinook salmon MSA in the Yukon River for years to come.

The large size and remote nature of the Yukon River renders Chinook salmon collections logistically difficult and extremely expensive. Our large sampling effort yields 19 collections that we consider of sufficient size to include in this report. Eight of these collections have the target sample size of 200 and the mean sample size for all 19 collections is 197, ranging from 55–450. Additional samples, not in this report due to time constraints, are available from 2004 collections to enhance the baseline at a later date. Further information on sampling Yukon River Chinook salmon is in Appendix 1.

2) Collect samples from the 2002–2003 Pilot Station test fishery and Russian Mission radiotagging project and conduct mixed-stock analysis on these samples using the allozyme baseline and on a portion of the samples with microsatellites.

Collection and analysis of samples using the allozyme baseline are completed, and Templin et al (2005) report on the results. The analysis of the 2003 Pilot Station samples by the ADFG lab using their suite of microsatellite loci to provide a comparison between allozymes, SNPs, and microsatellites is also completed, which the present study reports above.

3) Genotype baseline samples at a minimum of 10 microsatellite loci per lab, with overlap of 1–2 loci among labs to assist in scoring standardization.

As discussed above, the success rates vary across loci and labs. Constraints of time and money, limited quantities of DNA, variable DNA extraction chemistry, and the age of some extracted DNA prevent full data collection for some locus/stock combinations. However, all three labs have attempted to genotype a minimum of 10 loci for all 19 stocks. In the end, sufficient data exist for 30 of the 34 microsatellite loci for the 19 stocks used in this analysis. All labs are "certified" and standardized following CTC guidelines for the 13 CTC loci, precluding the need to standardize other loci overlapping among labs.

4) Investigate the performance of various allele binning procedures, maximum likelihood and Bayesian statistical approaches to mixture analyses.

Exploration and assessment of various statistical methods is commonplace for all three labs involved in this study. Specifically, during this project, the USFWS has developed a novel allele binning method (Bromaghin and Crane 2005). The ADFG has added two Bayesian allele frequency estimation methods (Rannala and Mountain 1997; Pella and Masuda 2001) to SPAM 3.7, the ADFG maximum likelihood (ML) stock composition estimation program, and the DFOC has developed a faster, more versatile, and user friendly version of the computer program Bayes, which implements Bayesian mixture modeling. Individuals from all three laboratories have collaborated on various aspects of the MSA herein and will continue to explore ways to improve MSA.

Mixed-stock analysis simulations have only been done using the ML method, the computational time that Bayesian mixture modeling (Pella and Masuda 2001) requires prevents

a comparison of the two methods. Previous works by the authors and others for different salmon species indicate that the Bayesian method provides greater accuracy than ML (Pella and Masuda 2001; Koljonen et al. 2005). The accuracy of Yukon River Chinook salmon MSA estimates using ML is already extremely high; Bayesian estimation accuracy would likely be similar as the room for improvement is small but should nonetheless be considered as a viable alternative to be explored in the future.

Conclusions

1) The collection of baseline samples is a primary success of this project because these samples will likely form the foundation for Chinook salmon MSA in the Yukon River for years to come.

2) Yukon River Chinook salmon exhibit a high degree of stock structure.

3) Accurate (> 90%) apportionment to country of origin and major geographic regions is possible with the current baseline.

4) Accurate apportionment to drainages and individual stocks analyzed here, and potentially others not yet identified, is likely, but should not be considered for such specific management applications until the baseline is adequately expanded.

Recommendations

1) Increase sample sizes for those stocks currently below 200 individuals.

2) Continue to identify, sample, and add additional stocks to the baseline.

3) Assay additional baseline samples and stocks with the final set of 13 CTC sanctioned microsatellite loci.

4) Work with interested management agencies to utilize the baseline to aid Yukon River Chinook salmon management.

Acknowledgments

Baseline samples were either collected or coordinated by staff in a variety of organizations, including the U.S. Fish and Wildlife Service, the Department of Fisheries and Oceans Canada, Bering Sea Fisherman's Association, Anvik River Lodge, and Yukon River Drainage Fisheries Association. Mixture samples from Pilot Station test fisheries were collected by the Alaska Department of Fish and Game. The Yukon River Drainage Fisheries Association facilitated the participation of the Department of Fisheries and Oceans Canada. The U.S. Fish and Wildlife Service, Office of Subsistence Management provided $561,360 in funding support for this project through the Fisheries Resource Monitoring Program, project 02-121. Support was also provided by the U.S. Fish and Wildlife Service, Alaska Department of Fish and Game, and the Department of Fisheries and Oceans Canada. Editing and formatting services were provided by Andrea Medeiros, Publication Specialist with the Office of Subsistence Management.

References

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

Apostal, B.L., W.C. Black IV, P. Reiter, and B.R. Miller. 1996. Stock genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. Heredity 76:325–334.

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 (*Oncorhynchus tshawytscha*). Journal of Heredity 90:281–288.

Banks, M.A., W. Eichert, and J.B. Olsen 2003. Which genetic loci have greater stock assignment power? Bioinformatics 19:1436–1438.

Banks, M.A., V.R. Rashbrook, M.J. Calavetta, C.A. Dean, and D. Hedgecock. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of Chinook salmon (*Oncorhynchus tschawytscha*). Canadian Journal of Fisheries and Aquatic Sciences 57:915–927.

Beacham, T.D., C.B. Murray, and R.E. Withler. 1989. Age, morphology, and biochemical genetic variation of Yukon River Chinook salmon. Transactions of the American Fisheries Society 118:46–63.

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

Beacham, T.D., J.R. Candy, K.J. Supernault, T. Ming, B. Deagle, A. Schulze, D. Tuck, K.H. Kaukinen, J.R. Irvine, K.M. Miller, and R.E. Wither. 2001. Evaluation and application of microsatellite and major histocompatibility complex variation for stock identification of coho salmon in British Columbia. Transaction of the American Fisheries Society 130:1116–1149.

Beacham, T.D., J.R. Candy, K.J. Supernault, Michael Wetklo, B. Deagle, K. Labaree, J.R. Irvine, K.M. Miller, R.J. Nelson, and R.E. Wither. 2003. Evaluation and application of microsatellites for stock identification of Fraser River Chinook salmon. Fisheries Bulletin 101:243–259.

Bergstrom, D.J., K.C. Schultz, V. Golembeski, B.M. Borba, D. Huttunen, L.J. Barton, T.L. Lingnau, RR. Holder, J.S. Hayes, K.R. Boeck, and W.H. Busher. 2001. Annual management report Yukon area, 1999. Regional Information Report No. 3A01-01, Alaska Department Fish and Game, Anchorage, Alaska.

Brabets, T.P., B. Wang and R.H. Meade. 2000. Environmental and hydrologic overview of the Yukon River Basin, Alaska and Canada. Water-Resources Investigations Report 99-4204, U.S. Geological Survey, Anchorage, Alaska.

Bromaghin J.F., and P.A. Crane. 2005. A method to bin alleles of genetic loci that maintains population heterogeneity. Canadian Journal of Fisheries and Aquatic Sciences 62:1570–1579.

Buccholz W.G, S.J. Miller, and W.J. Spearman. 2001. Isolation and characterization of chum salmon microsatellite loci and use across species. Animal Genetics 32:162–165.

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

Chakraborty, R. and O. Leimar. 1987. Genetic variation within a subdivided stock. Pages 89–120 *in* N. Ryman and F. Utter, editors. Population Genetics and Fishery Management. University of Washington Press, Seattle, Washington.

Cooper, D.W. 1968. The significance level in multiple tests made simultaneously. Heredity 23:614–617.

Debevec, E.M., R.B. Gates, M. Masuda, J. Pella, J.M. Reynolds, and L.W. Seeb. 2000. SPAM (Version 3.2): Statistics program for analyzing mixtures. Journal of Heredity 91:509–510.

Eiler, J.H., T.R. Spencer, J.J. Pella, M.M. Masuda, and R.R. Holder. 2004. Distribution and movement patterns of Chinook salmon returning to the Yukon River basin in 2000–2002. NOAA Technical Memorandum NMFS-AFSC-148, U.S. Department of Commerce.

Estoup, A., P. Presa, F. Krieg, D. Vaiman, and R. Guyomard. 1993. (CT)n and (GT)n microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). Heredity 71:488–496.

Ewens, W.J. 1983. The role of models in the analysis of molecular genetic data with particular reference to restriction fragment data. Pages 45–73 *in* B.S. Weir, editor. Statistical Analysis of DNA Sequence Data. Marcel Dekker, New York.

Flannery, B.G. 2004. Application of molecular markers to mixed-stock analysis of Yukon River fall chum salmon. Master's thesis. University of Alaska Fairbanks.

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

Goudet, J. 1995. FSTAT A program for IBM PC compatibles to calculate Weir and Cockerham's (1984) estimators of F-statistics (version 1.2). Journal of Heredity 86:485–486.

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

Hutchinson, D.W. and A.R. Templeton. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift of the distribution of genetic variability. Evolution 53:1898–1914.

Ihssen P.E., H.E. Brook, J.M. Casselman, J.M. McGlade, N.R. Payne, and F.M. Utter. 1981. Stock identification: materials and methods. Canadian Journal of Fisheries and Aquatic Sciences 38:1838–1855.

Joint Technical Committee (JTC). 1997. Review of stock identification studies on the Yukon River. The United States and Canada Yukon River Joint Technical Committee.

Joint Technical Committee (JTC). 2001. Yukon River salmon season review for 2001 and technical committee report. The United States and Canada Yukon River Joint Technical Committee.

Koljonen, M.L., J.J Pella, and M. Masuda. 2005. Classical individual assignments versus mixture modeling to estimate stock proportions in Atlantic salmon (*Salmo salar*) catches from DNA microsatellite data. Canadian Journal of Fisheries and Aquatic Sciences 62:2143–2158.

Kondzela, C.M., C.M. Guthrie, S.L. Hawkins, C.D. Russell, J.H. Helle, and A.J. Gharrett. 1994. Genetic relationships among chum salmon populations in Southeast Alaska and Northern British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 51(Supplement 1):50–64.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 12:1244–1245.

Lewis, C.M. 2002. Genetic association studies: design, analysis and interpretation. Briefings in Bioinformatics 3:146–153.

Myers, J.M., R.G. Kope, G.J. Bryant, D. Teel, L.J. Lierheimer, T.C. Wainwright, W.S. Grand, F.W. Waknitz, K. Neely, S.T. Lindley, and R.S. Waples. 1998. Status review of chinook salmon from Washington, Idaho, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-35, U.S. Department of Commerce.

Nei, M. and R.K. Chesser. 1983. Estimation of fixation indices and gene diversities. Annals of Human Genetics 47:253–259.

Nelson, R. J., and T.D. Beacham. 1999. Isolation and cross species amplification of microsatellite loci useful for study of Pacific salmon. Animal Genetics 30:228–229.

NRC (National Research Council). 1996. Upstream: salmon and society in the Pacific Northwest. National Academy Press. Washington, DC.

O'Connell, M., R.G. Danzmann, J.M. Cornuet, J.M. Wright, and M.M. Ferguson. 1997. Differentiation of rainbow trout stocks in Lake Ontario and the evaluation of the stepwise mutation and infinite allele mutation models using microsatellite variability. Canadian Journal of Fisheries and Aquatic Sciences 54:1391–1399.

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

Olsen, J.B., S. L. Wilson, E.J. Kretschmer, K.C. Jones, and J. E. Seeb. 2000. Characterization of 14 tetranucleotide microsatellite loci derived from sockeye salmon. Molecular Ecology 9:2185–2187.

Olsen, J.B., C. Habicht, J. Reynolds, and J.E. Seeb. 2004. Moderately and highly polymorphic microsatellites provide discordant estimates of stock divergence in sockeye salmon, *Oncorhynchus nerka*. Environmental Biology of Fishes 69:21–36.

O'Reilly, P.T., L.C. Hamilton, S. K. McConnell, and J. M. Wright. 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Sciences 53:2292–2298.

Pella, J. and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin 99:151–167.

Pella J.J. and G.B. Milner. 1987. Use of genetic marks. Pages 247–276 *in* N. Ryman and F. Utter, editors. Population Genetics and Fishery Management. University of Washington Press, Seattle, Washington.

Quinn, T.P. 1993. A review of homing and straying of wild and hatchery-produced salmon. Fisheries Research 18:29–44.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Science USA 94: 9197–9201.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

Scribner, K.T., P.A. Crane, W.J. Spearman, and L.W. Seeb. 1998. DNA and allozyme markers provide concordant estimates of stock differentiation: analyses of U.S. and Canadian stocks of Yukon River fall-run chum salmon (*Oncorhynchus keta*). Canadian Journal of Fisheries and Aquatic Sciences 55:1748–1758.

Scribner, K.T., J.R. Gust, R.L. Fields. 1996. Isolation and characterization of novel salmon microsatellite loci:cross-species amplification and stock genetic application. Canadian Journal of Fisheries and Aquatic Sciences 53:833–841.

Seeb, L.W. and P.A. Crane. 1999. Allozymes and mitochondrial DNA discriminate Asian and North American stocks of chum salmon in mixed-stock fisheries along the south coast of the Alaska Peninsula. Transactions of the American Fisheries Society 128:88–103.

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.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium stocks. Evolution 47:264–279.

Small M.P., T.D. Beacham, R.E. Whithler, and R.J. Nelson. 1998. Discriminating coho salmon within the Fraser River, British Columbia, using microsatellite DNA markers. Molecular Ecology 7:141–155.

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

Smith, C.T., W.D. Templin, J.E. Seeb, and L.W. Seeb. 2005a. 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.

Smith C.T., J.E. Seeb, P. Schwenke, and L.W. Seeb. 2005b. Use of the 5'-nuclease reaction for SNP genotyping in Chinook salmon. Transactions of the American Fisheries Society 134:207–217.

Smouse, P.E. J.C. Long, and R.R. Sokal. 1986. Multiple regression and correlation extensions of the mantel test of matrix correspondence. Systematic Zoology 35:627–632.

Smouse, P.E. and R.H. Ward. 1978. A comparison of the genetic infrastructure of the Ye'cuana and the Yanomama: a likelihood analysis of genotypic variation among stocks. Genetics 88:611–631.

Sokal, R.R. and F.J. Rohlf. 1995. Biometry. W.H. Freeman. New York.

Tallman R.F. and M.C. Healey. 1994. Homing, straying, and gene flow among seasonally separated stocks of chum salmon (*Oncorhynchus keta*). Canadian Journal of Fisheries and Aquatic Sciences 51:577–588.

Templeton, A.R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and stock history. Molecular Ecology 7:381–397.

Templin W.D., R.L. Wilmot, C.M. Guthrie III, and L.W. Seeb. 2005. United States and Canadian Chinook salmon stocks in the Yukon River can be segregated based on genetic characteristics. Alaska Fishery Research Bulletin 11:44–60.

Wattier R., C.R. Engel, P. Saumitou-Laprade, and M. Valero. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1ct in Gracilaria Gracilis (*Rhodophyta*). Molecular Ecology 7:1569–1573.

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

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

Wilmot, R.L., R.J. Everett, W.J. Spearman, and R. Baccus. 1992. Genetic stock identification of Yukon River chum and Chinook salmon, 1987–1990. Progress Report, U.S. Fish and Wildlife Service, Anchorage, Alaska.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138

Wright, S. 1978. Evolution and the genetics of stocks: Volume 4, variability within and among natural stocks. The University of Chicago Press, Chicago.

Zhivotovsky, L.A., A.J. Gharrett, A.J. McGregor, M.K. Glubokovsky, and M.W. Feldman. 1994. Gene differentiation in Pacific salmon (*Oncorhynchus sp.*): facts and models with reference to pink salmon (*O. gorbuscha*). Canadian Journal of Fisheries and Aquatic Sciences 51(Supplement 1):223–232.

Appendix 1

Baseline Sampling Summary

Samples for this project were obtained from two sources: archived collections and new collections, spanning the years 1985–2003. Fifty-two sites, 34 in Yukon Territory, Canada and 18 in Alaska, U.S. were identified as potential Chinook salmon baseline collections within the Yukon River drainage (Table 1). At the conclusion of the 2003 field season (end of allotted sampling time), 19 combined Canada/U.S. collections, covering key locations in the drainage, were of sufficient sample size (Table 1). Additional Chinook salmon baseline samples were collected in the 2004 but not incorporated into the analysis due to time constraints. These samples and additional future collections will be added to enhance the baseline at a later date.

The vast size and habitat diversity of the Yukon River basin along with limited knowledge of Chinook salmon ecology in the river drainage provided several logistical challenges during sampling. In addition, while run-timing and abundance are well documented for some tributaries (Barton 1984; Johnson et al. 2005a, 2005b; Osborne et al. 2002; Eiler et al. 2004), for other systems they are virtually unknown. However, generally stock abundance is relatively low, and 50% of the entire Yukon River Chinook salmon migration passes through the river mouth in a 7-day period with a mean date of June 20. The compressed nature of the run, small stock sizes, and remoteness of the Yukon River required multiple years and approaches to successfully collect samples.

Methods

The following approaches were used to obtain new adult Chinook salmon collections:

1) A crew was dropped off by helicopter or other transportation method on a tributary to float down the river to a pickup location collecting samples along the way. This was referred to as a drop-and-drift technique. When the crew located a spawning aggregate, they would place block nets above and below the reach and work fish towards one of the nets (50 ft in length, 8 feet in depth with 4.5 inch stretch mesh). The fish were captured by hand or with a large dip net and a small pelvic fin clip was taken and placed in a nalgene vial with 90% ETOH; the sampled fish was then released. In streams that were too deep for wading or too turbid, gill nets (from 50 to 100 feet in length, 11 feet in depth, 6.75 inch mesh with hang ratios of 2.4:1, 2.5:1 and 3:1) were set from watercraft. These nets were monitored and live fish released as soon as possible.

2) A crew was shuttled by helicopter between sampling sites. This technique, with favorable water conditions, allowed concentrations of Chinook salmon to be located by aerial observation prior to committing a sampling crew. This capability was especially valuable for tributaries where the locations of fish were unknown or uncertain. Gill nets were used as set nets on larger systems or as seines in the same manner as stated above.

3) Established weirs, counting towers, and sonar sites were used for collecting samples. For weirs, fish were sampled from live boxes. For counting towers and sonar sites, fish were sampled by net or by carcass survey.

4) Crews collected samples from carcasses while floating or walking along a river. Chinook salmon carcasses lying on the bottom of the river were snagged or forked out with a gigging spear.

Results

Most of the successful U.S. collections occurred at weirs, counting towers, or sonar sites. Collections at weirs, which were located 1.5–322 km upstream of their confluence with either the Yukon or Koyukuk Rivers, occurred on the following rivers: Andreafsky (43 km upstream), Gisasa (4 km), Henshaw (1.5 km), Tozitna (80 km), and Beaver (322 km). Collections at counting towers and sonar sites, located 5–73 km upstream, occurred through seine netting adult fish to collect age, sex and length (ASL) data or carcass surveys on the following rivers: Anvik (76 km), Chena (73 km), Salcha (5 km), and Chandalar (22 km). Additional samples from the Chandalar were collected by a subsistence fisherman from the village of Venetie (60 km) using a gill net.

The Canadian collections occurred mostly through accessing spawning grounds by road and/or boat and catching adult fish with seine or gill nets. The Chiandindu and Blind collections occurred at weirs that were approximately 3 km upstream of their confluences. The Whitehorse collection occurred at the hatchery when fish returned to spawn.

Discussion

Methods that tried to time the arrival of fish with the arrival of the sampling crew were generally not very successful. Run time variability was too great, successful sampling required a large commitment of time, thus, the reason for the success of the weirs, counting towers, and sonar sites.

The drop-and-drift technique approach required several days to sample a reach, though it permitted thorough coverage of that section of river. The sampling team was usually committed to the duration of the float trip once it was initiated even if no fish were present or high turbid waters conditions occurred from rains.

Shuttling crews by helicopter was somewhat limited by the operating range of the helicopter and weather conditions. Also, helicopter availability can be limited this time of year due to fire contracts. Extending the range required the additional logistics of caching aviation fuel at strategic locations. A number of sampling days were lost due to bad weather grounding the helicopter and crew. This grounding disrupted plans to sample multiple tributaries in succession during the relatively narrow window of spawning. However, the helicopter gave us the ability to land and pickup crews in sections of river not accessible by any other means.

Weirs, counting towers, and sonar sites were effective, and the genetic data indicated that each sample represented a single population. The established infrastructure, present throughout the run, assured successful sampling and reduced expenditures.

Carcass collections posed other problems. The sample had to be taken from a recently deceased fish otherwise the DNA was not suitable for lab analysis. Another unexpected drawback was that carcasses were sometimes difficult to find even on rivers with substantial runs (e.g. Chandalar and Sheenjek Rivers). We believe factors such as woody debris, turbidity, stream flows, scavengers, etc. influenced the success of this method.

Table 1A. Locations, years, and numbers of adult fish sampled.

Sample Location	Year Sampled	*N* Per Year	Total *N*
Canada			
Locations Analyzed			
Big Salmon	1987, 1997	76, 40	116
Blind	1997, 2003	1, 138	139
Chandindu	1998, 2001, 2003	123, 158, 85	366
Little Salmon	1987, 1997	20, 80	100
Mayo	1992, 1997, 2003	135, 32, 38	205
Pelly	1996, 1997	39, 113	152
Stewart	1997	99	99
Takhini	1997, 2002, 2003	63, 67, 38	168
Tatchun	1987, 1996, 1997, 2002, 2003	27, 200, 58, 36, 48	369
Whitehorse	1985, 1987, 1997	39, 89, 114	242
Total			1956
Locations Not Analyzed			
Bearfeed	1987	2	2
Big Campbell	2003	7	7
Big Kalzas	2003	33	33
Crow	2002	1	1
Earn	2003	36	36
Fifty-mile	2003	4	4
Fishing Branch	1982, 2002	6, 1	7
Gladys	2003	4	4
Glenlyon	2003	24	24
Hoole	2003	2	2
Janet	2003	7	7
Klondike	1995, 2001, 2003	6, 10, 70	86
Little Kalzas	2003	24	24
Mainstem/Minto	1987, 1997, 2002	8, 11, 19	28
Mica	1997	2	2
Michie	1994	47	47
Morley	1997, 2002, 2003	9, 8, 12	29
Nisutlin	1987, 1997	17, 39	56
Nordenskiold	2003	106	106
Ollie Lakes	2002	4	4
Porcupine	2002	12	12
Primrose	2003	3	3
Tincup	2003	32	32

Continued on next page

Table A1. *Continued.*

Sample Location	Year Sampled	*N* Per Year	Total *N*
Wolf	1995, 2003	50, 4	54
Total			610
	U.S.		
Locations Analyzed			
Andreafsky	2002, 2003	28, 209	237
Anvik	2002, 2003	75, 38	113
Beaver	1997	96	96
Chandalar	2002, 2003, 2004	4, 113, 61	178
Chena	2001	200	200
Gisasa	2001, 2004	368, 200	568
Henshaw	2001, 2004	150, 250	400
Salcha	2003, 2004	55, 100	155
Tozitna	2002, 2003, 2004	200, 250, 250	700
Total			2646
Locations Not Analyzed			
Archuelinguk	2002, 2003	28, 50	78
Atcheulinguk	2002	0	0
Barton	2002	0	0
Bonsila/Stuyahok	2002	0	0
Chatanika	2001	19	19
Kateel	2002	19	19
South Fork Koyukuk	2003	56	56
Melozitna	2003, 2004	28, 69	97
Sheenjek	2002, 2004	3, 11	14
Total			283

Appendix 1 References

Barton, L. 1984. A catalog of Yukon River salmon spawning escapement surveys. Technical Data Report No. 121, Alaska Department Fish and Game.

Eiler, J.H., T.R. Spencer, J.J., Pella, M.M. Masuda, and R.R. Holder. 2004. Distrubution and movement patterns of Chinook salmon returning to the Yukon River basin in 2000–2002. NOAA Technical Memorandum NMFS-AFSC-148, U.S. Department of Commerce.

Johnson, J., E. Weiss, and S. Maclean. 2005a. Catalog of waters important for spawning, rearing, or migration of anadromous fishes—Interior Region, effective January 15, 2005. Special Publication No. 04-04, Alaska Department of Fish and Game.

Johnson, J., E. Weiss, and S. Maclean. 2005b. Catalog of waters important for spawning, rearing, or migration of anadromous fishes—Western Region, effective January 15, 2005. Special Publication No. 04-08, Alaska Department of Fish and Game.

Osborne, B.M., and J.L. Melegari. 2002. Use of split-beam sonar for enumeration of Chandalar River fall chum salmon, 2000. Alaska Fisheries Technical Report 61, U.S. Fish and Wildlife Service.

Appendix 2

Table 2A. The number of individuals (*N*) per sample location successfully amplified at each microsatellite locus.

	***N* Per Locus**											
Sample Location	***Oke 2***	***Ots 3.1***	***Oki 10***	***Otsg 3***	***Oke 4***	***One 102***	***One 107***	***One 9***	***Ots 1***	***Ots 100***	***Ots 212***	***μ 73***
Andreafsky	204	205	204	197	204	191	194	203	198	200	202	193
Anvik	110	108	88	79	50	10	31	74	34	55	38	57
Gisasa	168	172	161	170	180	163	169	177	165	173	176	175
Henshaw	147	147	147	147	128	128	142	146	143	145	146	141
Tozitna	196	194	192	196	373	254	359	373	347	331	324	343
Chena	162	158	162	168	156	183	89	185	148	180	168	115
Salcha	16	20	38	30	49	53	53	54	50	52	38	54
Beaver	89	90	87	90	90	95	93	95	95	95	94	96
Chandalar	87	97	99	96	102	106	103	107	106	107	101	108
Chandindu	159	165	147	50	222	223	241	236	239	232	239	186
Stewart	84	77	36	90	90	97	85	90	98	93	99	93
Mayo	133	147	148	30	58	59	61	60	62	62	54	59
Pelly	116	125	114	125	104	109	110	105	110	98	107	100
Blind	128	126	103	67	114	120	129	127	129	125	132	127
Tatchun	119	122	104	115	132	118	110	135	132	123	120	100
Little Salmon	63	77	72	81	56	55	43	60	71	33	56	59
Big Salmon	75	90	86	92	102	97	92	103	100	101	102	105
Takhini	154	147	105	165	156	154	137	157	115	133	159	140
Whitehorse	138	146	106	118	189	190	196	186	188	188	190	187

Continued on next page

Table 2A. *Continued.*

Sample Location	*N* Per Locus									
	Ots 107	***Ots 101***	***Ots 104***	***Ogo 2***	***Ogo 4***	***Oki 100***	***Omy 325***	***Ots 2***	***Ots 9***	***Ssa 197***
Andreafsky	187	190	183	186	162	183	167	150	158	163
Anvik	57	62	58	49	71	67	73	57	49	52
Gisasa	174	174	170	165	167	177	167	155	153	176
Henshaw	138	142	140	141	142	123	135	145	140	140
Tozitna	401	417	398	417	426	415	423	424	414	407
Chena	135	159	142	151	149	154	160	154	146	137
Salcha	34	38	36	43	41	28	47	47	47	38
Beaver	88	88	84	86	88	89	89	90	90	89
Chandalar	92	89	86	95	91	94	102	97	97	88
Chandindu	340	335	337	323	333	318	338	329	327	336
Stewart	97	107	100	97	94	88	97	98	98	106
Mayo	184	178	178	160	169	147	177	161	158	183
Pelly	142	148	146	140	141	132	139	134	134	149
Blind	95	120	97	118	111	105	112	115	114	107
Tatchun	351	351	355	335	328	326	337	324	330	355
Little Salmon	88	85	88	82	84	79	86	84	83	87
Big Salmon	106	108	95	104	106	102	105	103	103	107
Takhini	161	162	142	163	164	158	161	157	157	155
Whitehorse	228	226	224	219	226	205	224	221	227	233