Development of a SNP Baseline for Yukon River Chum Salmon

R&M # 17-06

Project Proponents: Lisa W. Seeb, Nicholas A. Decovich, and William D. Templin, Alaska Department of Fish and Game, Gene Conservation Laboratory, Commercial Fisheries Division, 333 Raspberry Road, Anchorage, AK, 99518, lisa_seeb@fishgame.state.ak.us

1. Introduction:

Summary:

Chum salmon *Oncorhynchus keta* enter the Yukon River as two runs, summer and fall. While both runs are important resources for local communities, only the fall run spawns within Canada, and it is managed under an international agreement. Currently, the separation between the summer and fall runs is based on a single date, July 15, but considerable overlap occurs during the transition period. Genetic stock identification techniques based on several different marker types have been applied to chum salmon on the Yukon River, all of which provide concordant results. With the decline of the utility of allozyme markers, the Alaska Department of Fish and Game has chosen to use single nucleotide polymorphisms (SNPs), a high-throughput, inexpensive technology, to provide the replacement baseline. We report here the continued development of a SNP baseline to improve information for the management of these salmon.

Chum salmon *Oncorhynchus keta* enter the Yukon River as two runs, summer and fall. Together, summer- and fall-run chum salmon are a mainstay of the Yukon River ecosystem and are an important resource for communities along the Yukon River where subsistence and commercial fishing play a vital role in their economies and culture. Of the two runs, summer-run chum salmon begin to enter the river in late May, tend to be smaller, are not as "ocean bright" on river entry, and spawn in the lower portion of the drainage. Fall-run chum salmon begin to run early in July and by the middle of the month are present in greater numbers than summer-run chum salmon. However, there is considerable overlap in the both the run timing and physical appearance of summer- and fall-run chum salmon; and they cannot be distinguished by calendar date or visual inspection alone.

Currently, the separation for management purposes between the summer and fall runs is based on a single date, July 15, at the test fishery in Emmonak. Chum salmon harvested on or before this date are considered summer-run and fall-run after this date. While summer-run chum salmon spawn entirely within the United States (U.S.) portion of the drainage, fall-run chum salmon spawn in both the U.S. and Canada. The management of fisheries harvesting these populations is covered under the Yukon River Salmon Agreement between the U.S. and Canada. Information on the origin of chum salmon harvested in the subsistence and commercial fisheries on the Yukon River is important for the successful management of these fisheries.

Genetic stock identification methods have been applied to chum salmon on the Yukon River since the late 1980's using a variety of techniques including allozymes, microsatellites, introns, and mitochondrial DNA. The various techniques generally provide concordant results and indicate that chum salmon populations in the Yukon River segregate based on run timing and

geography rather than along political lines. Allozymes, AFLPs, and microsatellites have been shown to provide sufficient power to identify major stock groups in mixtures and are useful to investigate run timing and migration patterns of these aggregates (Wilmot et al. 1992; Scribner et al. 1998; Flannery et al. 2007).

In 1999, the Alaska Department of Fish and Game (ADF&G) implemented a study to use genetic stock identification to estimate the migration timing of summer- and fall-run chum salmon entering the Yukon River. The application used the allozyme genetic baseline described in Crane et al. (2001), a comprehensive baseline with over 8,000 individuals and 23 pooled groups. That study found that genetic stock identification methods can accurately and precisely discriminate summer- and fall-run chum salmon from the Yukon River (ADF&G 2003). Considerable annual variation was detected in the run timing of these two genetically diverse stocks. This was particularly apparent between even- and odd-years with the fall run arriving earlier in even years. In addition, fall-run chum salmon were estimated to contribute as much as 19% of the chum salmon passing the Pilot Station sonar prior to the week including July 15, while summer-run chum salmon contributed as much as 48% to the run the week following that date.

Prior to this analysis, the SNP baseline for the Yukon River included 15 populations (Table 1; Seeb et al. 2005). In this project we continued the development of a comprehensive SNP baseline in Yukon River chum salmon with the addition of both more populations and additional SNPs. The objective was to survey 10 additional populations for at least 36 SNPs. We added eight new populations, and an additional temporal sample from one existing population, Delta River. All populations in the baseline were surveyed for 53 SNPs. This baseline was then examined for its utility for use in mixed stock analysis applications for chum salmon fisheries in the Yukon River.

2. Methods:

COLLECTIONS

At the time that this project was designed, U.S. laboratory archives (ADF&G and United States Fish and Wildlife Service [USFWS]) contained collections from 45 populations spawning in the Yukon River drainage (Appendix 1, Figure 1). Of these, nine collections of chum salmon were selected from under-represented population groups in the SNP baseline and surveyed for individual genotypes at 53 SNPs.

LABORATORY METHODS

Genotyping assays were performed in 384-well reaction plates, with two wells in each plate used as negative controls (no-template) and two wells as positive controls (one for each allele) following the procedures of Smith et al. (2005a) and Elfstrom et al. (2007). Allelic discrimination analysis was performed on the ABI 7900HT using Sequence Detection Software 2.1 (ABI).

DATA COLLECTION

The SNP data collected were individual diploid genotypes for each locus. Genotype data were stored as output text files on a network drive. The data on this network was backed up nightly. Long-term storage of the data was in an Oracle database (*LOKI*) currently supported by ADF&G.

QUALITY CONTROL METHODS

Several measures were implemented to insure the quality of data produced.

1) Each sample that arrived in our laboratory was assigned a unique accession identifier. At the time DNA was extracted or analyzed from each sample, a sample sheet was created that linked each individual sample's code to a specific well number in a uniquely numbered 96-well plate. This sample sheet then followed the sample through all phases of a project, minimizing the risk of misidentification of samples.
2) Genotypes were assigned to individuals using a double-scoring system. Two observers independently produced allele scores for an entire project before the two data sets were compared. Discrepancies between the two sets of scores were then resolved with two possible outcomes: 1) one score was accepted and the other rejected, or 2) both scores were rejected and the score was blanked.
3) Approximately 8% of the individuals, eight samples from each 96-well DNA extraction plate, were reanalyzed for all loci. This insured that the data are reproducible and any errors created from the processing of individual plates were corrected.
4) The final data were checked for duplicated multi-locus genotypes for indication of errors caused prior to extraction of the DNA. When duplicate genotypes were found, the genotype was attributed to the first individual, and subsequent individuals with the same genotype were removed from the analysis to insure that any given individual does not appear more than once in the baseline.

STATISTICAL ANALYSIS

Individual genotype data were summarized as allele frequencies for each SNP in each population. Estimates of the population frequency of individual alleles for each SNP were calculated from the observed frequency of the allele in the representative sample. Observed and expected heterozygosity was calculated using *FSTAT* (version 2.9.3; Goudet 2001), and conformation of genotype frequencies to Hardy-Weinberg equilibrium (HWE) expected ratios was assessed using the exact test in *GENEPOP* (version 3.2, Raymond and Rousset 1995). The significance of departures from HWE for each SNP in each population was determined using α=0.05 adjusted for the number of SNPs (n=53) assayed in each population using the Bonferroni adjusted significance levels ($\acute{\alpha}= \alpha/n=0.0038$).

Some of the SNPs used in this analysis occur in portions of the DNA sequence that are relatively close, and a genotype at one SNP may not be independent of the genotype at another SNP. After accounting for non-random association of genotypes at surveyed SNPs, those remaining were considered independent genetic markers (loci). By definition, SNPs in the mitochondrial DNA are physically linked and were combined into a single locus. Independence between genotypes at each pair of SNPs was investigated by testing for gametic disequilibrium with *GENEPOP*. If genotypes in any pair of SNPs were not independent of each other (linkage), the SNP with the most complete genotypes was chosen for use in the remaining analyses and the other was excluded.

Two measures of population subdivision were calculated from allele frequency differences: Cavalli-Sforza and Edwards' chord distances (CSE; Cavalli-Sforza and Edwards 1967) and F_{ST} (Weir and Cockerham 1984). The chord distances were calculated in SPLUS (Insightful, Inc., Seattle, WA), and *FSTAT* was used to calculate F_{ST} values. Population structure was visualized

as a tree (unweighted pair-group method, Sneath and Sokal 1973) using the program *PHYLIP* version 3.6, (Felsenstein 2002) to view genetic similarities between populations reflected in the interpopulation chord distances.

BASELINE ANALYSIS

Simulations were conducted to evaluate the application of the SNP baseline to genetic stock identification of mixtures of chum salmon harvested in Yukon River fisheries. These simulations were used to help assess whether the baseline of allele frequencies at the available SNPs provided sufficient information to identify individual stocks or groups of stocks (reporting groups) in hypothetical mixtures.

Reporting groups for genetic stock identification of Yukon River chum salmon were defined based on a combination of genetic similarity, geographic features, run timing, and management applications. Populations were assigned to reporting groups based on three criteria: 1) run timing (summer or fall), 2) broad-scale groups, and 3) country of origin.

Simulations were performed using the Statistical Package for Analyzing Mixtures (SPAM version 3.6; Debevec et al. 2000). Baseline and mixture genotypes were randomly generated from the baseline allele frequencies assuming HWE. Each simulated mixture (N = 400) was composed 100% of the stock or reporting group under study. When a reporting group mixture was simulated, all stocks in the stock group contributed equally to the mixture. Average estimates of mixture proportions and 90% confidence intervals were derived from 1000 simulations. Reporting groups with mean correct estimates of 90% or better are considered highly identifiable in fishery applications. Reporting groups with mean correct estimates lower than 90% can still be considered identifiable in mixtures, but sources of misallocation should be considered when interpreting the results.

3. Results:

COLLECTIONS

As originally designed, 10 populations were to be analyzed originating from ADF&G, USF&WS, and CDF&O archives. Following the design outlined in the proposal, tissues from at least seven sites not held in ADF&G archives (California River, South Fork Koyukuk River Early, Jim Creek, and Chandalar, Minto, Kluane, and Teslin rivers) were requested from USFWS and DFO collections. However, during the study period, only ADF&G archives were available. For this reason it was decided to run nine collections from the ADF&G archives (Table 1), but analyze them for the 53 SNPs available for use in chum salmon (Table 2). Through this process the entire baseline was increased so that the current SNP baseline contains 23 populations with all 53 SNPs surveyed for coastwide use.

LABORATORY METHODS

A total of 2,174 individuals from 24 collections representing 23 populations (Table 1) were genotyped at the 53 available SNPs (Table 2). The overall failure rate for successfully assaying genotypes at the 53 SNP loci was 0.70%. The quality control checks employed demonstrated an error rate of 0.64 %.

The quality control checks revealed pairs of individuals in some populations that had identical multi-locus genotypes. The following populations had individuals with duplicate genotypes that

were found to match at 52 or 53 SNPs, a strong indication that the tissues sampled were actually from the same individual: Chulinak River (1 pair), Kantishna River (1 pair), and Donjek River (3 pairs). The second individual in the matching pair was removed from the analysis. All other genotype matches found involved seven or fewer loci, an occurrence that is much more probable by chance when surveying 53 loci. These individuals were not removed from the baseline.

STATISTICAL ANALYSIS

The 53 SNPs included six mitochondrial and 47 nuclear SNPs (Table 2). Significant gametic disequilibrium was found between six pairs of nuclear SNPs and one triplet of mtDNA SNPs. For each set of linked SNPs the locus with the greatest number of individual genotypes scored was kept and the remaining loci in the sets were removed prior to further analysis. The genotypes were kept for the following: *Oke_GHII-3129, Oke_IL8r-272, Oke_GnRH-527, Oke_GPH-105 Oke_U305-307, Oke_U401-143*. After removal of linked SNPs and combining the mtDNA SNPs, the final analyses included 41 independent nuclear loci and 1 mitochondrial locus (described by four SNPs).

After correcting for multiple tests, significant departures from HWE were not found in any collections. Basic statistics, observed heterozygosity and expected heterozygosity and F_{ST}, were calculated for all nuclear loci (Table 2).

Genetic differences between populations were measured using CSE distances calculated from allele frequencies at the 42 SNP loci that remain when using only one SNP from each set of linked loci. A cluster analysis of these interpopulation distances showed two major clusters of populations (Figure 2). The first cluster contains a relatively unstructured set of the summer-run populations from the lower Yukon River. The second cluster contains a more structured set of populations that can be broken further into four subclusters: 1) summer-run populations from the middle Yukon River, 2) fall-run populations from the middle Yukon River, 3) fall-run populations from the Upper U.S. and Canada, and 4) the fall-run population from the Donjek River.

BASELINE ANALYSIS

The initial results from the 23-population baseline of 42 SNP loci indicate that the SNP data are highly concordant with previous analyses using other marker types (Figure 2). Simulations using the current baseline indicate that the summer- and fall-runs of chum salmon can be distinguished from each other with a high degree of accuracy (mean = 95%, s.d. = 0.030). The major lineages within these runs can also be distinguished: 1) Lower Summer [mean=98%, s.d.=0.017], 2) Middle Summer [mean=85%, s.d.=0.051], 3) Middle Fall [mean=90%, s.d.=0.046], and 4) Canada Fall (including Donjek) [mean=88%, s.d.=0.051].

Table 1.– Collections of Yukon River chum salmon surveyed for 53 single nucleotide polymorphisms. Collections new to this study are denoted with asterisks (*).

				Available samples		
Run	**Country**	**Drainage**	**Location**	**Year(s)**	**Size**	**Analyzed**
Summer						
	U.S.					
		West Fork Andreafsky River		1993	100	95
		Chulinak River		1989*	100	93

		Innoko River		1993	88	86
		Anvik River				
			Beaver Creek	1993*	100	95
			Otter Creek	1993	100	96
		Nulato River		1994*	100	95
		Koyukuk River				
			Gisasa River	1994	100	95
			Henshaw Creek	1995	62	62
		Melozitna River		1994*	100	95
		Tozitna River		2003*	200	95
		Big Salt River		2001*	71	71
		Tanana River				
			Chena River	1994	100	95
			Salcha River	2001	85	85
Fall						
	U.S.					
		Tanana River				
			Delta River	1992, 1994*	100, 150	190
			Bluff Cabin	1992	100	95
			Kantishna River	2001*	161	95
			Toklat			
			Geiger Creek	1994	100	95
			Sushana River	1994*	100	95
		Porcupine River				
			Sheenjek River	1992	100	96
			Fishing Branch	1994	100	95
	Canada					
		Pelly River		1993	84	84
		Big Creek		1995	100	95
		White River				
			Donjek River	1994	76	76

Table 2.–Background information, observed heterozygosity (H_o), expected heterozygosity (H_e), and F_{ST} among Yukon River chum salmon stocks for 53 single nucleotide polymorphisms. Sets of linked SNPs are indicated. Only one SNP of each set was retained for analysis; dropped SNPs are noted. Heterozygosities and F_{ST} were not calculated for SNPs located in the mitochondrial DNA.

Locus	Linkage Set	Status	H_e	H_o	F_{ST}	Citation
Nuclear						
Oke_arf-319			0.317	0.332	0.004	2
Oke_BAMBI-116			0.011	0.012	0.009	2
Oke_CKS-389			0.418	0.415	0.008	1
Oke_copa-211			0.073	0.071	0.008	2
Oke_ctgf-105			0.299	0.299	0.005	3
Oke_DM20-548			0.494	0.475	0.034	1
Oke_eif4ebp2-64			0.049	0.051	0.007	2
Oke_FARSLA-242			0.011	0.011	0.004	3
Oke_GHII-2943	1	Dropped				3
Oke_GHII-3129	1		0.436	0.442	0.026	3
Oke_GnRH-373	2	Dropped				1
Oke_GnRH-527	2		0.291	0.277	0.014	1
Oke_GPDH-191			0.508	0.488	0.011	2
Oke_GPH-105	3		0.451	0.457	0.086	3
Oke_GPH-78	3	Dropped				3
Oke_HGFA-319			0.038	0.039	0.004	2
Oke_hsc71-199			0.035	0.036	0.004	2
Oke_hnRNPL-239			0.057	0.059	0.012	3
Oke_HP-182			0.382	0.370	0.018	3
Oke_HSP90BA-299			0.004	0.004	0.005	3
Oke_il-1racp-67			0.382	0.378	0.014	2
Oke_IL8r-272	4		0.186	0.185	0.005	1
Oke_IL8r2-406	4	Dropped				1
Oke_KPNA2-87			0.034	0.033	0.012	3
Oke_MARCKS-362			0.426	0.437	0.003	3
Oke_Moesin-160			0.053	0.053	0.018	2
Oke_ras1-249			0.325	0.328	0.022	3
Oke_RFC2-618			0.487	0.476	0.048	2
Oke_RH1op-245			0.029	0.030	0.002	2
Oke_SClkF2R2-239			0.384	0.396	0.011	2
Oke_serpin-140			0.478	0.494	0.006	2
Oke_TCP1-78			0.126	0.124	0.004	3
Oke_Tf-278			0.240	0.232	0.089	3
Oke_Tsha1-196			0.270	0.274	0.004	2
Oke_u1-519			0.191	0.190	0.012	1

Locus	Linkage Set	Status	H_e	H_o	F_{ST}	Citation
Nuclear						
Oke_u202-131			0.254	0.245	0.050	2
Oke_u212-87			0.015	0.015	0.004	2
Oke_u216-222			0.290	0.286	0.014	2
Oke_u217-172			0.493	0.492	0.014	2
Oke_u200-385			0.461	0.463	0.000	2
Oke_U302-195			0.401	0.370	0.010	3
Oke_U305-130	5	Dropped				3
Oke_U305-307	5		0.315	0.312	0.007	3
Oke_U305-330			0.050	0.051	0.010	3
Oke_U401-143	6		0.190	0.185	0.013	3
Oke_U401-220	6	Dropped				3
Oke_Zp3b-314			--	--	--	2
Mitochondrial						
Oke_Cr231			--	--	--	1
Oke_Cr30	7	Dropped				1
Oke_Cr386	7	Dropped				1
Oke_CR42			--	--	--	3
Oke_CR96			--	--	--	3
Oke_ND3-69	7		--	--	--	1

Note: Citations are as follows: 1) Smith et al. 2005a; 2) Smith et al. 2005b; 3) Elfstrom et al. 2007

.

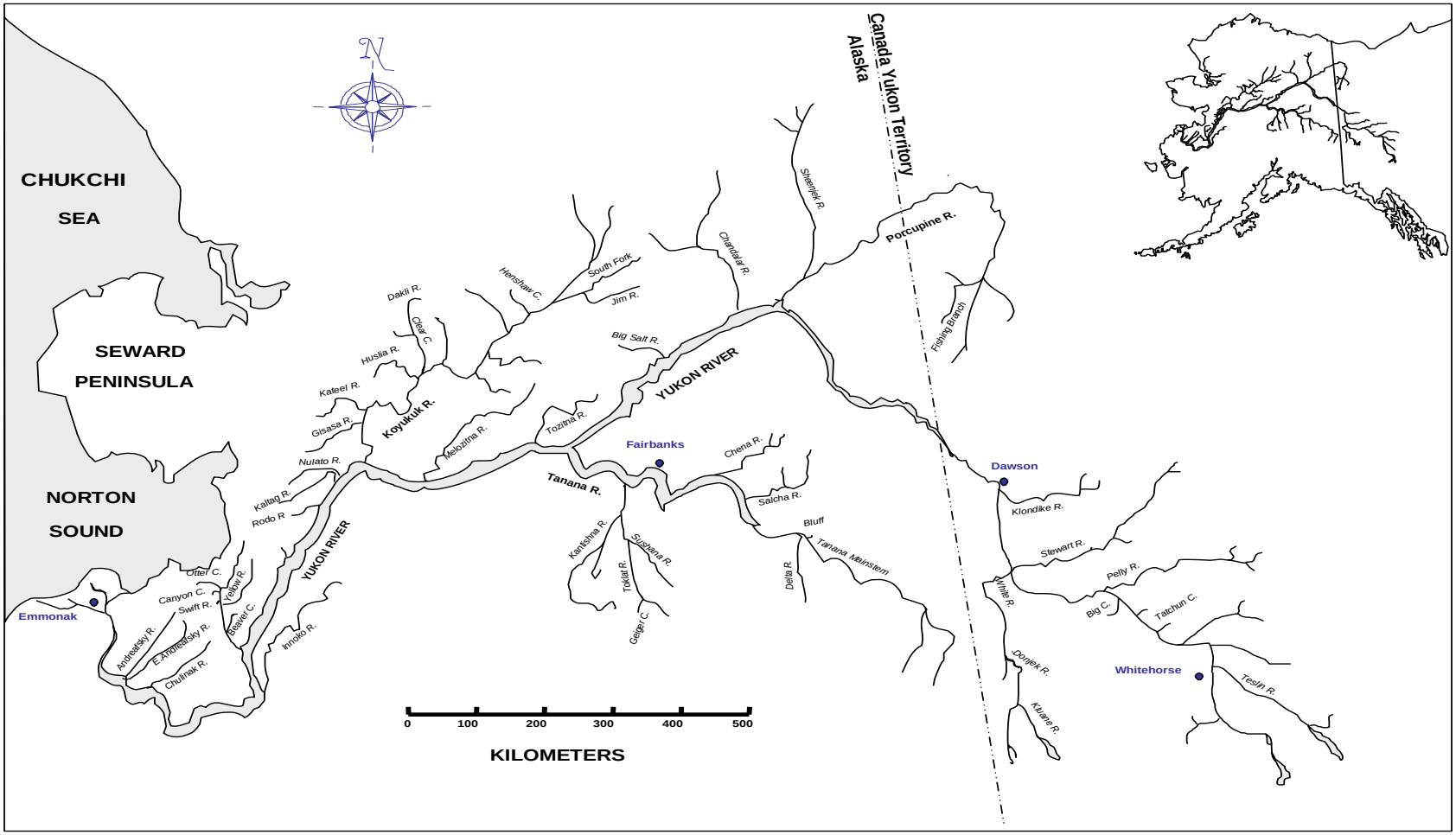

Figure 1. – Map of the locations of chum salmon collections in the Yukon River drainage.

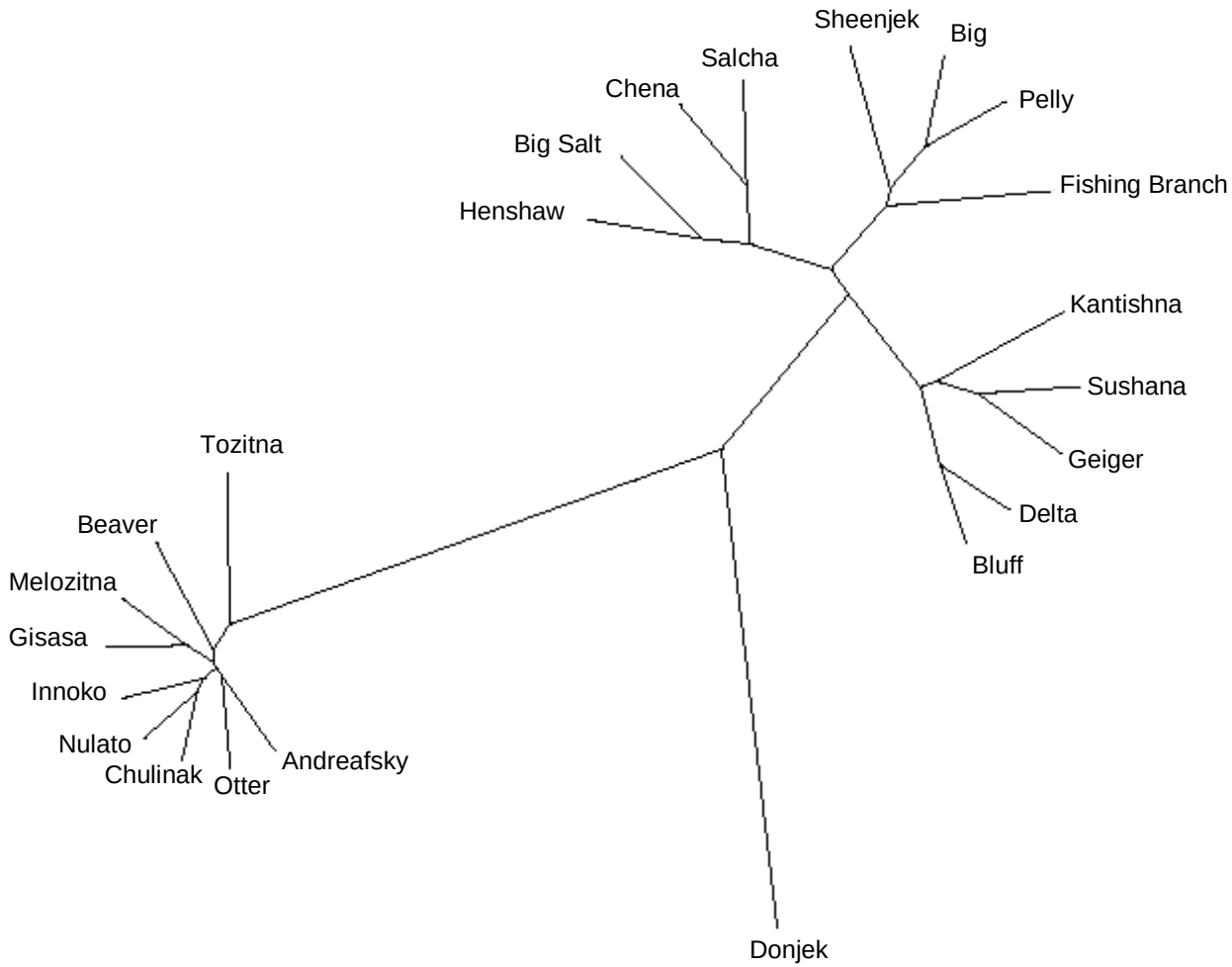

Figure 2. – Unweighted paired group-mean clustering tree based on CSE (Cavalli-Sforza and Edwards 1967) genetic distances between pairs of chum salmon populations in the Yukon River drainage.

4. Discussion:

The objectives of this project stated that 10 populations would be analyzed for 36 SNPs for addition to the SNP baseline, however, only nine populations were available for analysis. From this analysis, the number of SNPs at which the entire baseline was surveyed was increased to 53.

The results presented here demonstrate that the current baseline of SNP markers can discriminate among major regional and timing lineages of chum salmon. The groups identified and used for simulations in this study are more broad-scale than the units defined for chum salmon management in the Yukon River: 1) Lower Summer, 2) Tanana Summer, 3) Tanana Fall, 4) U.S. Border Fall, 5) Porcupine Canada, 6) Mainstem Canada, and 7) Upper Canada. However, the same hierarchical structure was maintained. This more conservative approach was taken based on the limited representation of fall-run populations in the current baseline; only 23 of the 45 populations available are included, of wh
ch 13 are from the Summer run. Greater potential resolution is indicated in the data, but sample sizes are insufficient to support use of smaller units for stock identification. For example, Donjek River is very different from the rest of the Canada populations (Figure 2), but it is a single collection of 76 individuals and is pooled with the rest of the Canada populations. Additional fall-run populations from the U.S. and Canada will need to be analysed before further resolution in genetic stock identification results is supportable.

Since SNP markers directly reflect the underlying DNA sequences, they are readily repeatable and combined with data collected across hardware, chemistry platforms, and laboratories. SNP data can be generated rapidly and are well suited to mixture and migratory studies of chum salmon for a variety of Yukon River applications, including in-season analysis. This is particularly important in the context of multi-jurisdictional fisheries management where both accuracy and repeatability are required.

The data collected in this study from the Yukon River will be combined with other populations from the Pacific Rim to allow for identification of Yukon River chum salmon in mixture analyses throughout the range of chum salmon (e.g. Seeb et al. 2004; Smith and Seeb 2008). We anticipate this expanded baseline will be extremely valuable for studies such as WASSIP (Western Alaska Stock Identification Project), the analysis of chum salmon captured in the walleye Pollock trawl bycatch, and high-seas studies of adults and juveniles (e.g. BASIS Bering Aleutian Salmon International Survey).

Acknowledgements:

Laboratory analysis was performed with the assistance of Carita Elfstrom, Andy Barclay, Zac Grauvogel, Gina Johnston, Beth McLain, and Heather Hoyt. Funding for this work was provided by the 2006 Yukon River Panel Research and Management Fund (Project URM 17-06) under agreement 70181-6-G400.

References:

ADF&G (Alaska Department of Fish and Game). 2003. Yukon River Salmon Negotiation Studies Completion Report. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 3A03-24, Anchorage.

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550-570.

Crane, P. A., W. J. Spearman, and L. W. Seeb. 2001. Yukon River chum salmon: Report for genetic stock identification studies, 1992-1997, Regional Information Report No. 5J01-08 Alaska Department of Fish and Game, Anchorage, Alaska.

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.

Elfstrom, C. M., C. T. Smith, and L. W. Seeb. 2007. Thirty-eight single nucleotide polymorphism markers for high-throughput genotyping of chum salmon. Molecular Ecology Notes 7(6):1211-1215.

Felsenstein, J. 2002. PHYLIP (Phylogeny Inference Package), version 3.6. Department of Genetics, University of Washington, Box 356360, Seattle, WA 98105, USA.

Flannery, B. G., J. K. Wenburg, and A. J. Gharrett. 2007. Variation of amplified fragment length polymorphisms in Yukon River chum salmon: Population structure and application to mixed-stock analysis. Transactions of the American Fisheries Society 136(4):911-925.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

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

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

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

Seeb, L. W, W. D. Templin, C. T. Smith, C. Elfstrom, S. Urawa, R. L. Wilmot, S. Abe, and, J. E. Seeb. 2005. SNPs provide an Easily-Standardized Baseline for NPAFC Studies of Chum Salmon (NPAFC Doc. 907) 12 p. Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK, USA 99518.

Smith, C. T., W. D. Templin, J. E. Seeb, and L. W. Seeb. 2005a. Single nucleotide polymorphisms (SNPs) 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. Baker, L. Park, L. W. Seeb, C. M. Elfstrom, S. Abe, and J. E. Seeb. 2005b. Characterization of 13 single nucleotide polymorphism markers for chum salmon. Molecular Ecology Notes 5: 259-262.

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

Smith, C. T., and L. W. Seeb. 2008. Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single-nucleotide-polymorphism (SNP) baselines for chum salmon. Transactions of the American Fisheries Society 137(3):751-762.

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman, San Francisco, California, 573 pp.

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

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

APPENDIX A: SAMPLING LOCATION, RUN TIME, YEAR COLLECTED, SAMPLE SIZE, AND GENETIC ASSAYS TO DATE FOR CHUM SALMON OF THE YUKON RIVER.

Appendix A1.– Sampling location, run time, year collected, sample size, status in the ADF&G tissue archive, and status for the SNP baseline for chum salmon of the Yukon River. Known collections not currently in the ADF&G archive are included.

Run	Country	Drainage/Location			Year(s)	Sample Size	ADF&G Archives	SNP Baseline
Summer								
	US							
		Archuelinguk River			1989	100	X	
		Andreafsky River						
			Andreafsky River		1987	61		
					200	4		
					2004	200	X	
			West Fork		1993	100	X	X
			East Fork		1993	100	X	
		Chulinak River			1989	100	X	X
		Innoko River						
			Innoko River		1993	88	X	X
			California River		1996	100	X	
			Tolstoi Creek		1997	100	X	
		Anvik River						
			Anvik River		1985	23		
					1988	100		
					1991	345	X	
					1992	7	X	
					2004	340	X	
			Beaver Creek		1992	100	X	
					1993	100	X	X
			Otter Creek		1993	100	X	X
		Kaltag River			1992	100	X	
		Rodo River			1989	78	X	
		Nulato River			1994	100	X	X
					2003	146	X	
		Koyukuk River						
			Gisasa River		1994	100	X	X
			Henshaw Creek		1995	62	X	X
					2004	200	X	
			Dakli River		1992	100	X	
			Huslia River		1993	100	X	
			Clear River		1995	100	X	
			East Fork		1993	100	X	
			South Fork Koyukuk					
				Early	1996	100	X	
				Late	1995	100	X	
					1996	100	X	
			Jim Creek		2002	215	X	
		Melozitna River			1994	100	X	X
					2003	146	X	
		Tozitna River			1992	71		
					2003	200	X	X
		Tanana River						

Run	Country	Drainage/Location		Year(s)	Sample Size	ADF&G Archives	SNP Baseline
			Chena River	1992	96	X	
				1994	100	X	X
			Salcha River	1992	107	X	
				1994	100	X	
				2001	85	X	X
			Clearwater Creek	1990	80	X	
		Big Salt River		2001	71	X	X
Fall							
		Tanana River					
			Tanana Mainstem	1992	97	X	
				1993	100	X	
				2003	50	X	
			Toklat River	1991	60	X	
				1992	155	X	
				1993	200	X	
				1994	400	X	
			Geiger Creek	1994	100	X	X
			Sushana River	1993	200	X	
				1994	100	X	X
			Clearwater River	1990	80	X	
			Delta River	1994	150	X	X
				1987	82		
				1991	100	X	
				1992	100	X	X
			Bluff Cabin	1992	100	X	X
			Kantishna River	2001	161	X	X
				2005	100	X	
		Chandalar River		1987	82		
				1988	73		
				1989	75		
				2001	200	X	
		Porcupine River					
			Black	1995	96	X	
			Sheenjek River	1987	108		
				1992	100	X	X
				1993	64	X	
	Canada		Fishing Branch	1987/89	481		
				1992	100		
				1994	100	X	X
				1997	160		
		Chandindu River		1998	32		
		Pelly River		1993	84	X	X
		Big Creek		1987	70		
				1992	100		
				1995	100	X	X
		Minto River		1989	100		
		Tatchun River		1992	98		

Run	Country	Drainage/Location		Year(s)	Sample Size	ADF&G Archives	SNP Baseline
		White River					
			Kluane River	1992	100		
				2001	200		
			Donjek River	1993	13		
				1994	76	X	X
			Teslin River	1992	100		