Western Alaska Salmon Stock Identification Project
Joint Meeting of Advisory Panel and Technical Committee
23 September, 2009
Hilton Hotel
Anchorage, Alaska

Chair: Eric Volk, ADFG

Call to Order: 08:45

Attending:
Eric Volk, ADFG
Bill Templin, ADFG
Tim Baker, ADFG
Tyler Dann, ADFG
Larry DuBois, ADFG
Steve Brown, CAMF
Pat Martin, CAMF
Dani Evenson, ADFG
Chris Habicht, ADFG
Guy Wade, BBSRI/LGL
Michael Link, BBSRI/LGL
Doug Eggers, ADFG
Andrew Munro, ADFG
Jim Jasper, ADFG
Jeff Estensen, ADFG
Jennifer Hooper, AVCP
Mark Witteveen, ADFG
Lisa Creelman, ADFG
Art Nelson, BSFA
Sam Cotton, AEB
Chuck McCallum, Lake & Peninsula Borough
Robin Waples, NOAA
Mike Sloan, Kawerak
Jill Klein, YRDFA – Arrived at 09:00
Milo Adkinson, UAF – Arrived at 09:30

Missing:
Tanana Chiefs Conference
Aleut Corporation
Federal Representative

Agenda
1. Welcome and introductions
2. Review and approval of agenda
3. Presentation of current project status
   a. Project budget update
   b. Highlights of 2009 sampling efforts
   c. Update on responses to TC and AP comments
      i. Additions to the baseline
      ii. Temporal variation within the baseline
      iii. Communicating results
      iv. Methods to increase precision and accuracy
      v. Marker development
      vi. Approach to estimation of harvest rates
   d. Update on project timeline
4. Review and approval of minutes from April, 2009 meeting
5. Meeting frequency and dates
6. Meeting attendance and quorum

1. Welcome and introductions

The meeting began with Eric Volk (chair) welcoming everyone and each person introducing themselves.

2. Review and approval of agenda

No additions to the agenda were made and the proposed agenda was approved.

3. Presentation of current project status

a) Project budget update
   As of last meeting, the project was roughly $1.5 million short.
   ADF&G has recently received $750,000 which is being used for sockeye SNP development. The remaining $750,000 has been requested.

b) Highlights of 2009 sampling efforts
   Chair addressed when 2009 sample numbers will be available. Preliminary numbers from the 2009 season were reviewed and a final report will be available as soon as data is compiled for all areas.

Norton Sound - Eric Volk
   Three specific areas were problems:
   1) Subdistricts 2&3 subsistence and commercial districts had very few chum salmon sampled because very poor runs of chum salmon precluded a directed chum fishery from occurring. Very few chum salmon were harvested incidentally in the coho fishery in August.

2) Pilgrim River had a record low sockeye salmon run. No commercial fishery occurred at Port Clarence and no samples were collected.
3) Unalakleet and Shaktoolik chum salmon had no targeted subsistence fishing and so no chum salmon were sampled.

Pat Martin: Is it possible to sample chum incidentally caught in the king salmon fishery?
Dani Evenson: Commercial king fishing didn't happen this year, but there could be a possibility of sampling some incidentally caught chum salmon from the king salmon subsistence fishery.

Yukon - Larry Dubois
A total of 951 summer chum and 610 fall chum were sampled from 10 commercial fishing periods. The summer chum commercial harvest had no directed king periods and so there was no possibility to sample commercially harvested chum from the Black River area. However, summer chum sample goals from the subsistence harvest were met and these included fish from Black River.

Kuskokwim - Jeff Estensen
Sockeye
Sample goals were met for commercial fisheries in districts 4, 5, and 1b.
Chum
Sample goals were met for commercial fisheries in districts 4, 5, 1b
Subsistence caught chum were hard to get and few samples were obtained.

Bristol Bay - Tim Baker
Exceeded goals overall but there were gaps temporally, as there are each year.

Pat Martin: What kind of continuing sampling is planned?
TB: As long as there's funding, sockeye will be sampled at a lower level, but not chum.

Alaska Peninsula - Mark Witteveen
Chignik outer areas weren't fished due to low Chignik River sockeye escapement.

SEDM early (during June) wasn't opened due to low numbers at Chignik.

Port Moller; the first sampling day was missed due to weather and logistics. The Three Hills section was not sampled because it was closed for most of the season due to low escapements at Sandy River. Otherwise most all other sampling goals were met.

Pat Martin: Were early Unimak chum sampled? Is all of Unimak represented equally; what divisions are there within the sampling?

MW: Yes, Unimak was sampled during all openings. Unimak was sampled in June as one area, and the Ikatan area as another section. Cape Pankof is the dividing line between the two areas.

Sam Cotton: Is there a problem getting samples from SEDM in June?
MW: The fishery was not open during June due to a low Chignik run. We would not want to test fish for samples under those conditions, because we'd be sampling in a scenario that a commercial fishery would not be allowed and that information would not be as useful.

- Pat Martin: Expressed appreciation for the efforts of many people in fishery sampling.

c) Update on responses to TC and AP comments

Sockeye salmon baseline - *Tyler Dann*

The sockeye baseline is continually being improved – there are currently 50K fish in the Pacific Rim baseline and 21K fish in the Western Alaska - Alaska Peninsula (WAAP) baseline. Currently, 42 nuclear and three mitochondrial SNP markers are being screened. An additional 51 SNP markers are being developed for incorporation into the baseline to improve the precision and accuracy of stock composition estimates. A review of quality control measures was outlined. A review of genetic variation within the Pacific Rim and the WAAP was presented. It is anticipated that these additional markers will be added to the baseline during the winter of 2009/2010. With the 45 SNPs, simulation and proof tests indicate that the 17 coast-wide reporting groups can be distinguished from each other with an acceptable degree of accuracy. Future plans include: increasing the sample sizes for collections which have existing samples available to genotype, incorporating collections gathered through 2009, continuing to investigate the potential impact of temporal variation in collections from similar locations over multiple years, and improving the methods for evaluating the baseline (i.e. increase sample sizes and use heterogeneous mixture compositions).

Questions and Concerns:

Robin Waples: Is there a QC procedure on extraction?

TD: No.

Pat Martin: For the non-geneticists, is this an acceptable error rate (~2%)?

RW: Error rates are difficult to quantify (from tissue sampling to genotyping) but this lab error rate is low.

RW: Is the age composition determined for fishery catches and for baseline samples? Populations which are dominated by a particular brood year could have divergence between generations.

TD/Bill Templin: Grouping baseline populations by reporting regions and the mixed nature of the fisheries would make this distinction difficult.

RW: *G* statistics are sensitive to uneven population sizes and a hierarchical F_{ST} test may be more appropriate.

TD: The majority of samples are of 95 fish and pooling was not biased among regions.

RW: Simulations vs. proof test clarification?

TD: In simulations, genotypes are generated based on the observed allele frequencies from populations and analyzed using SPAM. For 100% simulations, sets of genotypes were generated using allele frequencies from populations within reporting groups to produce mixtures that are representative of single reporting groups. In proof tests, individuals from the baseline populations (within a single reporting group) were taken out of the baseline and used as a mixture and analyzed with BAYES. This test is more conservative because it reduces the baseline and this method is not prone to the concerns of Anderson et al.

Chris Habicht: SPAM simulates while BAYES can't.

PM: Voiced concerns about what the appropriate methodology would be and that the decision will be made soon. What's best?

Jim Jasper: ONCOR is good, BAYES is good. We are looking into other programs.

RW: Problems with simulating populations where samples have more differences in allele frequencies than the actual population. This could create overly optimistic estimates of assignments. ONCOR has a leave-one-out technique which overcomes this problem and the proof tests in BAYES are more stringent (Pella and Masuda paper in 2001 with informative prior information).

BT: This may have been the BAYES2 version.

JJ: Discussed other methods for priors used in BAYES.

Concern: Selecting a method to determine a prior is an important decision, which needs to be carefully discussed before any analysis is completed.

Questions: MHC and the deficit of heterozygotes in the baseline?
      TD: We will be looking into why we see this deficit in heterozygotes.

      New SNPs to baseline in WASSIP region?
      CH: The new SNPs are being ascertained using collections from the WASSIP region.

Chum salmon baseline - *Bill Templin*

The chum salmon baseline is still in development, but currently is based on 16K fish from 202 collections for 153 populations throughout the Pacific Rim and 11K fish within the WAAP area. At present 58 nuclear and 3 mitochondrial markers are screened and 52 markers were retained in the analysis. A review of quality control measures was outlined. A review of genetic variation within the Pacific Rim and the WAAP was presented. Additional novel SNP markers are expected to be developed and incorporated into the baseline to increase the number of markers to 96. The increased number of markers in the baseline is expected to improve the precision and accuracy of stock composition estimates. Future plans also include increasing the sample sizes for collections which have existing samples available to genotype and to incorporate collections gathered through 2009. The utility of reducing the range of the baseline to include only populations likely to be present in WASSIP mixtures will also be investigated.

      RW: Do you have the reporting groups at the sort of level that will be useful to this group?
      PM: We are trying to get more but this is pretty good and these are manageable units.
      Tim Baker: This is much better than before.
      BT: ADF&G is looking for 96 SNPs for each species, which will improve resolution.
      Doug Eggers: Only applicable to stocks which are assessed by escapement and these reporting groups go beyond that.
      PM: How did we decide that SNPs are the best marker to use?
      BT: Previous stock separation methods only were 65% accurate, SNPs currently exceed that and are the best source for new markers. The need for high sample throughput is an important reason for using SNP's. Also, temporal stability is being assessed and is proving to be stable. These are preliminary assessments for chum salmon and I expect these numbers to improve.

Investigating temporal variation - *Bill Templin*

The WASSIP Technical Committee recommended evaluating temporal variation relative to spatial variation in allele frequencies. There is potential for allele frequencies to fluctuate naturally over time and this could affect the utility of the baselines for MSA,

which are composed of collections spanning a number of years. Analyses comparing the among-populations variation to the variation between repeat collections of the same populations at different times were conducted. For both sockeye and chum salmon, pair-wise F_{ST} values suggest that differences between temporal collections within populations are generally small relative to variation among populations. However, lack of variation among the Western Alaska chum salmon populations could be problematic for distinguishing these populations in mixed stock analyses with the current suite of baseline markers. The variation among sockeye populations is estimated to be 41 times the amount of variation between temporal collections and 39 times for chum salmon using ANOVA. In comparison, other sockeye studies have found the variation among populations to be about approximately 40 times greater than annual variation within populations. Temporal variation in allele frequencies within populations does not appear to be a major concern in the sockeye or chum salmon baselines. However, these analyses will be repeated when the full baseline sets are completed and many new temporal comparisons are possible.

RW: Temporal variation is more complicated than simply looking at generations. Overlapping generations and brood year relationships need investigation. We should expect that samples one year apart to be more different than samples collected five years apart.

BT: Among-year variation is minimal compared to among-population variation.

Milo Adkinson: That analysis compares all baseline populations. Variation between baseline populations which do not have much differentiation versus differences between years would be more interesting to compare.

BT: When this baseline is analyzed with the full panel of SNPs, smaller scale evaluation of temporal variation within and among populations will be done (cites Beacham).

RW: Developing markers that are not neutral may allow more differentiation between populations but may also increase the differences between temporal samples.

BT: We will be looking at SNPs that are under selection for specific populations, but that are not subject to selection temporally.

RW: Possible reasons for lack of differentiation between AYK chum?

BT: Glacial recession, straying, common river mouth historically.

PM: Through the ice ages (~9,000 years conservatively).

Asian chum in the baseline, ~35% of all chum in the sea, can't remove it from baseline for WASSIP analysis.

Estimating small proportions - *Jim Jasper*

Obtaining unbiased estimates of stocks in mixtures that appear in low proportions is difficult, especially when increasing sample sizes is not an option. One potential solution is to make use of multiple strata and increase the scope of the estimate. A simulated example using data from the Ugashik District where the contribution of North Peninsula stocks of sockeye salmon was set at 1.1% was presented. Results illustrated that the precision and accuracy of the estimate of North Peninsula stock improved as the scope of the estimate increased from separate estimates for each temporal strata within each year, to separate estimates for each year of data, and a single overall estimate for all years. Further analysis is required to fully investigate this approach. Jim Jasper also presented some preliminary results from a new analysis that allocates to regions rather than to populations that shows promise as far as reducing bias of misallocating fish to regions with more stocks.

   RW: Does the number of baseline populations in the reporting region affect the amount of misallocation that goes to each reporting region?

   JJ: That is a possibility and we can test that by reducing the number of baseline populations in each reporting region and then seeing if the same proportion is misallocating to that region.

   MA: Having a hard time following the "Allocation to Region" equation.

   Chair: Jim, Milo, and Robin need to get together and go over this and then they need to try to find a way to explain it to the rest of the AP.

   BT: Instead of assigning to the baseline populations and summing them up to the region level, this equation is trying to assign to the region.

New algorithm decreases the bias towards the region that has the greatest proportion of stock in the mixture, but the platform needs to be in something other than WinBUGS, maybe C++ (this is an issue with speed of running these simulations and WinBUGS is currently the limiting factor). Also, the algorithm needs to be tested with real data.

   Chair: Great example of the need for a technical committee. Milo, Robin, and Jim need to get together and see where this can lead.

   RW: Bias of not allowing a proportion to be centered around zero (the negative proportion problem) may be contributing to the assignment bias towards larger baseline reporting regions.

   RW: This bias will be applied to a fishery for two reasons. 1) The fish taken from the fishery are only a small proportion of the overall harvest and 2) a small proportion in the assignment will be biased as well.

MA: These uncertainties may be additive between the genetics and the sampling. The representation of a fishery by the samples may be optimistic.

Estimating harvest rate - *Doug Eggers*

Estimating harvest rates requires estimates of stock specific catches in terminal and non-terminal fisheries as well as good estimates of escapement. Estimation of stock specific catches has been difficult, but results from the WASSIP mixed stock analyses will improve the ability to characterize the composition of mixed catches. The precision and bias of escapement estimates vary by species, area and stock. Assessments of sockeye runs in the WASSIP area are generally complete and use precise and accurate escapement enumeration methods (weirs and towers). There are, however, some areas where assessments are incomplete or based on imprecise methods (e.g. aerial counts). In general, assessments of chum salmon runs in the WASSIP area are incomplete and based on less precise aerial counts. However, efforts to develop partial escapement enumeration using weirs, sonar, and mark/recapture experiments can be combined with longer-term escapement indices to update run reconstructions for some of these stocks.

PM: Could the uncalibrated escapements be remedied with some mark-recapture experiments between now and the time that the WASSIP analysis is complete?

DE: Some mark-recapture experiments were done for Chinook salmon in AYK.

Dani Evenson: There is interest in conducting more mark-recapture experiments for chum, but there were delayed mortality issues and possibly sampling bias by gear type (drift net or wheel). The experiments would not be on a relevant timeframe for WASSIP.

MA: Is there a way to use the harvest estimates to reconstruct run escapements?

DE: Somewhat. The auxiliary information that can be applied to escapement estimates is better than nothing.

PM: This is a very important issue that needs to be addressed between Doug and Milo, there is a lot of money that is going into the genetic components and this is equally important.

Project timeline - *Bill Templin*

1. The project is on schedule.
2. SNP development is continuing.
3. Doug Eggers is determining which samples to run and assign.

4. Jim Jasper will continue to work on algorithms.
5. June 2012 is the final deadline so the project report needs to be done earlier than that.

4. *Review of minutes from April 2009*

- Page 4. Pat Martin wanted to make sure that the issue of pooling samples is still open and in the discussion. In particular, do we restrict ourselves to the sampling strata that were originally designed? Also, will there be the opportunity to run additional samples either from the 2009 sampling season or from areas where more samples were collected than anticipated?
  o Bill Templin clarified that there are a set number of samples that are being planned for extraction and any samples goals that were not met could be used for other fish to be extracted and analyzed. Additional funding should be secured by April of 2010.

- Page 8. #4 September 24, **2008** correct

- Minutes were approved by the AP.

5. *Meeting frequency and dates.*

Eric Volk proposed pre-arranging approximate dates for future AP and TC meetings so that people can plan well in advance.
- Middle of the week in April and September was suggested, based on timing of the previous meetings.

- Pat Martin noted that it is very important to have the technical committee at the meetings, so they should be planned around their schedules.

Eric Volk proposed Wednesday April 21st 2010 and roughly September 22nd 2010 for the next two meetings.
- No objections were raised.

6. *Attendance and Quorum*

Eric Volk clarified that decisions are made by the AP members present at each meeting. Members may designate alternates in their absence.

There was agreement among AP members that this is acceptable.

7. *Meeting adjourned at 15:50.*