Report to the North Pacific Fishery Management Council

Genetic Stock Composition Analysis of Chum Salmon Bycatch Samples from the 2009 Bering Sea Trawl Fisheries

Andrew Gray Colby Marvin Chris Kondzela Tyler McCraney Jeffrey R. Guyon, PhD

Auke Bay Laboratories Alaska Fisheries Science Center NOAA Fisheries Ted Stevens Marine Research Institute 17109 Pt. Lena Loop Road Juneau, AK 99801

Submitted May 24, 2010

Table of Contents

List of Figures

List of Tables

Tables Page

Table 1[. Results from simulation studies in which 100% of a hypothetical mixture of 400 fish was](#page-9-0) derived from one region (columns) and reallocated back to the region (rows) with SPAM software. ____ 5

Table 2. [Regional SPAM and BAYES stock composition estimates for 1,437 chum salmon samples from](#page-10-0) [the bycatch of the 2009 season Bering Sea groundfish trawl fishery. ____________________________ 6](#page-10-0)

List of Appendices

[used in the analyses of this report. __ 11](#page-15-0)

Introduction

It is important to understand the stock composition of salmon caught in Bering Sea fisheries because this area is a known feeding habitat for multiple brood years of chum salmon (*Oncorhynchus keta*) from many different localities in North America and Asia. Determining the geographic origin of salmon caught in federally managed fisheries is essential to understanding whether management of federal fisheries could address conservation concerns. This report includes preliminary genetic stock identification results for a subset of chum salmon bycatch samples collected from the 2009 U.S. Bering Sea groundfish trawl fishery. The final analysis will be reported in a National Oceanic and Atmospheric Administration (NOAA) Technical Memorandum or other journal publication. National Marine Fisheries Servive (NMFS) geographical statistical areas associated with the groundfish fishery are shown in Figure 1 and are used later in the report to describe the spatial distribution of the chum salmon bycatch and genetic samples.

The goal of this report is to present a stock composition estimate for the 2009 chum salmon bycatch samples collected from the Bering Sea, but it is important to understand the limitations for making accurate estimates of the entire bycatch imposed by the sampling distribution and the genetic baseline. Hence, this report is divided into the following six sections: Introduction, Sample Distribution, Baseline Evaluation, Genetic Stock Composition, Comparison With Previous Estimates, and a Summary. For additional information regarding background and methodology, this report is intended to be supplemented with the chum salmon report prepared previously for the 2005 Bering Sea trawl fishery (Guyon et al., 2010). For the purpose of this report, the chum salmon genetic samples are designated as non-Chinook in the NMFS database since chum salmon comprise over 99.6% of the total non-Chinook bycatch (NPFMC, 2005).

Sample Distribution

Genetic samples were collected by the Alaska Fisheries Science Center's (AFSC) North Pacific Observer Program in 2009 for the Auke Bay Laboratories as a Special Project (designated "Salmon Genetic Project"). As opposed to previous years when samples were collected opportunistically, genetic samples were collected in 2009 as part of the observer's species composition analysis. Axillary processes for genetic analysis and scales for ageing were collected throughout the season and stored in coin envelopes which were labeled, frozen and shipped to the Auke Bay Laboratories.

In 2009, an estimated 46,617 chum salmon were taken in the bycatch of the Bering Sea trawl fishery (NMFS, 2010). This number is 69% less than the average of 147,472 non-Chinook salmon taken in the bycatch between 1994-2009 and 36% less than the median of 71,612 during the same time period (Figure 2). The final genetic sample set for the 2009 chum bycatch was 1,437 fish corresponding to an overall sampling rate of 3.1%.

Potential biases associated with the collection of genetic samples from the bycatch are well documented, and have the potential to affect resulting stock composition estimates (Pella and Geiger, 2009). Methods to collect representative samples are now being reviewed by the Alaska Fisheries Science Center and, when implemented, will reduce biases and improve defensibility of overall stock composition estimates. Potential biases associated with the 2009 chum salmon bycatch sample set were evaluated by comparing the genetic sample distributions with the overall bycatch estimates. To evaluate temporal bias, bycatch estimates and genetic samples were graphed by statistical week (week ending on Sunday) and a visual comparison of the two distributions showed they were comparable (Figure 3).

To evaluate the sample spatial distribution, the chum salmon bycatch was compared with the bycatch samples by statistical area over time (Figure 4). Spatial and temporal sample biases can become more apparent at these higher resolution scales. For example, while high levels of both bycatch and genetic samples were available from statistical area 509, statistical area 521 was overrepresented in the genetic sample set. In addition, spatial bias may be further exacerbated by the uncertainty of catch location for samples collected from shoreside deliveries in which the hauls are mixed and the location of the catch was taken from the first haul of a fishing trip.

Figure 3. Number of Bering Sea chum salmon bycatch and genetic samples from 2009 graphed by statistical week. Total numbers of chum salmon caught in the Bering Sea groundfish trawl fishery (top panel) compared with the available 1,437 genetic samples (bottom panel). Weeks 4-23 correspond to the groundfish "A" season, whereas weeks 24- 44 correspond to the "B" season, the demarcation of which is a vertical line.

Figure 4. Comparison of the chum salmon bycatch with the distribution of available genetic samples, by time and area. Not shown in the chum salmon bycatch are an estimated 342 fish from area 513, 8 from area 516, 33 from area 523, and 10 from area 541. Not shown from the genetic sample set are 2 fish from area 518. Weeks 4-23 correspond to the groundfish "A" season, whereas weeks 24-44 correspond to the "B" season, the demarcation of which is a vertical line.

Baseline Evaluation

Baseline allele frequencies from the published chum salmon microsatellite baseline (Beacham et al., 2009a) were downloaded from the Fisheries and Oceans Canada (DFO) Molecular Genetics web page (http://www-sci.pac.dfo-mpo.gc.ca/mgl/data_e.htm) and a SPAM (ADFG, 2003) baseline file was created within Excel. While this baseline has been used to identify over 50 regional groupings of salmon (Beacham et al., 2009b), our analysis used 6 broad regional groupings to analyze the chum salmon bycatch to (1) ensure the most reliable estimates for this contentious issue and (2) ensure that enough samples from this diverse collection were from a particular region to positively identify it. Regional groupings were similar to that reported previously (Guyon et al., 2010), except (1) all of southeast Alaska, Prince William Sound, British Columbia, and Washington were grouped into one region and (2) 6 populations were moved from east Asia to north Asia (Naiba, Kalininka, Amur, Tym, Udarnitsa, Tugur_River). The resulting six regional groupings are shown in Figure 5 and individual populations from each region are identified in Appendix 1.

(grey), north Asia (red), coastal western Alaska (blue), upper/middle Yukon (green), southwest Alaska (black), and the Pacific Northwest (magenta).

The DFO baseline contains 381 populations of chum salmon assayed for 14 microsatellite markers (Beacham et al., 2009a). For our analysis, 11 markers were used: *Oki100, Omm1070, Omy1011, One101, One102, One104, One114, Ots103, Ots3, Otsg68,* and *Ssa419. Oki2* and *One111* may be available in future analyses, pending optimization, while attempts to optimize the final locus, *Oke3*, have been unsuccessful. To evaluate the ability of the 11 loci to effectively separate the 6 regional groupings in mixed-stock analyses, 100% simulation studies were completed in which all samples of a hypothetical mixture were from one region and that mixture was re-evaluated against the baseline to determine the percentage reallocating back to the correct region. This analysis was completed in SPAM for all six regions (Table 1, top panel). East Asia, western Alaska, upper/middle Yukon, and the Pacific Northwest all allocated back to the correct region with 89-98% accuracy, whereas 85% correctly reallocated to the north Asia region and 82% correctly reallocated to the southwest Alaska region. In an effort to improve stock composition accuracy, the baseline was reevaluated with a subset of loci to increase resolution of the southwest Alaska region and the 100% simulation analyses were repeated (Table 1, bottom panel). With a suite of eight loci, the accuracies of estimates were improved with all simulations now at or above 87%. Based on these results, stock composition estimates are provided based on the allele frequencies for all loci and for the suite of eight loci. Further analyses of marker characteristics are currently ongoing to determine effects of allele number and frequency.

Table 1. Results from simulation studies in which 100% of a hypothetical mixture of 400 fish was derived from one region (columns) and reallocated back to the region (rows) with SPAM software. The fraction of fish from each region is designated.

Genetic Stock Composition

DNA was extracted from the axillary processes of the chum salmon bycatch genetic samples and microsatellite genotyping was performed as described previously (Guyon et al., 2010). Briefly, samples were genotyped for the following 11 microsatellite loci *Oki100* (Beacham et al., 2009c), *Omm1070* (Rexroad et al., 2001), *Omy1011* (Spies et al., 2005), *One101*, *One102*, *One104*, *One114* (Olsen et al., 2000), *Ots103* (Nelson and Beacham, 1999), *Ots3* (Banks et al., 1999), *Otsg68* (Williamson et al., 2002), and *Ssa419* (Cairney et al., 2000). Thermal cycling for the amplification of DNA fragments with the polymerase chain reaction (PCR) was performed on a dual 384-well GeneAmp PCR System 9700 (Applied Biosystems, Inc.). Samples from the PCR reactions were diluted into 96-well plates for analysis with a 16-capillary, 36 cm array on the ABI 3130xl Genetic Analyzer. Genotypes were double-scored with GeneMapper 4.0 software (Applied Biosystems, Inc.) and exported to Excel spreadsheets (Microsoft, Inc.) for further analysis.

A total of 1,563 samples from the 2009 chum salmon bycatch were analyzed, of which 1,442 samples were successfully genotyped for 8 or more of the 11 loci and analyzed in GenAlEx (Peakall and Smouse, 2006) for data integrity. Two duplicate samples (individuals) were removed. In addition, 3 of the remaining 1,440 samples (individuals) were removed due to an unusual excess of homozygotes (between 7 and 10 of the 11 loci were homozygous or not scored). Internal observations have suggested that unusually high homozygosity rates can potentially result from poor DNA quality, although it is unlikely that the affected 3 samples would impact the resulting stock composition estimates. The remaining 1,437 samples used in this analysis had genetic information for an average of 10.77 loci (out of 11). There were 1,226 samples with data for all 11 loci, 123 with 10 loci, 50 with 9 loci, and 38 with 8 loci. There were only 13 individual allele calls which referenced alleles not present in the chum salmon baseline; those alleles were pooled with a baseline allele nearest in size.

To generate the BAYES baseline, a program was written in C to convert the allele frequencies from the SPAM format into allele counts for the BAYES format. For the mixture files, allele designations were converted to match those in the baseline. Genotypes from converted mixtures were then exported from Excel as text files, and C programs were used to format the data into both SPAM and BAYES mixture files. Stock composition analysis was performed with both the SPAM and BAYES software by using previously published procedures (ADFG, 2003; Pella and Masuda, 2001). BAYES software uses a Bayesian algorithm to produce stock composition estimates and can account for missing alleles in the baseline (Pella and Masuda, 2001). BAYES stock composition estimates were derived using all available 11 loci in the mixture (Table 2, top panel). For each BAYES analysis, six Monte Carlo chains starting at disparate values of stock proportions were configured such that 95% of the stocks came from one designated region with weights equally distributed among the stocks of that region. The remaining 5% was equally distributed among remaining stocks from all other regions. For all estimates, a flat prior of 0.002625 (calculated as 1/381) was used for all 381 populations. The analyses were completed for a chain length of 10,000 with the first 5,000 deleted during the burn-in phase when determining overall stock compositions. Convergence of the chains to posterior distributions of stock proportions was determined with Gelman and Rubin shrink statistics, which were all 1.07 or less (Table 2, top panel) conveying strong convergence to a single posterior distribution (Pella and Masuda, 2001).

Table 2. Regional SPAM and BAYES stock composition estimates for 1,437 chum salmon samples from the bycatch of the 2009 season Bering Sea groundfish trawl fishery. BAYES estimates utilized information from all 11 loci whereas SPAM estimates were derived from both 11 and 8 informative loci. The BAYES mean estimates are provided with standard deviations (SD), 95% credible intervals, median estimate, and the associated Gelman and Rubin shrink statistic. Standard deviations and 90% nonsymmetric confidence intervals for the SPAM estimates were determined by the analysis of 500 bootstrapping resamplings of the mixture.

In contrast to the BAYES analysis, the SPAM software uses a maximum likelihood approach in which the mixture genotypes are compared directly with the baseline. Although Version 3.7 of the

SPAM software allows Bayesian modeling of baseline allele frequencies, these options were not utilized for the stock composition analyses. By comparing if there was a difference between the SPAM (which does not account for rare alleles) and BAYES (which can account for rare alleles) estimates, it was possible to determine if rare alleles could be adversely influencing the resulting stock composition estimates. Two SPAM estimates were provided using the maximum likelihood approach, one with all 11 and another with 8 loci for comparison purposes (Table 2, bottom panel). Convergence of the SPAM estimates was monitored with the "Percent of Maximum" value which was determined to be 92.2 for the 11 loci analysis and 90.3 for the 8 loci analysis, both exceeding the 90% guaranteed percent achievement of the maximal likelihood. While stock composition estimates for the two sets of loci were nearly identical (Table 2, bottom table, see overlapping 90% nonsymmetric bootstrap confidence intervals), it was interesting to note that the SPAM estimates derived from the 8 loci were closer to the BAYES estimates. In general, these 8 loci had lower numbers of alleles than the 3 unused loci, although the mechanism of this result is still under investigation.

Comparison With Previous Estimates

The stock composition results from the analysis of the 2009 chum salmon bycatch samples are in general agreement with previous estimates (Figure 6), particularly with the 2005 season. The primary difference in the stock composition of the chum salmon bycatch appears to be the higher contribution from east Asia and lower contribution from western Alaska in more recent years. However, caution must be used in comparisons across years as there are differences in where and when genetic bycatch samples were collected from year-to-year. The 1994-1995 chum bycatch estimates were produced with allozyme data (Wilmot et al. 1998), whereas the 2005 (Guyon et al. 2010) and 2009 chum salmon bycatch estimates were derived from DNA based microsatellite loci. The allozyme and microsatellite DNA baselines have data from many of the same populations, but there is some non-overlap.

Summary

Communities in western Alaska and elsewhere are dependent on salmon for subsistence and commercial purposes. Decreasing salmon returns to western Alaska have caused hardships in these communities. Salmon-dependent communities have expressed concern regarding the numbers of salmon caught as bycatch in the Bering Sea pollock fishery. Stock composition estimates of the salmon bycatch are needed for pollock and salmon fishery managers to understand whether the pollock fisheries may be impacting salmon returns. While stock composition estimates were developed for available 2009 chum salmon bycatch sample set, work remains before unbiased estimates of the entire bycatch can be produced. This report provides a stock composition analysis of a set of 1,437 individuals sampled from the 2009 chum salmon bycatch. The limitations and results of this analysis are summarized below.

Sampling issues:

We highlight the inherent spatial and temporal biases in the 2009 sample set (Figures 3 and 4), which may limit the application of the genetic sample stock composition estimate to the entire 2009 chum salmon bycatch. Through a collaboration with the Alaska Department of Fish and Game, Auke Bay Laboratories is also currently investigating methods for mitigating the effects of bias in a different salmon bycatch sample set. Methods developed through that collaboration may be applicable in future analyses of the chum salmon bycatch.

With regard to improved sampling protocols, NMFS recently published a proposed rule and notice of availability for Amendment 91 to the Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area (75 FR 14016, March 23, 2010). If approved, this rule would require that all salmon bycatch taken in the Bering Sea pollock fishery be sorted by species and counted to ensure compliance with the salmon bycatch caps for the pollock fishery. This may provide additional opportunity for observers to provide representative sampling of the salmon bycatch for genetic analysis, and improve the capability to characterize the origin of salmon taken as bycatch in the Bering Sea pollock fishery.

Evaluation of the baseline:

A chum salmon microsatellite DNA baseline developed by Dr. Beacham at the Fisheries and Oceans Canada was selected for this analysis and is the only publicly available baseline with known populations across the entire Pacific Rim (Beacham et al., 2009a). Genotype information from 11 loci provided discriminatory power to strongly identify the 6 stock distributions used in this analysis. These groupings were similar to regional groupings used in previous analyses using allozyme markers, thereby enabling comparison of the resulting estimates over time. A suite of eight loci selected to improve stock composition results for the southwest Alaska region showed improved accuracies in simulation analyses, although resulting stock composition estimates from this limited data set did not differ substantially from those produced using all 11 loci.

Stock composition estimates:

Overall, the genetic samples collected from the chum salmon bycatch were predominantly from Asian stocks (64%) although substantial contributions were also from western Alaska (13%) and the Pacific Northwest (18%). These are in general agreement with previous estimates; however, there appeared to be a higher contribution from east Asia and lower contribution from western Alaska in more recent years. Given the differences in where and when genetic bycatch samples were collected from year-to-year, caution must be used in comparisons across years. In addition, potential biases in the genetic sample set can adversely affect the stock composition estimates; therefore, estimates derived from these samples should be viewed as estimates of the sample set rather than estimates of the entire chum salmon bycatch.

Application of these estimates:

The extent to which any salmon stock is impacted by the bycatch of the Bering Sea trawl fishery is dependent on many factors including (1) the overall size of the bycatch, (2) the age of the salmon caught in the bycatch, (3) the age of the returning salmon, and (4) the total escapement of the affected stocks taking into account lag time for maturity and returning to the river. As such, a higher stock composition estimate one year does not necessarily infer greater impact than a smaller estimate the next. Efforts to better understand these relationships and their impacts are the subject of a NPRB proposal from Drs. Criddle and Adkison for which Auke Bay Laboratories is collaborating.

Acknowledgements

DNA was purified by Hanhvan Nguyen (NMFS). The baseline used for these analyses was obtained through a web portal sponsored by Fisheries and Oceans Canada and developed in the Molecular Genetics Laboratory with genetic loci identified in a number of laboratories. This document has been peer reviewed by internal and external reviewers.

References

ADFG (Alaska Department of Fish and Game). (2003). SPAM Version 3.7: Statistics Program for Analyzing Mixtures. Alaska Department of Fish and Game, Commercial Fisheries Division, Gene Conservation Lab, Anchorage, AK.

Banks, M.A., Blouin, M.S., Baldwin, B.A., Rashbrook, V.K., Fitzgerald, H.A., Blankenship, S.M., and Hedgecock, D. (1999). Isolation and Inheritance of Novel Microsatellites in Chinook Salmon (*Oncorhynchus tschawytscha*). Journal of Heredity *90*, 281-288.

Beacham, T.D., Candy, J.R., Le, K.D., and Wetklo, M. (2009a). Population structure of chum salmon (*Oncorhynchus keta*) across the Pacific Rim, determined from microsatellite analysis. Fishery Bulletin *107*, 244-260.

Beacham, T.D., Candy, J.R., Sato, S., Urawa, S., Le, K.D., and Wetklo, M. (2009b). Stock origins of chum salmon (*Onchorhynchus keta*) in the Gulf of Alaska during winter as estimated with microsatellites. North Pacific Anadromous Fish Commission Bulletin *5*, 15-23.

Beacham, T.D., Le, K.D., Wetklo, M., McIntosh, B., Ming, T., and Miller, K.M. (2009c). Population structure and stock identification of chum salmon from western Alaska determined with microsatellite and major histocompatibility complex variation. Pages 141-160. Pacific Salmon: ecology and management of western Alaska's populations (C.C. Krueger and C.E. Zimmerman, eds.). In American Fisheries Society, Symposium 70 (Bethesda, MD).

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.

Guyon, J.R., Kondzela, C., McCraney, T., Marvin, C., and Martinson, E. (2010). Genetic Stock Composition Analysis of Chum Salmon Bycatch Samples from the 2005 Bering Sea Groundfish Fishery, Report to the North Pacific Fisheries Management Council. (Juneau, AK, National Marine Fisheries Service, Alaska Fisheries Science Center, Auke Bay Laboratories), pp. 31.

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

NMFS (National Marine Fisheries Service). (2010). BSAI chum salmon mortality estimates, 1991 present, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Regional Office, Juneau, AK.

http://www.fakr.noaa.gov/sustainablefisheries/inseason/chum_salmon_mortality.pdf

NPFMC (North Pacific Fishery Management Council). (2005). Environmental Assessment/Regulatory Impact Review/Initial Regulatory Flexibility Assessment for Modifying Existing Chum and Chinook Salmon Savings Areas: Amendment 84, Secretariat Review Draft. North Pacific Fishery Management Council, Anchorage.

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

Peakall, R., and Smouse, P.E. (2006). Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes *6*, 288-295.

Pella, J., and Geiger, H.J. (2009). Sampling considerations for estimating geographic origins of chum salmon bycatch in the Bering Sea pollock fishery. Alaska Department of Fish and Game Special Publication No. SP 09-08.

Pella, J., and Masuda, M. (2001). Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin *99*, 151-167.

Rexroad, C.E., Coleman, R.L., Martin, A.M., Hershberger, W.K., and Killefer, J. (2001). Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics *32*, 317- 319.

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

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

Wilmot, R.L., Kondzela, C.M., Guthrie, C.M., and Masuda, M.M. (1998). Genetic stock identification of chum salmon harvested incidentally in the 1994 and 1995 Bering Sea trawl fishery. North Pacific Anadromous Fish Commission Bulletin No. 1, 285-299.

Appendices

Appendix 1. Chum salmon populations in the DFO microsatellite baseline with the regional designations used in the analyses of this report.

AGENDA C-1(c)(1) June 2010