

Genetic analysis of juvenile Chinook salmon collected in White Salmon River

United States Fish & Wildlife Service Abernathy Fish Technology Center Report

January 15, 2009

Christian T. Smith
U.S. Fish and Wildlife Service
Abernathy Fish Technology Center
Applied Program in Conservation Genetics
Longview, WA 98632

Sherri Baker
U.S. Fish and Wildlife Service
Dexter National Fish Hatchery & Technology Center
7116 Hatchery Rd.
Dexter, NM 88230

and

Rod Engle and Doug Olson
U.S. Fish and Wildlife Service
Columbia River Fisheries Program Office
Hatchery Assessment Team
Vancouver, WA 98683

Background

Condit Dam is located on the White Salmon River, approximately three miles from the confluence with the Columbia River in Washington State. Construction of the dam was completed in 1913, and other than short-term, temporary trapping activities, has blocked the migration of anadromous fishes since. Several species of anadromous salmonids (*Oncorhynchus* spp.) presently spawn and rear in the three miles of habitat below Condit Dam although, prior to the present work, it was not clear that persistent populations of Chinook salmon (*Oncorhynchus tshawytscha*) existed there.

Due to the lack of fish passage and the high cost of installing a fish passage system, it has been decided by settlement agreement between PacifiCorp, agencies, tribes and environmental groups that Condit Dam will be removed. The original agreement was for removal as early as October 2006 but removal has been delayed to 2009 or later. Removal of the dam is expected to release sediment that has piled-up behind the dam over nearly a century and is expected to temporarily inundate spawning habitat for Chinook salmon listed as Threatened under the Endangered Species Act, and other species in the lower White Salmon River. Determination of appropriate measures to be taken by fisheries co-managers required information regarding 1) the population structure of Chinook salmon spawning in White Salmon River and 2) the relationships among these populations and those in adjacent hatcheries.

This report describes methods used to genotype and perform mixture analyses of Chinook salmon samples collected in White Salmon River. Our focus is on samples collected in 2008, but we also present comparative analyses of data from samples collected in 2006 and 2007 (Smith et al. 2007).

Methods and Materials

Juvenile Chinook salmon were captured by U.S. Geological Survey (USGS) personnel using a rotary trap in the lower portion of the White Salmon River from March through June in 2008 (Brady Allen, USGS, personal communication). Fin clips were taken from a subset of the captured individuals (n=646) and stored in 100% ethanol prior to DNA extraction (4 to 7 months).

Microsatellite Analysis

DNA was extracted from a small (~2mm²) piece of each fin clip using a DNAeasy-96 Tissue Kit (QIAGEN). The polymerase chain reaction (PCR) was used to amplify 13 microsatellite loci (Appendix 1) from each DNA sample. Loci were amplified in 10µl reaction volumes consisting of 5.0µl 2x QIAGEN Multiplex PCR Master Mix (final concentration of 3 mM MgCl₂), and 0.2µl oligonucleotide PCR primer mix. Primer mix compositions and thermal cycling profiles are listed in Appendix 2. Liquid handling was performed using a JANUS Automated Workstation (Perkin Elmer). PCR products were size-fractionated using an AB3130 DNA Sequencer (Applied Biosystems), and raw microsatellite data (electropherograms) were analyzed using GENEMAPPER 4.0. Amplified products were binned into alleles used in the standardized coastwide Chinook salmon baseline (Seeb et al. 2007). All genotypes were scored by two independent readers (double-scoring).

Following completion of the data collection, 10% of all samples were re-analyzed as part of Abernathy Fish Technology Center's (AFTC's) QA/QC protocol. The Microsoft Excel add-in Microsatellite Toolkit (Park 2001) was used to scan the dataset for individuals with identical genotypes.

Columbia River genetic baseline

The Columbia River portion of the standardized multi-agency baseline was used for this work. Additionally, we used unpublished data for several Columbia River populations provided to us for this analysis by Columbia River Inter-Tribal Fish Commission (CRITFC) as well as data for Hood River (collected in 2007 under the present project). In total, the baseline used here contained samples from 54 populations (Appendix 3).

Sub-division of White Salmon River samples

In order to evaluate the possibility of multiple populations within the White Salmon River samples, we used ONCOR (V 4.24.2008; <http://www.montana.edu/kalinowski/Software/ONCOR.htm>) to calculate the probability that the multi-locus genotype of each individual originated from each of the 54 baseline populations. Assigned samples were sorted by collection date and we examined the data for discontinuities associated with the time during which few samples were collected (the first week of May; Figure 1).

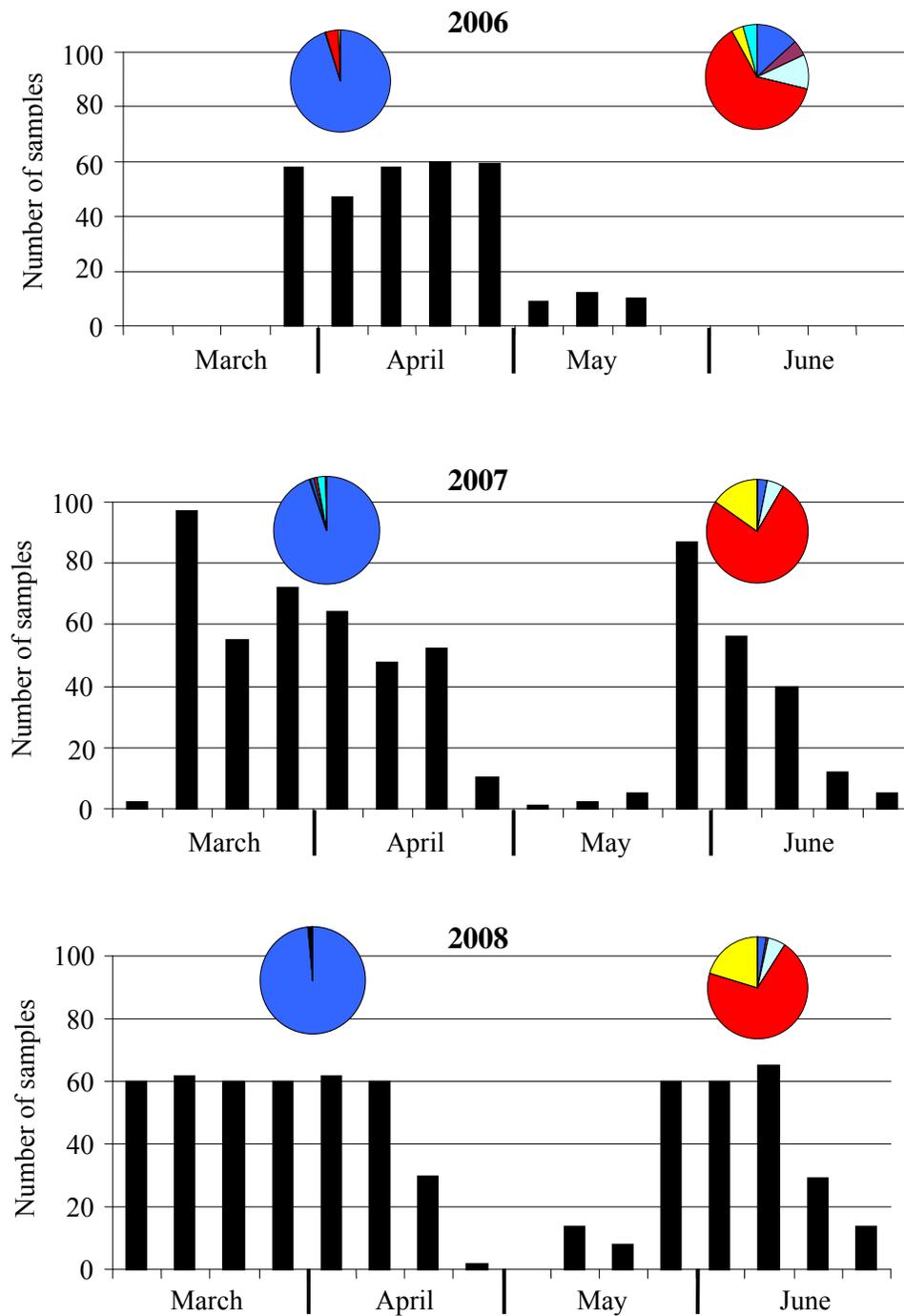


Figure 1. Number of juvenile Chinook salmon analyzed by collection week in each of three years. In each year collections were split into “early” (prior to May 8) and “late” categories. Pie diagrams indicate proportional assignment results of each period in each year (blue = LowCol, red = midupColOT, yellow = SnakeOT, light blue = DeschutesOT as described in Appendix 3).

Genetic diversity observed in White Salmon River

Allelic richness (number of alleles per population, corrected for sample sizes) was calculated for each locus in each collection using the program FSTAT (Goudet 2000). We compared allelic richness observed in the White Salmon River collections to that observed in other samples in the Columbia River baseline.

Genetic divergence among populations

Correspondence Analysis was performed using the program GENETIX (Belkhir et al. 2004) to reduce the genotype matrix to two dimensions and allow visual inspection of the baseline data. Divergence among populations in the Columbia River and between each of the stocks (spring and fall-run) from Little White Salmon National Fish Hatchery (NFH) and their neighbors was examined using a test for allele frequency heterogeneity in ARLEQUIN (Excoffier et al. 2005). The fixation index F_{ST} (θ ; Weir and Cockerham 1984) was calculated between each pair of populations using ARLEQUIN. Statistical significance of pairwise F_{ST} estimates was tested using a permutation procedure with 10,000 replicates.

Mixture analysis

Mixture samples were proportionally assigned to baseline populations and reporting groups following the method described by Anderson et al. (2008), as implemented in the program ONCOR. Prior to performing mixture analysis on the samples collected at White Salmon River, we tested the accuracy of the baseline using 1) simulations and 2) fish of known origin. The simulations involved generating a mixture of 200 fish from one population and then performing mixture analysis on those 200 fish and observing how many assigned back to the population used to generate them. If the baseline were powerful enough to allow perfect mixture analysis, then all 200 fish (100%) would assign back to the correct population. This was repeated 1,000 times for each population, and the mean proportion assigned back to the correct population was recorded.

Simulations may provide optimistic estimates of accuracy, so it is desirable to also test the baseline using “samples of known origin”, or, fish which are not in the baseline but for which true population of origin is known. For this purpose we analyzed an additional plate of 95 fish which consisted of 40 fall Chinook from Spring Creek NFH, 39 fall Chinook from Little White Salmon NFH and 16 spring Chinook from Little White Salmon NFH. Individual assignment and

proportional assignment were performed on these fish as described above in order to empirically evaluate the accuracy of our baseline for correctly assigning fish to these populations of primary interest.

Finally, proportional assignment was performed on the juvenile samples collected from White Salmon River. Samples collected in 2006 and 2007 were re-assigned using identical methods in order to make results comparable across years.

Results

Microsatellite Analysis

Of 646 samples analyzed from White Salmon River in 2008, 5 samples failed to amplify at all thirteen loci and another 17 samples failed at four to ten loci each. These 22 samples were removed from the data yielding 625 samples for mixture analysis (distribution shown in Fig 1). None of the samples of known origin failed to amplify at more than two loci, so all 95 samples were included in the mixture analysis. The PCR failure rate in the 720 (625+95) samples included in the mixture analysis was 0.3%. Of 949 QA/QC genotype comparisons (13 loci x 73 fish) we observed 7 conflicts (~0.7%), all of which appeared due to allelic dropout in either the original data or the QC data. Assuming dropout was equally likely in the original and QC runs, this suggests a genotyping error rate of (~0.4%).

Sub-division of White Salmon River samples

Of the 625 White Salmon River juveniles analyzed using individual assignment, 484 (77%) were assigned to a baseline reporting group with $\geq 90\%$ probability. In collections prior to May 8, 2008, 364/366 (>99%) of these fish assigned to Spring Creek NFH and other populations in the LowCol reporting group. Conversely, only 6/118 (~5%) of fish collected from May 8 through June 26, 2008, were assigned to these groups. These results were similar to those observed for collections taken from the rotary trap in 2006 and 2007, and provided further support for the hypothesis that two distinct populations are migrating out of the White Salmon River (Smith et al. 2007). Based on this genetic disjunction and a reduction in the number of samples collected during the first week of May in all three years, we divided samples from each year into early (March through May 7) and late (May 8 through the end of June) components prior to performing proportional assignment (graphical depiction in Fig 1).

Genetic diversity observed in White Salmon River

The number of alleles per locus in the 2008 samples from White Salmon River ranged from 5 at *Ots9* in the early component to 44 at *Omm1080* in the late component. In total 396 unique alleles were observed in the White Salmon River collections. Allelic richness was generally higher in the late component than it was in the early component (Table 1). Compared to other populations in the Columbia River baseline, the White Salmon River collections did not exhibit exceptionally high or low allelic richness (mean rank from 44th to 89th percentile; Table 1). Allelic richness estimates for all loci and collections are listed in Appendix 4.

Table 1. Allelic richness observed at thirteen microsatellite loci in six collections of Chinook salmon from White Salmon River. Mean Percentile indicates the percentile of the corresponding collection compared to the fifty-four Chinook salmon populations in the Columbia River baseline (Appendix 3).

	2006		2007		2008	
	Early	Late	Early	Late	Early	Late
<i>Ssa408</i>	10.2	15.1	10.9	14.5	9.8	14.9
<i>Ots3M</i>	6.5	6.7	6.8	7.7	6.1	7.5
<i>Ogo4</i>	7.7	9.4	8.4	8.2	7.4	8.6
<i>Ots211</i>	12.6	16.7	13.2	15.5	12.0	16.6
<i>Ots201</i>	15.4	18.9	15.2	17.1	14.0	17.5
<i>Ots212</i>	14.8	16.4	14.4	17.3	14.4	15.9
<i>Ots9</i>	3.5	4.5	3.1	4.3	3.0	4.2
<i>Ogo2</i>	8.3	10.4	7.3	10.2	6.2	10.0
<i>OtsG47</i>	8.4	7.3	8.7	7.2	8.2	7.8
<i>Ots213</i>	16.7	18.4	16.1	19.4	15.2	19.4
<i>Ots208</i>	21.0	20.9	20.4	21.8	20.2	21.5
<i>Oki100</i>	17.5	18.2	17.0	18.7	17.0	18.6
<i>Omm108</i>	21.0	18.0	19.5	21.1	19.7	21.8
Mean Percentile	57.9%	84.0%	54.9%	87.3%	44.0%	88.7%

Genetic divergence among populations

Correspondence analysis clustered the Chinook salmon baseline samples into three broad groups, including 1) lower Columbia fall / hatchery, 2) mid-upper Columbia River fall – summer, and 3) spring run Chinook salmon (Figure 2). Five spring runs (Kalama Hatchery, McKenzie Hatchery, North Santiam Hatchery, and Klickitat River) did not fit inside these groups, illustrating a broad amount of variability among spring run Chinook salmon. This analysis revealed similarities between each of the White Salmon River samples and an adjacent hatchery stock. Early White Salmon River samples from all three years clustered near the Spring Creek NFH sample. Late White Salmon River samples from 2007 and 2008 sample clustered with the Mid-Columbia Fall Chinook salmon (which included the Little White Salmon NFH fall Chinook salmon). The late White Salmon River sample from 2006 fell further from the center of this group, however, no strong interpretations were made of this based on the low sample size (n=22).

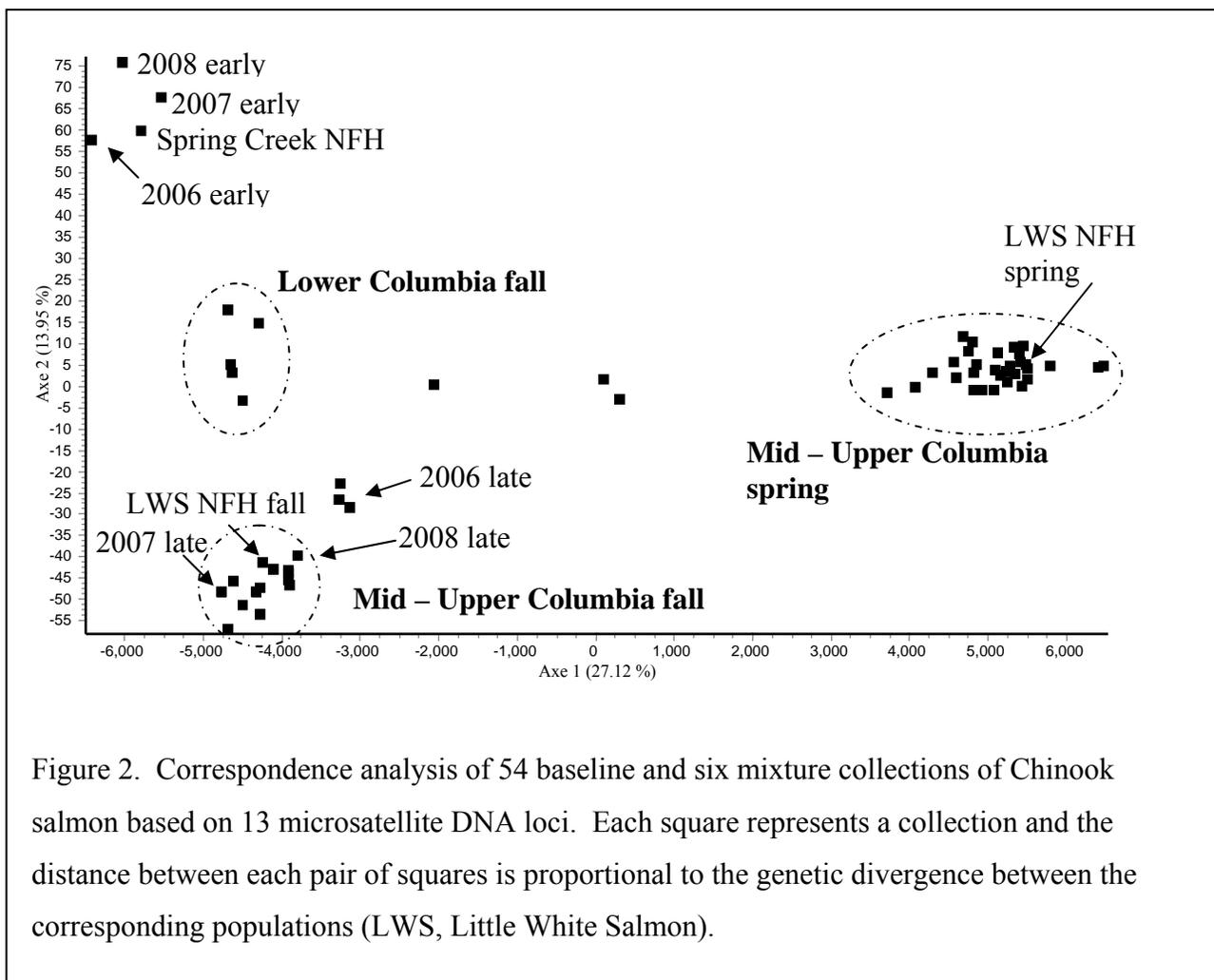


Figure 2. Correspondence analysis of 54 baseline and six mixture collections of Chinook salmon based on 13 microsatellite DNA loci. Each square represents a collection and the distance between each pair of squares is proportional to the genetic divergence between the corresponding populations (LWS, Little White Salmon).

Heterogeneity tests revealed significant pairwise differences in allele frequencies between the three early White Salmon River collections and between each of these and all baseline populations. Conversely, allele frequencies in late White Salmon River collections from 2006 and 2008 were not significantly different from one another. Further, allele frequencies were not significantly different between the late collections and several mid-upper Columbia River and Snake River collections (Table 2).

Pairwise F_{ST} s were significant between the early White Salmon River collections and all other populations except Spring Creek NFH fall. Similar to the heterogeneity results, pairwise F_{ST} s indicated similarity between pairs of late White Salmon River collections and between these and several mid-upper Columbia River and Snake River collections.

Table 2. Results of tests of significance of genetic divergence among collections from White Salmon River and collections from known populations of Columbia River Chinook salmon. Results are shown for populations that were NOT significantly different from one or more of the smolt collections. Comparisons for which heterogeneity tests were significant but F_{ST} was not greater than zero are indicated by *. Comparisons for which heterogeneity tests were not significant AND F_{ST} was not greater than zero are indicated by **. Population descriptions are provided in Appendix 3.

Collection	Collections & populations not significantly different
2006 early	Spring Creek NFH*
2006 late	2007 late*, 2008 late**, Hood R*, LWS fall**, Ldeschut*, KLCKsu**, KLCKfa*, HanfordR**, PRH*, LFH*, CWFCH**
2007 early	Spring Creek NFH*
2007 late	2006 late*, 2008 late*, KLCKsu**, KLCKfa**, HanfordR*, PRH*
2008 early	Spring Creek NFH*
2008 late	2006 late**, 2007 late*, Ldeschut*, KLCKsu**, KLCKfa**, HanfordR*, PRH*, CWFCH*, NPTH*

Mixture analysis

Mixture analysis of simulated fish indicated that the mean assignment accuracy to population was 70.7% (range 50.0% - 99.0%; Appendix 3). When populations were pooled based on genetic similarity into 15 reporting groups, mean accuracy rose to 92.1% (range 48.4% - 99.9%). An example of a group of populations for which mixture analysis accuracy was increased by pooling was the Mid-Columbia Fall Chinook reporting group (MidupColOT). In this case, accuracy to each population was low (mean = 32.8%), but simulated accuracy to the group was over 85% for each population. Assignment to the hatchery populations of primary interest to the present work suggested >90% accuracy to population for Spring Creek NFH and <90% accuracy for the two Little White Salmon NFH stocks. Accuracy to reporting group was >90% for Spring Creek NFH and 88.2% and 89.4% for spring and fall Chinook, respectively, from Little White Salmon NFH. Simulated accuracies were generally lower than those reported last year. The difference seems likely due to the more realistic simulation method described this year (Anderson et al. 2008).

Even under the new more realistic simulation method, several assumptions remained necessary, including that all populations contributing to the mixture were represented in the baseline, and that the samples in the baseline were representative of the populations they represented. These simplifying assumptions generally make simulation accuracies high relative to assignment of real fish. By analyzing “fish of known origin”, for which population of origin is known and which are independent of the baseline, we were able to empirically evaluate the baseline. Results for mixture analysis of samples of known origin in the present study are listed in Table 3. Assignment of fish from Spring Creek NFH was accurate to both population (93.1%) and reporting group (99.5%). Accuracy in assigning Little White Salmon NFH fall run fish to reporting group was also high (89.5%), however, assignment of the same fish to population (24.1%) was dramatically less accurate than predicted by the simulations. Assignment of spring run fish from Little White Salmon NFH was very inaccurate to both population (24.9%) and reporting group (46.4%).

A pattern that we observed in the results of our analysis of fish of known origin is that mixture analysis becomes less accurate as anthropogenic fish transfers from other baseline populations becomes more pronounced. The Spring Creek NFH fall Chinook stock was originally derived from the geographically adjacent White Salmon River, and these fish were assigned with the

Table 3. Mixture results for samples of known origin. Contribution indicates the proportion of the 95 samples that were collected from the national fish hatchery belonging to each reporting group. For the LowCol reporting group this was the Spring Creek NFH fall stock, for MidupColOT this was the Little White Salmon NFH fall stock, and for upColST this was Little White Salmon NFH spring stock.

Reporting Group	Contribution	Estimate	
		To Reporting Group	To NFH*
LowCol	0.421	0.423	0.392
LewisHsp		0.019	
upColST	0.168	0.078	0.042
MidupColOT	0.411	0.367	0.099
SnakeOT		0.025	
RapCWST		0.087	

* This mixture analysis incorrectly allocated 46.7% of the fish to non-NFH baseline populations.

greatest accuracy. The Little White Salmon NFH fall stock was founded from populations destined upstream of The Dalles Dam (Celilo Falls), and accuracy in assignment of these fish was intermediate. Finally, the spring run at Little White Salmon River was founded from the “Carson stock” which historically included a mix of founders from upper Columbia River and Snake River populations, and assignment of these fish was the least accurate of those included. Assignment of some of these fish to SnakeOT and RapCWST (Table 3) reveals the complex legacy of the Little White Salmon NFH spring stock.

Based on the simulations and assignment of fish of known origin we are confident in mixture analysis results that assign fish to reporting groups and are even reasonably confident in proportional assignments to Spring Creek NFH but are less confident in the available microsatellite baseline to produce accurate proportional assignments to the Little White Salmon NFH.

Proportional assignment of the fish collected in White Salmon River closely mirrored our expectations based on the individual assignment results described above. Assignment of the early run in all three years was predominantly (94.7%-98.6%) to the LowCol reporting group and assignment of the late run was predominantly (63.3%-76.0%) to the MidupColOT reporting group

(Table 4). The majority (73.4% – 86.4%) of early fish assigned to Spring Creek NFH, while a substantially smaller proportion (8.3%-13.5%) of late fish were assigned to Little White Salmon NFH. This last result could reflect either a small contribution of Little White Salmon NFH fall Chinook or our inability to distinguish Little White Salmon NFH fall Chinook from Mid-Columbia Fall Chinook stocks represented in the baseline. Coded-wire tag recoveries have indicated that a substantial component (>50%) of late returning adults in White Salmon River are Little White Salmon NFH strays, suggesting that the later case may be more likely.

Table 4. Proportions of juvenile Chinook salmon caught at White Salmon River allocated to NFH populations and reporting groups in the Columbia River baseline. Reporting groups are defined in Appendix 3. Numbers in parentheses indicate 95% bootstrap confidence intervals based on 1000 permutations. Early components (a) were taken between March and the first week of May, and late components (b) were taken between May 8 and June 30.

a)

	2006 early (n=291)		2007 early (n=401)		2008 early (n=378)	
Population						
LWS NFH fall	0.008	(0.000-0.029)	0.002	(0.000-0.017)	0.000	(0.000-0.000)
LWS NFH spring	0.000	(0.000-0.000)	0.007	(0.000-0.015)	0.000	(0.000-0.005)
Spring Creek NFH	0.736	(0.562-0.749)	0.734	(0.549-0.743)	0.864	(0.703-0.860)
Reporting Group						
LowCol	0.949	(0.881-0.964)	0.947	(0.908-0.961)	0.986	(0.960-0.996)
Willamette	0.000	(0.000-0.014)	0.003	(0.000-0.017)	0.004	(0.000-0.021)
LewisHsp	0.005	(0.000-0.019)	0.000	(0.000-0.004)	0.000	(0.000-0.008)
DeschutesOT	0.000	(0.000-0.013)	0.001	(0.000-0.011)	0.000	(0.000-0.003)
YakimaST	0.000	(0.000-0.000)	0.000	(0.000-0.000)	0.002	(0.000-0.007)
upColST	0.000	(0.000-0.000)	0.011	(0.000-0.032)	0.003	(0.000-0.011)
MidupColOT	0.042	(0.019-0.087)	0.009	(0.000-0.030)	0.002	(0.000-0.015)
SnakeOT	0.005	(0.000-0.039)	0.002	(0.000-0.017)	0.003	(0.000-0.012)
RapCWST	0.000	(0.000-0.000)	0.025	(0.006-0.041)	0.000	(0.000-0.008)
upSalST	0.000	(0.000-0.000)	0.003	(0.000-0.012)	0.000	(0.000-0.000)

b)

	2006 late (n=22)		2007 late (n=207)		2008 late (n=247)	
Population						
LWS NFH fall	0.135	(0.000-0.388)	0.112	(0.032-0.172)	0.083	(0.023-0.143)
LWS NFH spring	0.000	(0.000-0.000)	0.000	(0.000-0.000)	0.000	(0.000-0.000)
Spring Creek NFH	0.037	(0.000-0.136)	0.000	(0.000-0.001)	0.000	(0.000-0.004)
Reporting Group						
LowCol	0.133	(0.000-0.299)	0.032	(0.002-0.082)	0.027	(0.009-0.090)
Willamette	0.048	(0.000-0.150)	0.000	(0.000-0.009)	0.007	(0.000-0.021)
LewisHsp	0.000	(0.000-0.000)	0.000	(0.000-0.000)	0.000	(0.000-0.014)
DeschutesOT	0.105	(0.000-0.305)	0.053	(0.001-0.108)	0.057	(0.014-0.113)
YakimaST	0.000	(0.000-0.000)	0.000	(0.000-0.000)	0.000	(0.000-0.000)
upColST	0.000	(0.000-0.000)	0.000	(0.000-0.000)	0.000	(0.000-0.000)
MidupColOT	0.633	(0.352-0.891)	0.760	(0.600-0.824)	0.706	(0.569-0.775)
SnakeOT	0.036	(0.000-0.287)	0.156	(0.100-0.307)	0.203	(0.123-0.311)
RapCWST	0.046	(0.000-0.136)	0.000	(0.000-0.000)	0.000	(0.000-0.000)
upSalST	0.000	(0.000-0.091)	0.000	(0.000-0.000)	0.000	(0.000-0.000)

Conclusions

Simulations indicated that the standardized Chinook salmon microsatellite baseline provides marginal (~49%) to highly (~99%) accurate estimates of mixture composition to 15 reporting groups within the Columbia River. Of particular relevance to the present study, the Lower Columbia Fall Chinook (LowCol) and Mid-Columbia Fall Chinook (MidupColOT) groups were estimated to within a few percent of true values. Accuracy of mixture analysis to upper Columbia stream type (upColST) was lower, as was accuracy to individual populations within the 15 reporting groups. Analysis of fish of known origin supported accurate estimation of proportions of fish from Little White Salmon NFH and Spring Creek NFH to reporting groups, as well as to population in the case of Spring Creek NFH.

The present data support the existence of two populations of Chinook salmon in the White Salmon River. One population, which we have designated the “early component” based on the relative out-migration time, appears genetically similar to Lower Columbia River Fall Chinook runs and in particular to the fall stock at Spring Creek NFH. The second population, which we have called the “late component”, appears genetically similar to fall stocks from the middle and upper Columbia River. Diversity, measured here as allelic richness, within each of these populations is comparable to that in other Columbia River Chinook salmon populations. Divergence between the early and late stocks in White Salmon River is substantial relative to the total divergence of fall Chinook salmon within the Columbia River (Fig 2).

Little variation was observed among years in the early component in White Salmon River. This could reflect stability of the population (i.e. effective population size large enough to prevent major allele frequency changes due to drift) and a substantial influence on this stock by Spring Creek NFH.

In the three years of samples examined here, transition between the two populations (Lower Columbia River Fall Chinook and Mid-Columbia Fall Chinook) for out-migrating smolts took place in the first two weeks of May.

Acknowledgements

Financial support for this study was provided by the U.S. Fish and Wildlife Service. The results described here represent one part of an ongoing collaborative study, and a broader account is expected to follow. Samples from White Salmon River were collected by a United States Geological Survey team lead by Pat Connolly and Brady Allen. Laboratory assistance was provided by Amanda LaGrange and Lindsay Godfrey. Don Campton, Larry Marchant and Speros Doulos provided insight and support throughout the study. We are grateful to Shawn Narum (Columbia River Inter-Tribal Fisheries Commission) for providing unpublished data that we included in the baseline used for this project. The authors are grateful for constructive comments made on earlier drafts of this report by Denise Hawkins and Patty Crandell.

The views expressed here are those of the authors and do not necessarily reflect those of the United States Fish and Wildlife Service.

References:

- Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences* 65(7):1475-1486.
- 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(2):281-288.
- Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
- Cairney, M., J. B. Taggart, and B. Hoyheim. 2000. Characterisation of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar*) and cross-species amplification in other salmonids. *Molecular Ecology* 9:2175-2178.
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.
- Goudet, J. 2000. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.1). Available from <http://www.unil.ch/izea/software/fstat.html>. Updated from Goudet (1995).
- Greig, 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(3):376-379.
- Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. *Mol Ecol* 7(8):1087-9.
- Park, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection [Ph.D. thesis (in prep.)]. University of Dublin.
- Rexroad, C. E., R. L. Coleman, A. M. Martin, W. K. Hershberger, and J. Killefer. 2001. Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). *Animal Genetics* 32(5):317-319.
- Seeb, L. W., A. Antonovich, A. A. Banks, T. D. Beacham, A. R. Bellinger, S. M. Blankenship, A. R. Campbell, N. A. Decovich, J. C. Garza, C. M. Guthrie, T. A. Lundrigan, P. Moran, S. R. Narum, J. J. Stephenson, K. J. Supernault, D. J. Teel, W. D. Templin, J. K. Wenburg, S. E. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. *Fisheries* 32(11):540-552.
- Smith, C. T., A. Lagrange, and W. Ardren. 2007. Genetic analysis of juvenile Chinook salmon collected in White Salmon River. United States Fish and Wildlife Service, Abernathy Fish Technology Center Annual Report.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38(6):1358-1370.
- Williamson, K. S., J. F. Cordes, and B. May. 2002. Characterization of microsatellite loci in Chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. *Molecular Ecology Notes* 2(1):17-19.

Appendix 1. Thirteen microsatellite loci standardized by the Genetic Analysis of Pacific Salmonids (GAPS) consortium (Seeb et al. 2007).

Locus	Primer sequence (5' to 3')	Citation
<i>Ots201b</i>	F- CAGGGCGTGACAATTATGC R- TGGACATCTGTGCGTTGC	OSU unpublished
<i>Ots208b</i>	F- GGATGAACTGCAGCTTGTTATG R- GGCAATCACATACTTCAACTTCC	(Greig et al. 2003)
<i>Ots211</i>	F - TAGGTTACTGCTTCCGTCAATG R - GAGAGGTGGTAGGATTTGCAG	(Greig et al. 2003)
<i>Ots212</i>	F- TCTTCCCTGTTCTCGCTTC R- CCGATGAAGAGCAGAAGAGAC	(Greig et al. 2003)
<i>Ogo4</i>	F- GTCGTCACTGGCATCAGCTA R- GAGTGGAGATGCAGCCAAAG	(Olsen et al. 1998)
<i>Ogo2</i>	F- ACATCGCACACCATAAGCAT R- GTTTCTTCGACTGTTTCCTCTGTGTTGAG	(Olsen et al. 1998)
<i>Ots3M</i>	F- TGTCACTCACACTCTTTCAGGAG R- GAGAGTGCTGTCCAAAGGTGA	(Banks et al. 1999)
<i>Ots213</i>	F- CCCTACTCATGTCTCTATTTGGTG R- AGCCAAGGCATTTCTAAGTGAC	(Greig et al. 2003)
<i>Omm1080</i>	F- GAGACTGACACGGGTATTGA R- GTTATGTTGTCATGCCTAGGG	(Rexroad et al. 2001)
<i>Ssa408UOS</i>	F- AATGGATTACGGGTACGTTAGACA R- CTCTTGTGCAGGTTCTTCATCTGT	(Cairney et al. 2000)
<i>Ots9</i>	F- ATCAGGGAAAGCTTTGGAGA R- CCCTCTGTTACAGCTAGCA	(Banks et al. 1999)
<i>OtsG474</i>	F- TTAGCTTTGGACATTTTATCACAC R- CCAGAGCAGGGACCAGAAC	(Williamson et al. 2002)
<i>Oki100</i>	F- CCAGCACTCTCACTATTT R- CCAGAGTAGTCATCTCTG	CDFO unpublished

Appendix 2. Oligonucleotide PCR primer mixes and thermal cycling conditions for microsatellite markers.

A. Primer mixes

MSA	μL	MSB	μL	MSC	μL
<i>Ots3M</i> (6FAM)	5.0	<i>Ots213</i> (NED)	20.0	<i>Ots208b</i> (NED)	30.0
<i>Ots211</i> (VIC)	10.0	<i>Ots212</i> (VIC)	4.0	<i>Omm1080</i> (VIC)	15.0
<i>Ogo4</i> (PET)	6.0	<i>OtsG474</i> (PET)	30.0	<i>Oki100</i> (6FAM)	40.0
<i>Ssa408</i> (PET)	30.0	<i>Ots9</i> (6FAM)	4.0		
<i>Ots201b</i> (NED)	8.0	<i>Ogo2</i> (6FAM)	8.0		
dH ₂ O	41.0	dH ₂ O	34.0	dH ₂ O	15.0
Total	100.0	Total	100.0	Total	100.0

B. Thermal cycler profiles

MSA and MSB (°C/min.)	MSC (°C/min.)	
95.0/15:00	95.0/15:00	} 29 cycles
95.0/0:30	95.0/0:30	
59.0/1:30	54.0/1:30	
72.0/1:00	72.0/1:00	
60.0/20:00	60.0/20:00	

Appendix 3. Populations and reporting groups in Columbia River genetic baseline. Mixture analysis accuracy is the proportion of simulated fish, in mixture analyses of simulated fish from that population, that was correctly assigned back to that population and the associated reporting group. Simulations included 200 fish each, and the numbers listed for each population are averages over 1000 simulations.

Population	Reporting Group	Mixture analysis accuracy	
		To population	To reporting group
1 Cowlitz Hat. (fall)	LowCol	0.810	0.987
2 Lewis R. (fall)	LowCol	0.584	0.989
3 Sandy R. (fall)	LowCol	0.818	0.984
4 Cowlitz Hat. (spring)	LowCol	0.933	0.978
5 Kalama Hat. (spring)	Willamette	0.890	0.894
6 Lewis Hat. (spring)	LewisHsp	0.864	0.864
7 McKenzie Hat. (spring)	Willamette	0.886	0.998
8 Hood River (fall)	LowCol	0.862	0.918
9 N. Santiam Hat. (spring)	Willamette	0.869	0.997
10 Little White Salmon NFH (fall)	MidupColOT	0.517	0.894
11 Little White Salmon NFH (spring)	upColST	0.723	0.882
12 Spring Cr. Hat. (fall)	LowCol	0.956	0.999
13 upDeschutes R. (summer)	DeschutesOT	0.864	0.974
14 lowDeschutes R. (fall)	DeschutesOT	0.576	0.700
15 Carson Hat. (spring)	upColST	0.787	0.907
16 Warm Springs Hat. (spring)	midColST	0.984	0.989
17 Klickitat R. (spring)	midColST	0.476	0.484
18 Klickitat R. (summer)	MidupColOT	0.005	0.883
19 Klickitat R. (fall)	MidupColOT	0.075	0.936
20 Shitike Cr. (spring)	midColST	0.980	0.982
21 John Day R. (spring)	midColST	0.856	0.858
22 Yakima Hat. (spring)	YakimaST	0.987	0.987
23 Wenatchee R. (spring)	upColST	0.826	0.935
24 Methow R. (spring)	upColST	0.624	0.831
25 Entiat R. (spring)	upColST	0.974	0.984
26 Hanford Reach (fall)	MidupColOT	0.519	0.853
27 Priest Rapids Hat. (fall)	MidupColOT	0.323	0.923
28 Wells Hat. (fall)	MidupColOT	0.471	0.967
29 Methow R. (summer)	MidupColOT	0.391	0.980
30 Tucannon R. (spring)	TucST	0.990	0.990
31 Imnaha (spring)	SFSalST	0.896	0.904
32 Minam R. (spring)	RapCWST	0.760	0.928
33 Lostine R. (spring)	LostST	0.986	0.986
34 Catherine Cr. (spring)	RapCWST	0.757	0.967
35 Lyons Ferry Hat. (fall)	SnakeOT	0.404	0.831
36 Clearwater R. (fall)	SnakeOT	0.228	0.768

	Population	Reporting Group	Mixture analysis accuracy	
			To population	To reporting group
37	Nez Perce Tribal Hat. (fall)	SnakeOT	0.325	0.831
38	Lolo Cr. (spring)	RapCWST	0.481	0.936
39	Newsome Cr. (spring)	RapCWST	0.577	0.991
40	Dworshak Hat. (spring)	RapCWST	0.469	0.895
41	Red River (spring)	RapCWST	0.486	0.908
42	Powell Trap (spring)	RapCWST	0.539	0.968
43	S. Fork Clearwater R. (spring)	RapCWST	0.691	0.989
44	Rapid River Hat. (spring)	RapCWST	0.791	0.995
45	Big Creek a (spring)	MFSalST	0.823	0.930
46	Big Creek b (spring)	MFSalST	0.846	0.961
47	Johnson Cr. (spring)	SFSalST	0.819	0.982
48	Secesh R. (spring)	SFSalST	0.926	0.951
49	McCall Hat. (spring)	SFSalST	0.689	0.967
50	Sawtooth Hat. (spring)	upSalST	0.919	0.952
51	W. Fork Yankee Fork (spring)	upSalST	0.904	0.976
52	E. Fork Yankee Fork (spring)	upSalST	0.884	0.969
53	Pahsimeroi Hat. (spring)	upSalST	0.912	0.956
54	Marsh Cr. (spring)	MFSalST	0.626	0.655

Appendix 4. Allelic richness observed in each locus and collection. Collection is listed as year and component (“e” for early and “l” for late) for White Salmon River collections. Baseline collections are listed as numbers which correspond to Appendix 3.

Collection	Ssa408	Ots3M	Ogo4	Ots211	Ots201	Ots212	Ots9	Ogo2	OtsG47	Ots213	Ots208	Oki100	Omm108
2006 e	10.2	6.5	7.7	12.6	15.4	14.8	3.5	8.3	8.4	16.7	21.0	17.5	21.0
2006 l	15.1	6.7	9.4	16.7	18.9	16.4	4.5	10.4	7.3	18.4	20.9	18.2	18.0
2007 e	10.9	6.8	8.4	13.2	15.2	14.4	3.1	7.3	8.7	16.1	20.4	17.0	19.5
2007 l	14.5	7.7	8.2	15.5	17.1	17.3	4.3	10.2	7.2	19.4	21.8	18.7	21.1
2008 e	9.8	6.1	7.4	12.0	14.0	14.4	3.0	6.2	8.2	15.2	20.2	17.0	19.7
2008 l	14.9	7.5	8.6	16.6	17.5	15.9	4.2	10.0	7.8	19.4	21.5	18.6	21.8
1	12.9	7.7	8.0	14.7	15.2	13.0	3.9	8.4	8.9	18.2	21.3	19.1	19.8
2	12.7	7.5	8.4	15.5	14.3	15.5	4.3	8.1	9.7	18.5	22.3	18.0	20.9
3	14.3	8.6	8.9	16.4	13.9	17.0	5.1	10.3	8.3	18.5	21.7	17.3	20.4
4	12.1	6.9	7.8	14.4	15.0	15.0	3.5	8.3	9.2	17.4	18.0	18.4	18.0
5	13.6	7.2	8.9	16.6	13.1	14.0	4.2	8.0	7.8	18.3	18.9	16.0	18.9
6	13.1	7.1	9.6	14.9	14.6	13.1	4.4	8.7	7.5	18.0	18.1	17.0	20.3
7	13.9	6.3	6.1	13.4	12.3	10.8	2.3	8.1	6.2	13.5	16.1	14.7	18.0
8	14.6	7.8	8.6	16.7	17.3	16.5	4.8	10.2	8.5	19.4	22.1	18.3	21.9
9	13.1	5.9	5.8	14.7	12.8	10.2	2.2	7.9	7.2	14.1	15.5	15.8	19.9
10	13.8	7.1	8.7	16.0	17.3	16.2	4.5	10.6	7.2	18.2	21.4	17.0	20.4
11	10.6	5.4	7.4	14.3	13.2	11.5	4.2	6.4	2.5	13.1	14.5	14.7	19.6
12	9.8	6.6	6.5	13.0	14.9	14.8	2.7	6.4	7.5	14.9	20.7	16.6	18.9
13	13.4	5.7	7.6	14.8	14.1	15.1	4.2	8.6	10.5	16.9	19.2	16.5	18.1
14	14.5	7.1	7.5	15.9	15.9	15.5	4.5	8.3	9.2	18.8	20.2	17.7	20.6
15	11.9	4.4	7.7	14.6	13.8	11.5	3.8	7.2	3.1	12.5	16.2	13.3	19.7
16	11.2	3.9	6.9	12.1	10.7	10.5	3.5	5.3	3.0	10.5	12.5	15.0	16.6
17	13.5	6.8	10.6	16.4	16.4	14.4	3.9	10.1	6.9	15.9	16.8	18.3	20.8
18	14.5	7.9	8.4	16.6	18.4	16.7	4.5	11.0	8.5	15.5	20.4	18.4	22.6
19	14.5	7.8	7.3	17.6	18.2	15.8	4.8	10.4	7.2	18.9	19.8	18.4	20.8
20	12.1	4.5	8.5	12.1	13.8	11.1	3.7	5.7	1.9	14.4	15.6	13.4	18.2
21	13.7	4.3	8.6	14.4	14.8	11.8	4.1	7.9	2.6	14.5	17.5	15.2	19.6
22	13.1	5.8	7.9	14.7	13.3	12.1	4.1	7.1	4.2	15.0	16.1	13.9	19.8
23	12.5	5.2	8.1	15.0	15.3	13.6	4.1	7.8	2.4	13.9	17.4	14.4	19.2
24	12.1	4.4	8.4	15.4	16.2	11.0	4.0	6.4	2.4	15.1	17.2	14.4	21.0
25	11.4	4.7	7.0	13.2	12.6	10.7	4.0	7.1	2.4	12.5	15.0	13.3	17.9
26	14.5	7.0	8.0	16.2	17.9	16.6	4.8	10.7	7.3	19.4	21.3	18.3	21.9
27	13.4	7.7	7.3	16.1	16.1	15.8	4.2	10.5	7.1	18.5	21.3	17.9	21.8
28	13.0	6.7	6.1	15.5	18.7	14.6	3.9	10.0	6.4	17.8	20.1	18.0	20.1
29	13.3	6.7	5.6	15.7	18.8	15.1	3.9	9.7	7.0	17.8	21.1	17.9	21.0
30	12.2	3.8	7.7	11.7	13.3	10.3	3.5	6.1	3.9	13.5	14.5	12.7	16.8
31	11.3	3.8	7.4	13.5	15.4	11.9	4.2	7.3	2.5	14.1	16.3	15.5	17.6
32	13.7	4.6	7.8	14.9	16.2	10.8	4.3	7.3	2.6	15.4	17.1	15.5	18.6
33	10.2	3.3	7.0	10.4	12.3	8.9	3.4	8.1	3.0	13.9	13.6	12.8	16.5
34	12.9	3.7	7.9	15.1	15.1	11.0	4.2	7.1	2.3	13.8	16.3	14.7	17.4
35	13.1	6.4	7.6	15.7	16.0	16.1	4.1	10.3	8.0	18.9	21.3	17.8	20.9
36	13.8	7.2	8.0	16.7	16.1	16.2	3.9	10.3	7.3	17.8	22.1	19.0	21.3
37	13.2	7.5	6.9	15.8	15.2	16.0	4.1	9.9	7.8	18.2	21.0	18.6	20.9
38	13.1	4.7	7.9	15.6	15.9	11.6	4.1	6.9	2.2	16.1	16.7	16.2	19.2
39	11.2	3.8	6.9	13.6	13.8	8.6	4.4	7.3	2.2	14.9	14.8	14.2	17.9

Collection	Ssa408	Ots3M	Ogo4	Ots211	Ots201	Ots212	Ots9	Ogo2	OtsG47	Ots213	Ots208	Oki100	Omm108
40	13.0	4.0	7.8	15.8	15.6	10.3	3.8	7.8	2.6	16.1	15.9	15.8	21.0
41	13.3	4.0	8.0	14.5	12.9	11.7	4.2	6.9	2.9	13.9	17.0	14.6	20.4
42	13.6	4.2	7.9	14.7	14.8	11.3	4.3	7.8	2.6	15.2	16.1	15.1	18.7
43	12.4	4.1	7.4	15.1	13.7	10.4	4.4	7.8	2.9	14.9	14.8	14.4	19.8
44	10.3	3.5	7.2	13.3	13.5	9.5	4.4	6.8	2.8	14.4	13.3	13.1	16.5
45	8.5	3.0	6.6	11.9	14.9	8.7	4.0	8.3	1.2	13.8	16.9	13.3	18.4
46	9.5	3.0	7.8	11.6	14.0	10.0	4.0	7.3	1.8	13.4	14.2	13.2	17.5
47	9.4	5.1	8.9	14.3	13.9	9.4	3.9	6.9	1.2	14.2	14.5	14.0	17.3
48	9.2	3.3	8.2	14.0	13.9	10.2	4.2	6.3	1.4	14.0	15.5	13.7	19.8
49	9.7	4.7	9.6	13.6	12.0	10.5	4.4	7.2	1.7	15.1	15.8	15.4	16.7
50	11.3	4.6	7.5	13.4	15.5	10.2	3.9	8.4	1.4	14.6	17.2	16.5	20.1
51	10.4	3.4	6.8	10.4	13.5	8.2	3.0	7.5	1.0	13.5	13.4	12.0	14.3
52	11.2	3.7	6.3	13.7	14.4	10.7	3.4	8.0	1.4	13.1	16.0	14.7	18.9
53	9.9	4.3	7.4	12.2	15.4	9.4	4.2	7.3	1.3	14.1	14.1	13.5	16.7
54	10.6	5.1	6.6	13.7	15.2	9.6	4.4	5.9	1.6	16.2	14.4	14.5	17.5
mean	12.3	5.6	7.7	14.5	15.0	12.7	4.0	8.1	5.0	15.8	17.8	15.9	19.3
sd	1.6	1.6	1.0	1.6	1.8	2.6	0.5	1.5	2.9	2.2	2.8	2.0	1.7