2010 Alaska Sustainable Salmon Fund

Yukon River Chum Salmon Baseline Augmentation

Final Report for Study 45753

Blair Flannery and Randal Loges

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

September 2011

Abstract

A total of 4,055 Yukon River chum salmon were genotyped at 13 microsatellite loci and added to the existing baseline. Mixed-stock simulation results generally improved. For some reporting groups, 100% simulation accuracy decreased slightly, which was likely due to the new samples showing closer genetic relationships among regions than previously detected. However, an expanded baseline should improve real fishery composition estimation because the genetic relationships are more accurately depicted and estimation accuracy improves with the inclusion of potentially contributing stocks. This is supported by the increased accuracy of the expanded baseline in multi-region simulations, which imitate actual fisheries, and thus are better indicators of baseline performance than the 100% simulations.

Introduction

Management of the Yukon River salmon fishery is complex because of the difficulty of determining stock specific abundance and timing, overlapping multi-species salmon runs, and the immense size of the drainage. Salmon fisheries within the Yukon River may harvest stocks that are up to several weeks and hundreds of miles from their spawning grounds. Because the Yukon River fisheries are largely mixed-stock fisheries, some tributary populations may be under- or over-exploited in relation to abundance.

Determining stock structure and the relative contribution of stocks to harvests are essential for effective management (Larkin 1981). This is a difficult task, though greatly simplified through the use of genetic mixed-stock analysis (MSA; Cadrin et al. 2005). Genetic MSA of Yukon River chum salmon began in 2004. Consistent with the Yukon River Salmon Agreement, fall chum salmon, a shared resource between the U.S. and Canada, were the focus of the MSA. However, the instability of the summer chum salmon stock composition over the past decade has recently prompted MSA of summer chum salmon (Flannery et al. 2009).

Genetic baselines constantly evolve to increase statistical power in order to provide more accurate and precise stock composition estimates and address new management questions (e.g. summer chum salmon). Bias and precision of stock composition estimates are directly related to baseline representation and knowledge of stock genotypic frequencies. Bias is reduced and precision increased when stock genotypic frequencies are known rather than estimated, which is the equivalent of very large samples sizes (Pella and Milner 1987). Significant improvement in MSA accuracy has been realized with previous baseline augmentations. For example, MSA simulation accuracy improved from 79% to 91% for Fishing Branch when its sample size was increased from 96 to 481 (Flannery et al. 2007). Because stocks are more likely to be subdivided geographically than temporally, it is recommended that sampling of stocks be repeated over time and that these samples be pooled in order to better describe genotypic frequencies (Waples 1990).

The current Yukon River chum salmon baseline is weighted (1.5:1) toward fall chum salmon due to the initial concern with this run (Table 1). In this study, we genetically assay over 4,000 samples at 13 microsatellite loci, nearly doubling the baseline size. This includes the addition of seven new stocks, of which five are summer, and increases in sample sizes for many critical

summer and fall chum salmon stocks (Table 2). These additions will bring balance to the baseline and may increase the accuracy and precision of stock composition estimates for Yukon River summer and fall chum salmon.

Methods

Sample collection and laboratory analysis

Tissue samples (*N* = 4,055) from adult chum salmon were collected from 26 spawning locations in the Yukon River (Table 2). Total genomic DNA was extracted from the fish tissue (~25mg) using proteinase K with the Dneasy™ DNA isolation kit (Quiagen Inc. Valencia, CA), quantified with fluorometry and diluted to a standard concentration. The following microsatellite loci were assayed for genetic variation: *Oke3* (Buchholz et al. 2001), *Oki1, Oki2* (Smith et al. 1998); *Oki100* (Miller unpublished); *Omy1011* (Spies et al. 2005); *One102, One103, One104,* and *One114* (Olsen et al. 2000); *Ssa419* (Cairney et al. 2000); *OtsG68* (Williamson et al. 2002); *Ots103* (Beacham et al. 1998). Polymerase Chain Reaction (PCR) DNA amplification was done in 10 µl volumes; general conditions were: 2.5 mM $MgCl_2$, 1X PCR buffer (20 mM Tris-HCl pH 8.0, 50 mM KCl), 200 µM of each dNTP, 0.40µM fluorescently labeled forward primer, 0.40 µM unlabeled reverse primer, 0.008 units *Taq* polymerase, and 1 µl of DNA (30ng/µl). An MJResearch DNA Engine® thermal cycler was used to perform PCR. Standard thermal cycling conditions were: initial denaturation cycle of 94°C for 3 min, followed by 94°C for 1 min, 50-62°C for 1 min (locus-specific annealing temperature), 72°C for 1 min, with a final single cycle of 72°C for 10 min. One µl of PCR product was electrophoresed and visualized with the Applied Biosystems 3730 Genetic Analyzer utilizing a polymer denaturing capillary system. The sizes of bands were estimated and scored by the computer program GeneMapper® version 4.0. Applied Biosystems GeneScan™-600 LIZ® size standard, 20-600 bases, were loaded in all lanes as an internal lane standard. All scores were verified manually. Alleles were scored by two independent researchers, with any discrepancies being resolved by replicating the analysis for the samples in question and repeating the double scoring process until scores match.

Statistical analysis

The baseline data were checked for duplicated genotypes using the program MICROSATELLITE TOOLKIT (Park 2001), with any duplicates removed. Where multiple tests of the same hypothesis were performed, a sequential Bonferroni method was used to maintain the overall alpha at 0.05 (Rice 1989). The stocks and loci were analyzed for Hardy-Weinberg and gametic phase equilibrium using the program FSTAT 2.9.3 (Goudet 1995) and GENETIX 4.05 (Belkhir et al. 2004), respectively. Stocks with multiple collection years were checked for temporal variation of allele frequencies using GENEPOP 4.1 (Raymond and Rousset 1995). The program PHYLIP 3.57 (Felsenstein 1993) was used to calculate a stock pairwise chord distance (Cavalli-Sforza and Edwards 1967) matrix from allele frequencies and produce a neighbor-joining (Saitou and Nei 1987) dendrogram.

Simulation analyses were used to evaluate the accuracy and precision of stock composition estimates derived from the baseline. The fishery management regions established by the Joint Technical Committee (JTC; Figure 1) were used for apportioning mixtures. A hundred mixtures

(N = 400) were simulated for each region, and then, with bootstrap resampling of the baseline, the stock compositions of the mixtures were estimated by conditional maximum likelihood (CML) using SPAM 3.7 (Debevec et al. 2000). Region mixtures were comprised of equal proportions of the stocks within the region. A Bayesian method (Rannala and Mountain 1997) was used to estimate allele frequencies in order to prevent sampling zeros. Mean stock composition estimates and their standard deviations were tabulated. Simulations of multi-region mixtures were also conducted using likely proportions in a lower river fishery.

Results and Discussion

All stocks and loci were in Hardy-Weinberg and gametic phase equilibrium. Temporal variation of allele frequencies was observed among collection years within the Tozitna River and the Tanana River mainstem. Within the Tozitna River, temporal allele frequency variation existed between distant but not proximate collection years. Neighbor-joining analysis of CSE distances among the stocks revealed a spatial and temporal component to the distribution of genetic diversity (Figure 2). Summer stocks were subdivided into lower and middle river regions while fall stocks were subdivided into Tanana River, border area, White River, and Teslin River regions.

Mixed-stock simulation results generally improved. For some reporting groups, 100% simulation accuracy decreased slightly (Table 3), which is likely due to the clinal or gradient nature of genetic diversity in Yukon River chum salmon (Figure 2). In that the included samples revealed closer relationships between certain regions than previously detected in the baseline. While this may affect simulation accuracy, real fishery composition estimation should improve with a more descriptive baseline because a source of inherent error is reduced. The genetic relationships are more accurately depicted and estimation accuracy improves with the inclusion of potentially contributing stocks (Pella and Milner 1987). This is supported by the increased accuracy of the expanded baseline in multi-region simulations (Table 4), which imitate actual fisheries, and thus are better indicators of baseline performance than the 100% simulations.

The objective of this study was to survey additional samples to better describe the genetic diversity of Yukon River chum salmon in an attempt to improve accuracy and precision of stock composition estimates. This study fully met the objective by adding 4,055 samples and seven new stocks to the baseline and improving overall simulation accuracy of stock composition estimates. We anticipate that this expanded baseline will be beneficial for in-season stock composition analysis of Yukon River chum salmon.

Acknowledgements

The Alaska Sustainable Salmon Fund provided substantial financial support for this project. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Alaska Department of Fish and Game, National Oceanic and Atmospheric Administration, and U.S. Fish and Wildlife Service.

References

Beacham, T. D., L. Margolis, and R. J. Nelson. 1998. A comparison of methods of stock identification for sockeye salmon (*Oncorhynchus nerka*) in Barkley Sound, British Columbia. North Pacific Anadromous Fish Commission Bulletin 1:227-239.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Catch. 2004. GENETIX 4.05, software under Windows™ for the genetics of populations. University of Montpellier, Montpellier, France.

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

Cadrin, S. X, K. D. Friedland, J. R. Waldman. 2005. Stock Identification Methods: Applications in Fishery Science. Elsevier, Burlington, Massachusetts.

Cairney, M., J. B. Taggart, and B. Hoyheim. 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 A. W. F. Edwards. 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 3.5c. University of Washington, Seattle.

Flannery, B. G., T. D. Beacham, R. R. Holder, E. J. Kretschmer, and J. K. Wenburg. 2007. Stock structure and mixed-stock analysis of Yukon River chum salmon. U.S. Fish and Wildlife Service, Alaska Fisheries Technical Report 97, Anchorage. Available from: http://www.r7.fws.gov/fisheries/genetics/reports.htm

Flannery B. G., R. R. Holder, G. F. Maschmann, E. J. Kretschmer, and J. K. Wenburg. 2009. Application of mixed-stock analysis for Yukon River fall chum salmon, 2007. U.S. Fish and Wildlife Service, Office of Subsistence Management, Fisheries Resource Monitoring Program, Annual Report for Study 06-205, Anchorage, Alaska.

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

Larkin, P. A. 1981. A perspective on population genetics and salmon management. Canadian Journal of Fisheries and Aquatic Sciences 38:1469-1475.

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.

Park, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. Doctoral dissertation. University of Dublin, Ireland.

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.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceeding of the National Academy of Sciences 94:9197-9201.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2) population genetic software for exact tests and ecumenicism. Journal of Heredity 86: 248-249.

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

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

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.

Spies, I. B., D. J. Brasier, P., T. L. O'Reilly, T. R. Seamons, and P. Bentzen. 2005. Development and characterization of novel tetra-, tri-, and dinucleotide microsatellite markers in rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology Notes 5:278-281.

Waples, R.S. 1990. Temporal changes of allele frequency in Pacific salmon: 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.

Table 1. Current Yukon River chum salmon baseline identifying stock location, seasonal race, management region, country, sample collection years, and number of fish sampled (*N*) from 29 putative stocks.

Stock	Seasonal race	Management Region	Country	Year	*N*
Andreafsky	Summer	Lower	U.S.	1987, 2004	261
Chulinak	Summer	Lower	U.S.	1989	100
Anvik	Summer	Lower	U.S.	1988	100
California	Summer	Lower	U.S.	1997	50
Nulato	Summer	Lower	U.S.	2003	100
Gisasa	Summer	Lower	U.S.	2003	200
Henshaw	Summer	Lower	U.S.	2003	200
Jim	Summer	Lower	U.S.	2002	160
S.F. Koyukuk Early	Summer	Lower	U.S.	1996	100
S.F. Koyukuk Late	Summer	Lower	U.S.	1996	100
Melozitna	Summer	Lower	U.S.	2003	146
Tozitna	Summer	Lower	U.S.	2002	200
Chena	Summer	Tanana	U.S.	1992, 1994	186
Salcha	Summer	Tanana	U.S.	1994, 2001	185
Delta	Fall	Tanana	U.S.	1990	80
Kantishna	Fall	Tanana	U.S.	2001	161
Toklat	Fall	Tanana	U.S.	1990, 1994	250
Big Salt	Fall	U.S. Border	U.S.	2001	71
Chandalar	Fall	U.S. Border	U.S.	1989, 2001	338
Sheenjek	Fall	U.S. Border	U.S.	1987, 1988, 1989	263
Black	Fall	U.S. Border	U.S.	1995, 2001	112
Fishing Branch	Fall	Porcupine	Canada	1987, 1989, 1992, 1994, 1997	481
Big Creek	Fall	Mainstem	Canada	1992, 1995	200
Minto	Fall	Mainstem	Canada	1989, 2002	166
Pelly	Fall	Mainstem	Canada	1993	84
Tatchun	Fall	Mainstem	Canada	1987, 1992	175
Donjek	Fall	White	Canada	1994	72
Kluane	Fall	White	Canada	1987, 1992, 2001	462
Teslin	Fall	Teslin	Canada	1992, 2001	143

Table 2. Samples added to the current Yukon River chum salmon baseline with the following identified: stock location, seasonal race, management region, country, sample collection years, and number of fish sampled (*N*) for each stock. Stocks that are new to the baseline are highlighted grey.

Stock	Seasonal race	Management Region	Country	Year	*N*
Anvik-Beaver	Summer	Lower	U.S.	1992, 1993	200
Anvik-Yellow	Summer	Lower	U.S.	1992	100
Anvik-Swift	Summer	Lower	U.S.	1992, 1993	200
Anvik-Canyon	Summer	Lower	U.S.	1993	50
Rodo	Summer	Lower	U.S.	1989	78
Kaltag	Summer	Lower	U.S.	1992	100
Nulato	Summer	Lower	U.S.	1987, 1988, 1994	282
California	Summer	Lower	U.S.	1997	50
Tolstoi	Summer	Lower	U.S.	1997	100
Gisasa	Summer	Lower	U.S.	1988	100
Clear	Summer	Lower	U.S.	2002	204
Dakli	Summer	Lower	U.S.	1992	100
Henshaw	Summer	Lower	U.S.	1987, 2004	243
Jim	Summer	Lower	U.S.	1987	101
SF Koyukuk	Summer	Lower	U.S.	1990	75
Melozitna	Summer	Lower	U.S.	1989, 2003, 2004	104
Tozitna	Summer	Lower	U.S.	1989, 1992, 2003	371
Toklat Mainstem	Fall	Tanana	U.S.	1991, 1992, 1994	315
Toklat Sushana	Fall	Tanana	U.S.	1993, 1994	300
Delta	Fall	Tanana	U.S.	1991, 1992, 1994	300
Bluff Cabin	Fall	Tanana	U.S.	1992	100
Tanana Mainstem	Fall	Tanana	U.S.	1992, 1993	197
Sheenjek	Fall	U.S. Border	U.S.	1992, 1993	164
Chandalar	Fall	U.S. Border	U.S.	1988	73
Donjek	Fall	White	Canada	1993	53
Teslin	Fall	Teslin	Canada	1989	95

Table 3. Results of apportionments from mixed-stock analysis of simulated mixtures developed from region, run, and country-of-origin baseline data collected at 13 microsatellite loci. Each simulated mixture (*N*=400) was composed of equal proportions of stocks within the management group. The mixture and baseline were bootstrapped 100 times. Apportionment estimates and standard deviations are given for the current and expanded baselines.

	Current	Expanded
Summer	0.97 (0.01)	0.97 (0.01)
Lower	0.94 (0.02)	0.96 (0.02)
Tanana	0.91 (0.02)	0.86 (0.03)
Fall	0.95 (0.02)	0.96 (0.02)
Tanana	0.92 (0.02)	0.94 (0.02)
U.S. Border	0.80 (0.04)	0.83 (0.03)
Canada Border	0.91 (0.03)	0.89 (0.03)
Mainstem	0.87 (0.03)	0.83 (0.03)
Porcupine	0.91 (0.03)	0.90 (0.03)
Upper	0.96 (0.01)	0.97 (0.01)
White	0.98 (0.01)	0.98 (0.01)
Teslin	0.94 (0.02)	0.96 (0.01)
Fall U.S.	0.84 (0.03)	0.89 (0.03)
U.S.	0.97 (0.02)	0.98 (0.01)
Canada	0.95 (0.02)	0.92 (0.02)
U.S. Border + Canada	0.96 (0.01)	0.95 (0.02)
Mainstem + Upper	0.91 (0.03)	0.89 (0.02)

Table 4. Results of apportionments from mixed-stock analysis of simulated multistock mixtures that may be encountered in a lower Yukon River fishery. Each management group represented in the simulated multistock mixtures was composed of equal proportions of stocks within the management group. Mixture sample size was 400. The mixture and baseline were bootstrapped 100 times. Apportionment estimates (SD) are given for the current and expanded baselines.

	Mix 1			Mix 2		
	Expected	Current	Expanded	Expected	Current	Expanded
Summer	0.15	0.21 (0.03)	0.19 (0.03)	0.43	0.47 (0.03)	0.46 (0.03)
Lower	0.13	0.17 (0.03)	0.17 (0.03)	0.38	0.40 (0.03)	0.41 (0.03)
Tanana	0.02	0.04 (0.02)	0.02 (0.01)	0.05	0.07 (0.02)	0.05 (0.02)
Fall	0.85	0.79 (0.03)	0.81 (0.03)	0.57	0.53 (0.03)	0.54 (0.03)
Tanana	0.32	0.27 (0.03)	0.30 (0.03)	0.12	0.10 (0.02)	0.12 (0.02)
U.S. Border	0.25	0.22 (0.03)	0.23 (0.03)	0.27	0.22 (0.03)	0.23 (0.03)
Canada Border	0.17	0.19 (0.03)	0.18 (0.03)	0.10	0.13 (0.03)	0.11 (0.03)
Porcupine	0.07	0.09 (0.02)	0.09 (0.02)	0.03	0.06 (0.02)	0.05 (0.02)
Mainstem	0.10	0.10 (0.02)	0.09 (0.02)	0.07	0.07 (0.02)	0.06 (0.02)
Upper	0.11	0.11 (0.02)	0.11 (0.02)	0.08	0.08 (0.02)	0.08 (0.01)
White	0.10	0.10 (0.02)	0.10 (0.02)	0.06	0.06 (0.01)	0.06 (0.01)
Teslin	0.01	0.01 (0.01)	0.01 (0.01)	0.02	0.02 (0.01)	0.02 (0.01)
Fall U.S.	0.57	0.49 (0.03)	0.52 (0.03)	0.39	0.32 (0.03)	0.35 (0.03)
U.S.	0.72	0.70 (0.03)	0.72 (0.03)	0.82	0.79 (0.03)	0.81 (0.03)
Canada	0.28	0.30 (0.03)	0.28 (0.03)	0.18	0.21 (0.03)	0.19 (0.03)
U.S. Border + Canada	0.53	0.52 (0.03)	0.51 (0.03)	0.45	0.43 (0.03)	0.42 (0.03)
Mainstem + Upper	0.21	0.21 (0.03)	0.20 (0.03)	0.15	0.15 (0.03)	0.14 (0.03)

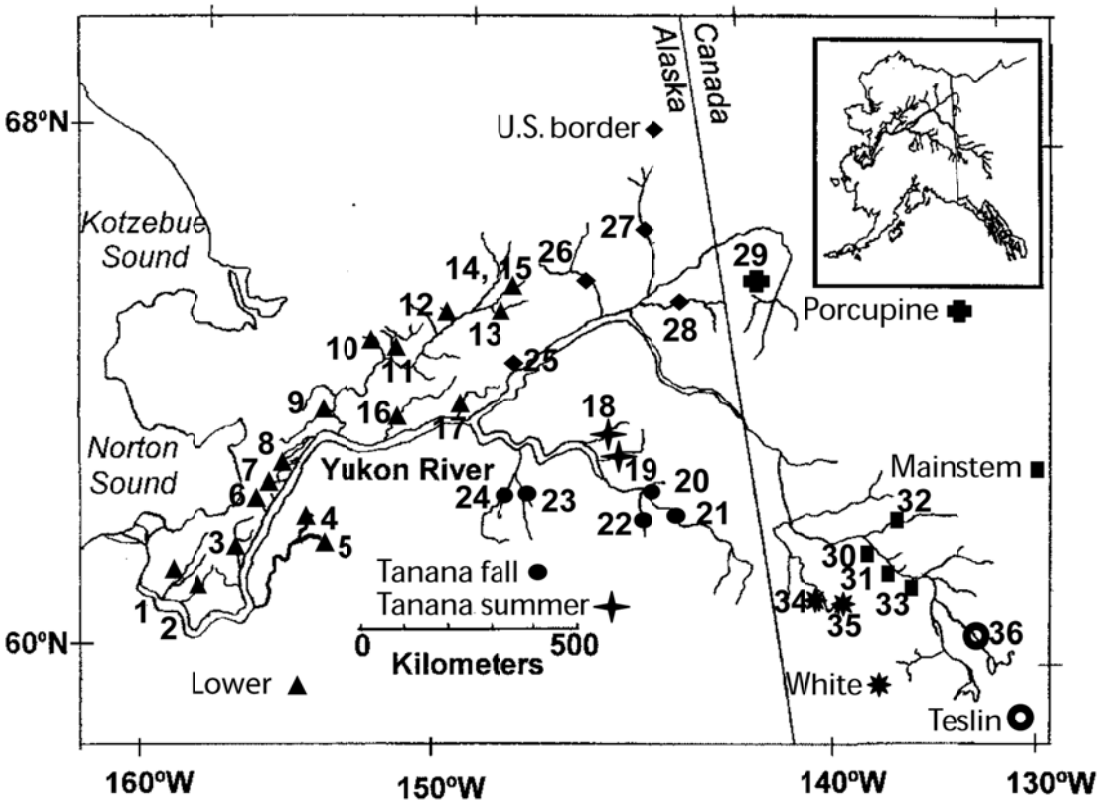

Figure 1. Expanded baseline sampling locations, 1 = Andreafsky, 2 = Chulinak, 3 = Anvik, 4 = California, 5 = Tostoi, 6 = Rodo, 7 = Kaltag, 8 = Nulato, 9 = Gisasa, 10 = Dakli, 11 = Clear, 12 = Henshaw, 13 = Jim, 14 = South Fork Koyukuk Early, 15 = South Fork Koyukuk Late, 16 = Melozitna, 17 = Tozitna, 18 = Chena, 19 = Salcha, 20 = Bluff Cabin, 21 = Tanana Mainstem, 22 = Delta, 23 = Toklat, 24 = Kantishna, 25 = Big Salt, 26 = Chandalar, 27 = Sheenjek, 28 = Black, 29 = Fishing Branch, 30 = Big Creek, 31 = Minto, 32 = Pelly, 33 = Tatchun, 34 = Donjek, 35 = Kluane, and 36 = Teslin.

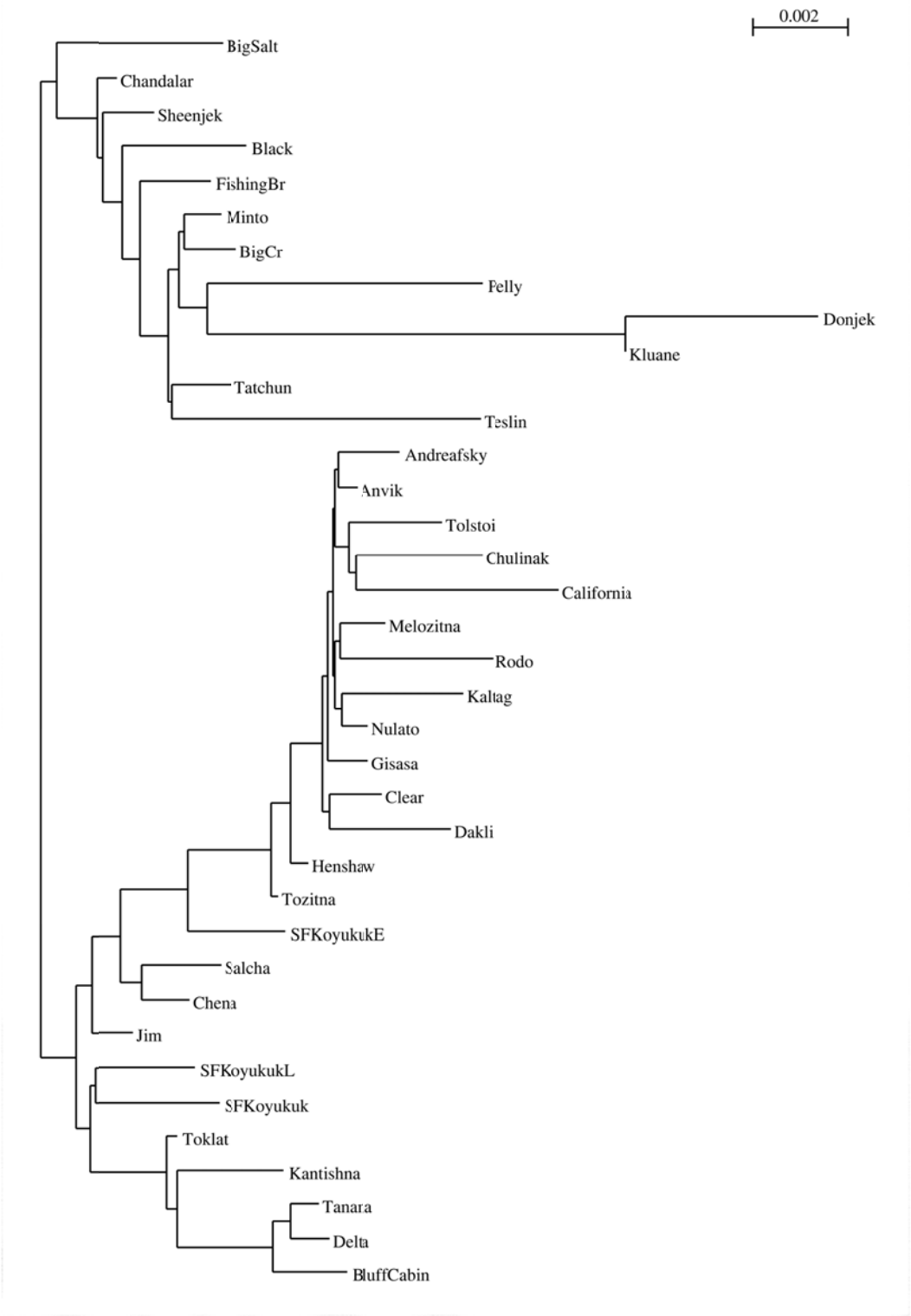

Figure 2. Neighbor-joining dendrogram of Cavalli-Sforza and Edwards (1967) chord distances calculated from allele frequencies at 13 loci.