

22nd Pink and Chum workshop 2005, Ketchikan, AK

Building a collaborative chum salmon microsatellite baseline for the Yukon River genetic stock id.

Candy, J.R.[†], Flannery, B.G.[‡], and Beacham, T.D.[†]

[†]**Molecular Genetics Laboratory,
Department of Fisheries and Oceans,
Pacific Biological Station,
Nanaimo, B.C.
V9T 6N7**

[‡]**Conservation Genetics Laboratory,
U.S. Fish and Wildlife Service,
1011 E Tudor Rd, Anchorage,
AK 99503**

Abstract

A baseline of 17 Yukon River chum populations common to both labs was screened using microsatellite genetic markers in a collaborative project between the Conservation Genetics Laboratory, U.S. Fish and Wildlife Service (USFWS) and the Molecular Genetics Laboratory, Canadian Department of Fisheries and Oceans Canada (CDFO). Each laboratory developed and screened a suite of microsatellite loci for the same baseline fish samples as well as analyzing 1200 mixed-stock fish from the lower river (Pilot Station). A combined marker set of 22 loci was used to define 9 reporting regions: Lower River summer, Tanana summer, Koyukuk summer, Tanana falls, Upper Alaska, Canadian Porcupine, White River, Teslin River and Canadian mainstem. The degree these reporting regions are identifiable in the mixtures was tested with 100% single population simulations. Correct mean allocations from single population simulations achieved the 90% threshold for 7 of the 9 regional groups, analysis of accuracy and precision indicate that ~300 alleles are required to exceed an average correct allocation greater than 90% for all reporting regions with a standard deviation of less than 3%. The Pilot Station mixed-stock sample indicates that both CDFO and USFWS baselines provide very similar regional estimates of stock composition over the 6 sampling intervals. Summer run chum contributed ~40% of the first sampling interval. Following sampling intervals showed an increase in border Canada and border US fall run populations followed by an increase in Tanana Fall populations by the end of August.

Introduction

Yukon River chum salmon undertake the longest freshwater migration of chum salmon in North America, some individuals traveling over 2,700 kilometers upstream (Beacham et al. 1988). Effective management of fisheries with this major drainage requires the knowledge of exploitation rates on specific population or population groups. Accurate and timely estimates of stock composition from the lower river are required to meet target

allocations and escapement goals for both Canada and the US under the Yukon River Salmon Agreement (YRSA). For Yukon chum, lower river migration timing and abundance estimates are determined at Pilot Station using acoustic counters and a gillnet drift test fishery (Maxwell and Huttunen 1998).

Previous work using allozymes on Yukon River chum (Beacham et al. 1988, Wilmot et al. 1992, ADFG unpub. data) indicate five regional groups can be identified; these are Lower Summers, Tanana Fall, Boarder US/Canada, White River and Teslin River. However, allozyme analysis failed to discriminate the border populations well enough for management needs. Microsatellite variation in a number of species has been useful to provide increased level of stock discrimination for salmonid populations passing through lower river mixed stock fisheries (Beacham et al. 1999, Beacham et al. 2000, Parken et al. in prep). In this paper we look at the discrimination power of a combined microsatellite dataset from the USFWS and CDFO labs for Yukon River chum salmon and compare results from two independent baselines for the same lower river mixed stock fishery samples.

Methods

Microsatellite loci

The CDFO lab screened baseline populations with 13 microsatellite loci with 11 to 87 alleles per loci and the USFWS lab screened for 11 microsatellite loci with between 2 to 36 alleles per loci (Table 1). Two loci *Oke3* and *Ots3* were run by both labs resulting in a combined maker set of 22 loci.

DNA was extracted from the samples either as described by Withler et al. (2000) or by using proteinase K with the Dneasy™ DNA isolation kit (Quiagen Inc. Valencia, CA). At the CDFO laboratory, PCR products from 13 microsatellite loci: *Ots3* (Banks et al. 1999), *Oke3* (Buchholz et al. 1998), *Oki2* (Smith et al. 1998), *Oki100* (Miller et al. unpub), *Omy 1011* (Bentzen, unpub.), *One101*, *One102*, *One103*, *One104*, *One111*, and *One114* (Olsen et al. 2000), *Ssa419* (Cairney et al. 2000), and *OtsG68* (Williamson et al. 2002) were size fractionated on denaturing polyacrylamide gels with the ABI 377 automated DNA sequencer. Allele sizes were determined with Genescan 3.1 and Genotyper 2.5 software (PE Biosystems, Foster City, CA).

The USFWS laboratory, analyzed PCR products from 11 microsatellite loci: *Oke3*, *Oke4*, *Oke8*, *Oke11* (Buchholz et al. 1998), *Oki1*, *Oki23.1* (Smith et al. 1988.), *Ots2.1*, *Ots3.1* (Banks, et al. 1999), and *Ots103* (Small et al. 1998). One µl of PCR product was electrophoresed and visualized on a denaturing 6% polyacrylamide gel using a Li-Cor IR2® DNA scanner. The sizes of bands were estimated and scored by the computer program Saga GT version 3.1 (Li-Cor, Lincoln, NE). Li-Cor size standards (50bp – 350bp) and allele ladders were run every sixteen lanes to ensure consistency of allele scores. All scores were verified by eye. Alleles were scored by two independent researchers, with any discrepancies being resolved by re-running the samples in question and repeating the double scoring process until scores match.

Baseline populations

The USFWS lab analysed samples from 18 populations, while the CDFO lab analysed samples from 23 populations providing a baseline of 17 populations in common between the two labs (Table 2). Due to poor sample quality and low DNA amplification success the Big Salt sample was left out of the analysis. The reporting regions differed slightly between the two labs. CDFO defined reporting regions on a finer scale than the USFWS. This is due to higher level of discrimination from the CDFO loci and greater number of populations present in the CDFO baseline. Allele frequencies for each population were derived from combining annual sampling using methods of Waples (1990). Sampling locations for the 23 populations can be seen in Figure 1.

Analysis of the combined USFWS/CDFO 17 population baseline include visualization of genetic structure using cord distance (Cavalli-Sforza and Edwards, 1967) and neighbour joining tree from the program PHYLIP (Felsenstein, 1993). The ability to discriminate to reporting groups was assessed using 100% single population simulations using SPAM version 3.7 (Debevec et al. 2000). The program was run with the Rannala and Mountain (1997) data augmentation routine to avoid having fish in the mixture with alleles not observed in the baseline. The number of alleles required for accurate estimation to reporting group was also tested using 100% single population mixtures where increasing cumulative number of alleles varied for each run of the estimation.

Pilot Station samples

Every chum salmon caught in Pilot Station sonar test fisheries from July 19 –Aug 31, 2004 was sampled (Figure 1). For management purposes fish entering the river after July 19 are assumed to be “fall run” fish (Bue et al. 2004). The Pilot Station samples were stratified by run pulse; 200 fish per strata were analyzed. The periods were designated “build-up” (July 19- Aug 2), “pulse 1” (Aug 3 – 9), “pulse 2” (Aug 10 – 15), “pulse 3” (Aug 16- 21), “pulse 4” (Aug 22 – 26), and “pulse 5” (Aug 27-31). The tissue samples from the Pilot Station test fisheries were sub-sampled, proportional to the daily sonar passage estimate, at the USFWS lab and sent to the CDFO lab. The samples were genotyped by each lab using the respective suites of loci used in the two baselines. The USFWS lab provided stock composition estimates for each sample strata using Bayes mixture model (Pella and Masuda, 2001), and an 18 population baseline. The CDFO lab analyzed the same mixture samples using CBayes (a C++ rewrite of the Bayes program by CDFO), and a 23 population baseline. Results were reported by USFWS reporting regions.

Results and Discussion

The USFWS lab identified 6 regional groups are: Lower River Summer, Middle Summer, Tanana fall, Border USA, Border Canada and Upper Canada (Table 2). Nine reporting regions were identified using the combined baseline dataset. These are Lower River summer, Tanana summer, Koyukuk summer, Tanana falls, Upper Alaska, Canadian Porcupine, White River, Teslin River, and Canadian mainstem. Figure 2 shows the unrooted dendrogram of genetic distances between the 17 populations using the combined dataset of 22 loci, and an overlay of the 9 regional groups. The lower left of

the dendrogram consists of populations geographically located in the lower river and the upper right of the dendrogram consists of upper river fall run populations. Figure 1 shows the location of these sample sites and geographic relationships among these regional groupings.

Mean allocation from simulated mixtures composed of 100% single reporting regions are thought to perform well if the mean contribution is greater than 90%. Reporting regions that had correct allocations greater than 90% were Lower River summer, Tanana summer, Tanana falls, Upper Alaska, Canadian Porcupine, White River, Teslin River, (Table 3). Two regional group contributions were slightly less than 90%; Koyukuk summer at 89.7% with allocation lost to Lower summer, and Canadian mainstem (84.2%) with allocation lost to Upper Alaska. Inclusion of new baseline populations and increasing sample sizes of existing populations should allow better characterization of these two under performing reporting groups.

The relationship between number of alleles present in the analysis and accuracy to reporting region is non-linear; estimates approach 100% asymptotically as the allele number increases (Figure 3). Using just the USFWS loci with 106 alleles average accuracy to regional group was 87% with a standard deviation of 4%. By using the additional 400 CDFO loci alleles the accuracy to regional group increased 92% with a standard deviation of 1.7 %. It took approximately 300 alleles to exceed the 90% threshold thought to be required for accurate stock identification in mixture analysis.

Analysis of the Pilot Station mixture sample indicated that both the USFWS baseline and CDFO baseline produced vary similar results (Figure 4). Analyses of samples indicated that summer chum salmon are still a significant contributor during the early portion of the fall management season, comprising ~40 % of the build-up stratum, but dropped precipitously in subsequent strata. Border Canada, Border US and Upper Canada reporting groups showed an increase in the early August, presumably as headwater populations moved through the lower river. Toward the end of August the Tanana Fall run fish became the single biggest contributor.

Figure 1 – A map of sampling sites and regional groups of Yukon River chum salmon determined from 22 microsatellite loci using the USFWS/CDFO baseline.

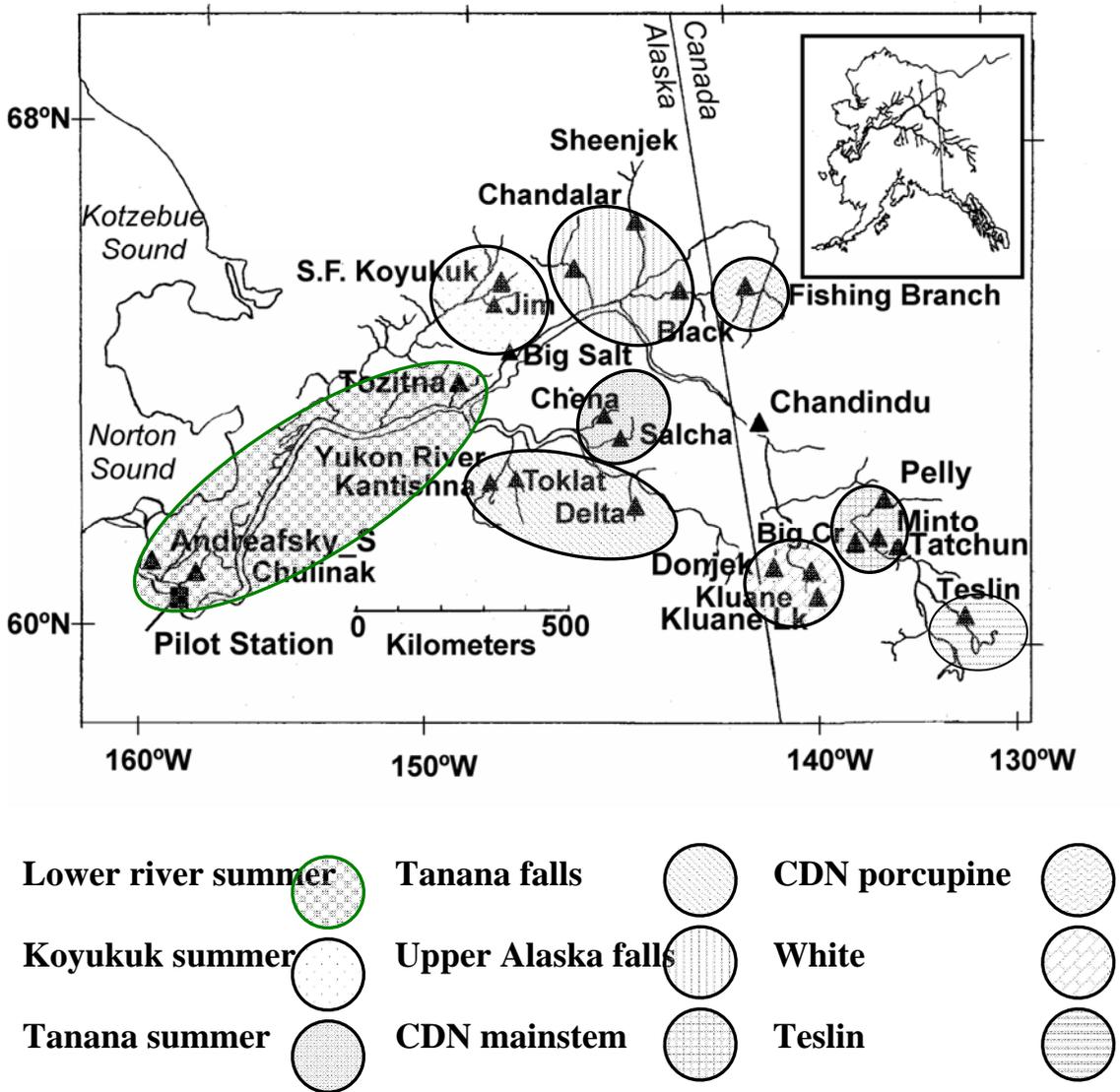


Figure 2 – Unrooted dendrogram of genetic distances (Cavalli-Sforza and Edwards 1967) for 17 Yukon River chum salmon populations using 22 microsatellite loci from the combined USFWS/CDFO baseline.

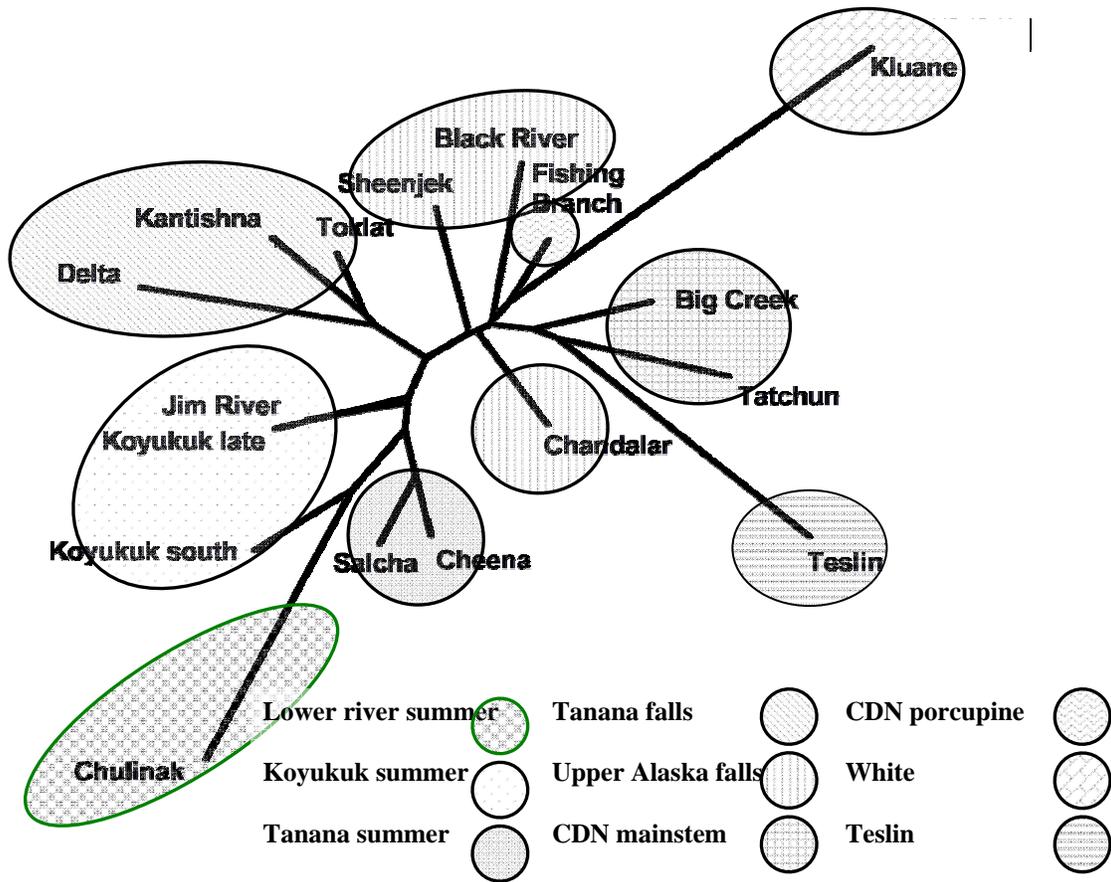


Figure 3 –Average regional estimates (9 reporting regions) of accuracy and precision plotted against cumulative allele counts for USFWS (diagonal slash) and CDFO (stippled) loci. Estimates to region assuming mixture with 100% single population, where more than one population exists in region, average values for all run used for the regional estimate.

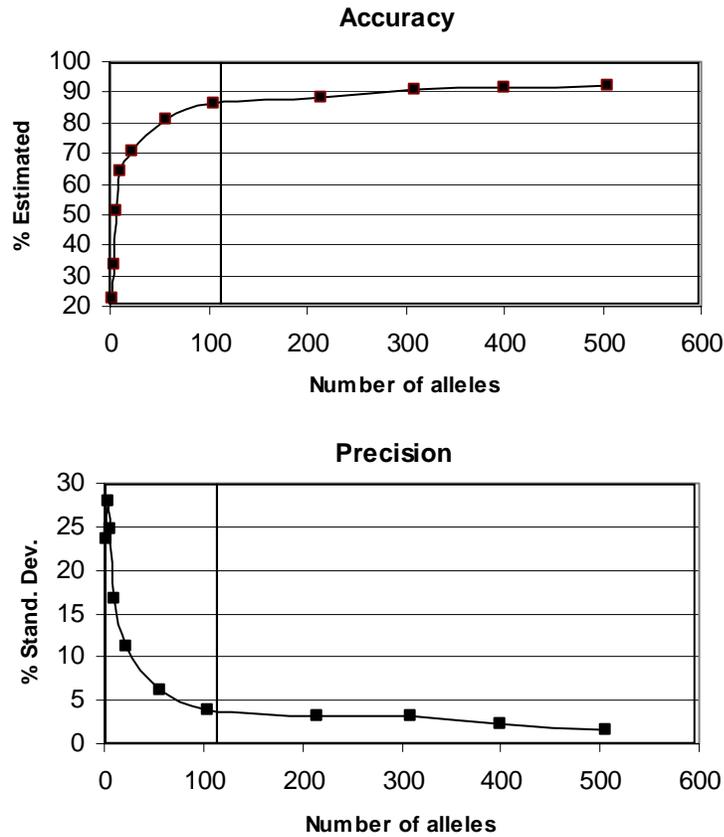


Figure 4 – Comparison of Pilot Station mixture analysis on the same fish using the USFWS (back slash) and the CDFO (stippled) baseline.

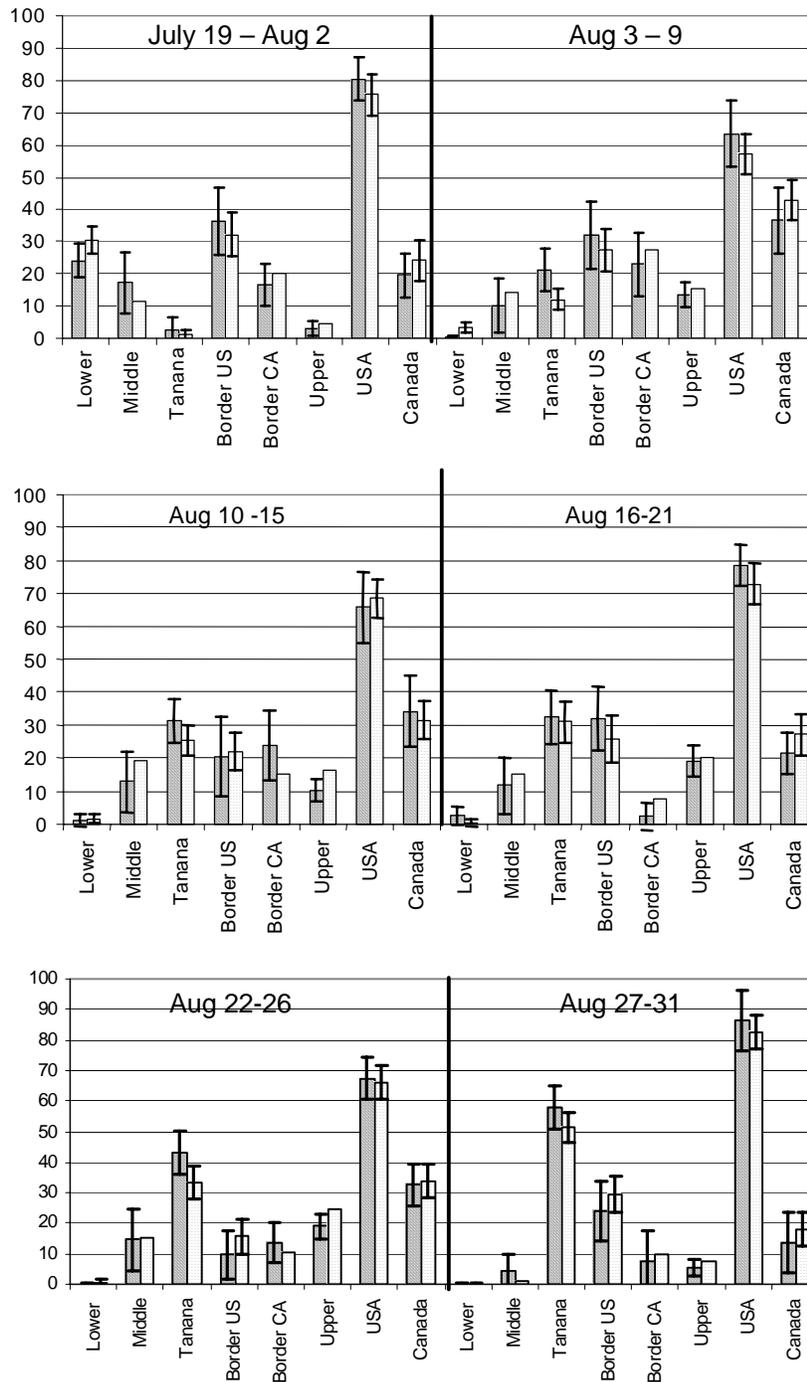


Table 1. – List of loci and number of alleles found in Yukon River chum, where A is the USFWS lab loci and B is the CDFO lab loci.

A		B	
Locus	# alleles	Locus	# alleles
Oke3	8	Oke3	11
Oke4	5	Oki100	24
Oke8	2	Oki2	19
Oke11	5	Omy1011	32
Oki1L	3	One101	39
Oki1U	18	One102	37
Oki23.1	4	One103	31
Ots2.1L	3	One104	28
Ots2.1U	6	One111	87
Ots 3.1	16	One114	35
Ots103	36	Ots3	24
		Otsg68	39
		Ssa419	17

Table 2 – List of Yukon Chum populations, sampling year, and number of fish surveyed by each lab for baseline, where A is the USFWS baseline and B is the CDFO baseline.

A

Population/Regional Group	N
Lower Summer	
Chulinak 1989	100
Middle Summer	
Chena 1992, 1994	186
Salcha 1994, 2001	185
Jim 2002	160
S F Koyukuk Early 1996	100
S F Koyukuk Late 1996	100
Big Salt 2001	71
Tanana Fall	
Delta 1990	80
Toklat 1994	200
Kantishna 2001	161
Border USA	
Chandalar 1989, 2001	250
Sheenjek 1988, 1989	154
Black 1995	112
Border Canada	
Fishing Branch 1989, 1992	150
Tatchun 1992	100
Big Creek 1992	100
Upper Canada	
Kluane 1992, 2001	250
Teslin 1992	100

B

Population/Regional Group	N
Lower Summer	
Chulinak 1989	100
Andreafsky 1987	61
Tozitna 2002	200
Koyukuk Summer	
Jim 2002	159
S F Koyukuk Early 1996	100
S F Koyukuk Late 1996	100
Tanana Summer	
Chena 1992, 1994	186
Salcha 1994, 2001	185
Tanana Fall	
Delta 1990	80
Toklat 1994	200
Kantishna 2001	161
Upper Alaska	
Big Salt 2001	71
Chandalar 2001	200
Sheenjek 1987, 1988, 1989	263
Black 1995	95
Canadian Porcupine	
Fishing Branch 1987, 1994, 1997	331
Chandindu River	
Chandindu	55
White	
Kluane 1987, 1992, 2001	462
Teslin 1992, 2001	143
Donjek 1994	72
Teslin	
Teslin 1992, 2002	143
Mainstem Yukon	
Tatchun 1987	75
Big Creek 1995	100
Pelly 1993	84
Minto 1989, 2002	166

Table 3. Region estimates assuming 100% single population in mixture (N=400) using 22 loci from the combined USFWS/CDFO baseline. *indicates less than 90% allocation required for accurate estimate to region.

	Estimate	Std. Dev.
Lower river summer	93.3	(1.5)
Koyukuk summer	89.7*	(2.1)
Tanana summer	93.7	(1.7)
Tanana fall	93.2	(1.6)
Upper Alaska fall	91.4	(2.0)
Canadian mainstem	84.2*	(2.6)
White	98.7	(0.6)
Canadian Porcupine	93.5	(1.8)
Teslin	94.2	(1.3)

References

- 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 (*Oncorhynchus tshawytscha*). *Journal of Heredity* 90:281-288.
- Beacham, T.D., C.B. Murry, and R.E. Withler. 1988. Age, morphology, developmental biology, and biochemical genetic variation of Yukon River fall chum salmon, *Oncorhynchus keta*, and comparisons with British Columbia populations. *Fishery Bulletin* 86:663-674.
- Beacham, T.D. and C.C. Wood 1999. Application of microsatellite DNA variation to estimation of stock composition and escapement 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., C.C. Wood, R.E. Withler, K.D. Le, and K.M. Miller. 2000 Application of microsatellite DNA variation to estimation of stock composition and escapement of Skeena River sockeye salmon (*Oncorhynchus nerka*). *North Pacific Anadromous Fish Commission Bulletin* 2:263-276.
- Buchholz, 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.
- Bue, F.J., B.M. Borba, D.J. Bergstrom. 2004. Yukon River fall chum salmon stock status and action plan. Regional Information Report No. 3A04-05. Alaska Department of Fish and Game, Division of Commercial Fisheries, 333 Raspberry Road, Anchorage, AK 99518
- Cairney, M., Taggart, J. B., and Hoyheim, B. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. *Molecular Ecology* 9:2175-2178.
- Cavalli-Sforza, L.L. and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 21:550-570.
- Debevec, E. M., R.B. Gates, M. Masuda, J. Pella, J. Reynolds, and L.W. Seeb. 2000 SPAM (Version 3.2): Statistics Program for Analyzing Mixtures. *Journal of Heredity* 91:509-510.
- Felsenstein, J. 1993. PHYLIP (phylogeny inference package) version 3.5c edition. Department of Genetics, SK-50, University of Washington, Seattle.

- Maxwell, S.L. and D.C. Huttunen. 1998. Yukon River sonar project report 1997. Regional Information Report No. 3A98-12. Alaska Department of Fish and Game, Division of Commercial Fisheries, 333 Raspberry Road, Anchorage, AK 99518.
- 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 Atlantic salmon. *Molecular Ecology* 9: 2155-2234.
- Parken, C.K., J.R. Candy, J.R. Irvine, and T.D. Beacham. (in prep) Application of individual-based approach to estimate the relative abundance and timing of populations of chinook salmon moving through a mixed stock freshwater salmon fishery.
- Pella, J. and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. *Fishery Bulletin* 99:151-167.
- Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Science USA* 94:9197-9201.
- Small, M.P., T.D. Beacham, R.E. Withler, and R.J. Nelson. 1998. Discriminating coho salmon (*Oncorhynchus kisutch*) populations 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* 11:1614-1616
- Withler, R.E., K.D. Le, R.J. Nelson, K.M. Miller, and T.D. Beacham. 2000. Intact genetic structure and high levels of genetic diversity in bottlenecked sockeye salmon, *Oncorhynchus nerka* populations of the Fraser River, British Columbia, Canada. *Canadian Journal of Fisheries and Aquatic Sciences* 57:1985-1998.
- Waples, R. S. 1990. Temporal changes of allele frequency in Pacific salmon populations: implications for mixed-stock fishery analysis. *Canadian Journal of Fisheries and Aquatic Sciences* 47:968-976.
- Williamson, K. S., J.F. Cordes, and B.P. May. 2002. Characterization of microsatellite loci in chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. *Molecular Ecology Notes* 2:17-19.