

Western Alaska Salmon Stock Identification Program

Technical Document:¹
14

1
2 **Title:** Sockeye salmon baseline based upon 96 SNPs **Version:** 2.0
3 **Authors:** T. Dann, C. Habicht, J. Jasper, L. Fox, H. Hoyt, H. Hildebrand, E. Lardizabal, P. Kuriscak, Z.
4 Grauvogel, and W. Templin
5 **Date:** January 27, 2012

6 **Abstract**

7 Uncertainty about the magnitude, frequency, location, and timing of the non-local harvest of
8 sockeye and chum salmon was the impetus for the Western Alaska Salmon Stock Identification
9 Program. The program was designed to use genetic data in mixed stock analysis to reduce this
10 uncertainty. A baseline of allele frequencies is required for use in mixed stock analysis to
11 estimate the stock of origin of harvested fish. This technical document describes the
12 methodology we used to understand the population genetic structure among sockeye salmon
13 populations within the study area and to build and test a baseline for use in mixed stock analysis
14 of sockeye salmon. Of the 41,406 individuals from 472 collections selected to be genotyped, the
15 final baseline was composed of 38,193 individuals from 439 collections representing 290
16 populations. Average population sample size was 131 individuals. Correct allocations for proof
17 tests averaged 0.96, ranged from 0.82 to 0.99, and 22 of the 24 proof tests met our goal of 90%
18 correct allocation. Correct allocations for escapement tests averaged 0.93, ranged from 0.49 to
19 0.99, and 26 of the 30 escapement tests met our goal of 90% correct allocation. We believe this
20 baseline will provide accurate and precise estimates of stock composition in WASSIP fisheries.

21 **Introduction**

22 Sockeye salmon (*Oncorhynchus nerka*) are Alaska's most commercially valuable salmonid, and
23 the majority originate from and are harvested in Western Alaska (Eggers and Carroll 2011;
24 Bugaev et al. 2008). The combination of sockeye salmon life history, migratory pathways and
25 the geography of Western Alaska create the potential for the harvest of non-local populations as
26 they return to natal streams. While a majority of the harvest of sockeye salmon in Western
27 Alaska occurs in terminal fisheries, where non-local harvest is minimal (e.g., Bristol Bay; Dann
28 et al. 2009), the harvest of non-local populations does occur and can bias estimates of total run
29 and stock productivity. The relative impact of this bias depends on population size: less
30 abundant populations are more impacted by their non-local harvest than very abundant
31 populations. Uncertainty about the magnitude, frequency, location, and timing of this non-local
32 harvest was the impetus for the Western Alaska Salmon Stock Identification Program

¹ This document serves as a record of communication between the Alaska Department of Fish and Game Commercial Fisheries Division and the Western Alaska Salmon Stock Identification Program Technical Committee. As such, these documents serve diverse ad hoc information purposes and may contain basic, uninterpreted data. The contents of this document have not been subjected to review and should not be cited or distributed without the permission of the authors or the Commercial Fisheries Division.

33 (WASSIP). WASSIP was designed to use genetic data in mixed stock analysis (MSA) to try to
34 reduce this uncertainty.

35 The use of MSA to estimate the stock composition of harvested fish requires a comprehensive
36 baseline of genetic information from the stocks involved, generally allele frequencies. This is
37 accomplished by comparing genotypes of fish of unknown origin from the harvest to the allele
38 frequencies in spawning populations that potentially contribute to the harvest. A baseline is
39 defined by two components: populations of individuals and the genetic markers for which they
40 have been genotyped. This document describes the baseline the Gene Conservation Laboratory
41 (GCL) has built for sockeye salmon for use in WASSIP. It comprises populations ranging from
42 Salmon Lake on the Seward Peninsula to Bering Lake near Cape Suckling (~6,000 km), and 96
43 single nucleotide polymorphisms (SNPs) chosen specifically for WASSIP (Technical Document
44 6).

45 This baseline differs from the baseline previously presented to the WASSIP Advisory Panel (AP)
46 and Technical Committee (TC; Technical Document 5) in three primary ways: 1) the number and
47 range of populations included; 2) the SNPs assayed in these populations; and 3) the methods
48 used to build the baseline. In an effort to reduce uncertainty in stock composition estimates, the
49 AP stipulated that the baseline contain 96 SNPs. As part of this stipulation and to meet
50 budgetary constraints we genotyped a subset of populations from our library of tissues. This
51 subset was chosen in an attempt to gain the greatest representative value from our genotyping
52 efforts for a given cost. The original baseline included populations from throughout the Pacific
53 Rim, whereas the new baseline has more representation within Western Alaska, but has reduced
54 range. The original baseline contained 45 SNPs that were ascertained from various sources
55 while the new baseline contains 96 SNPs chosen specifically for WASSIP because they indicated
56 greater potential for distinguishing reporting groups in the WASSIP study area (Technical
57 Document 6). We have also developed new methods of analyzing a baseline, particularly for the
58 way we handle potentially linked loci, and testing the reporting groups suggested in Technical
59 Document 11 and agreed upon at the September 2010 meeting of the AP and TC. We are
60 currently reanalyzing this baseline for quality control, and it is possible, but not expected, that
61 some results presented here may change.

62 Three goals of this technical document are to describe: 1) the methodology we used to
63 understand the population genetic structure and build a baseline for use in MSA; 2) the genetic
64 structure among sockeye salmon populations within the WASSIP study area; and 3) the
65 performance of the baseline for WASSIP.

66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104

Methods

Tissue Sampling

Baseline collections

We collected baseline samples from spawning populations of sockeye salmon ranging from Russia to the Columbia River drainage to compile our library of tissues. Target sample size for baseline collections was 95 individuals to achieve acceptable precision for estimating allele frequencies (Allendorf and Phelps 1981; Waples 1990a) and to accommodate our genotyping platform. From this broad range of collections we chose a subset to represent sockeye salmon baseline populations for WASSIP.

Selection of baseline collections to genotype

We selected a subset of collections to include in the WASSIP baseline to reflect four goals: to represent 1) population abundance; 2) genetic diversity (Technical Document 5); 3) geographic coverage of populations; and 4) among-year variation of allele frequencies within populations. We attempted to find the greatest representative value with an economical approach by balancing the maximum representation and cost of choosing fish from every collection and the minimum representation and cost of choosing fish from a very small subset of collections. We restricted the range of collections for WASSIP to Cape Prince of Wales to Cape Suckling because we do not expect individuals from populations outside this range to be present in WASSIP area sockeye salmon fisheries.

Escapement collections

We collected samples from the escapement of sockeye salmon to rivers within the WASSIP area to test the baseline with. These were commonly collected at escapement enumeration sites, which were generally located well below spawning grounds, but above the tidal influence in each system and most likely only captured fish destined to spawn within the river.

Laboratory Analysis

Assaying genotypes

We extracted genomic DNA from tissue samples using a DNeasy® 96 Tissue Kit by QIAGEN® (Valencia, CA). We screened 96 SNP markers (Technical Document 6) using Fluidigm® 96.96 Dynamic Arrays (<http://www.fluidigm.com>). The Fluidigm® 96.96 Dynamic Array contains a matrix of integrated channels and valves housed in an input frame. On one side of the frame are 96 inlets to accept the sample DNA from individual fish and on the other are 96 inlets to accept the assays for 96 SNP markers. Once in the wells, the components are pressurized into the chip using the IFC Controller HX (Fluidigm). The 96 samples and 96 assays are then systematically combined into 9,216 parallel reactions. Each reaction is a mixture of 4µl of assay mix (1x DA Assay Loading Buffer (Fluidigm), 10x TaqMan® SNP Genotyping Assay (Applied Biosystems), and 2.5x ROX (Invitrogen)) and 5µl of sample mix (1x TaqMan® Universal Buffer (Applied Biosystems), 0.05x AmpliTaq® Gold DNA Polymerase (Applied Biosystems), 1x GT Sample Loading Reagent (Fluidigm) and 60-400ng/µl DNA) combined in a 7.2nL chamber. Thermal cycling was performed on an Eppendorf IFC Thermal Cycler as follows: 70°C for 30 min for

105 “Hot-Mix” step, initial denaturation of 10 min at 96°C followed by 40 cycles of 96° for 15 s and
106 60° for 1 min. The Dynamic Arrays were read on a Fluidigm® EP1™ System or BioMark™
107 System after amplification and scored using Fluidigm® SNP Genotyping Analysis software.

108 Assays that failed to amplify on the Fluidigm system were reanalyzed on the Applied Biosystems
109 platform. Each reaction on this platform was performed in 384-well reaction plates in a 5µL
110 volume consisting of 5-40ng/µl of template DNA, 1x TaqMan® Universal PCR Master Mix
111 (Applied Biosystems), and 1x TaqMan® SNP Genotyping Assay (Applied Biosystems).
112 Thermal cycling was performed on a Dual 384-Well GeneAmp® PCR System 9700 (Applied
113 Biosystems) as follows: an initial denaturation of 10 min at 95°C followed by 50 cycles of 92°C
114 for 1s and annealing/extension temperature for 1 min. The plates were scanned on an Applied
115 Biosystems Prism 7900HT Sequence Detection System after amplification and scored using
116 Applied Biosystems’ Sequence Detection Software (SDS) version 2.2.

117 Genotypes produced on both platforms were imported and archived in the Gene Conservation
118 Laboratory Oracle database, LOKI.

119 *Laboratory quality control*

120 We conducted a quality control analysis (QC) to identify laboratory errors and to measure the
121 background discrepancy rate of our genotyping process. The QC analyses were performed by
122 staff not involved in the original genotyping. We applied four methods to the QC depending on
123 the type of collection and when it was genotyped. We have termed these the “Old”, “Assay”,
124 “39” and “New” QC methods.

125 The “Old” QC method was how we conducted QC prior to WASSIP. This method consists of re-
126 genotyping 8% of the fish genotyped in the original project using the same DNA extraction for
127 the same SNPs assayed in the original project. Discrepancy rates were calculated as the number
128 of conflicting genotypes, divided by the total number of genotypes compared. These discrepancy
129 rates describe the difference between original project data and QC data for all SNPs and are
130 capable of identifying assay plate errors but cannot detect DNA extraction plate errors (rotations,
131 etc.) since they are based upon the same extractions.

132 The “39” QC method compared new and old genotypes for the 39 SNPs common to our current
133 and previous baselines (Technical Document 5). Since we assayed collections for all 96 SNPs at
134 once, we were able to compare genotypes for these SNPs for 100% of individuals in a collection.
135 Discrepancy rates were calculated as above; these rates describe the difference between our old
136 data for these 39 SNPs and new data for these same SNPs and are capable of identifying errors
137 associated with these SNPs but cannot detect DNA extraction errors since they are based upon
138 the same extractions.

139 The “Assay” QC method compared all 96 SNPs for original project genotypes with QC
140 genotypes based upon the same DNA extraction. We instituted this QC method as a complement
141 to the “39” method since the “39” method is incapable of detecting errors associated with the 57
142 new SNPs on each assay plate. Errors associated with these new loci were detected by
143 genotyping the 96 loci from previously-extracted DNA on one chip of 96 previously-genotyped
144 and QC’ed individuals every time an assay tray was assembled. The new genotypes from these
145 96 fish were then compared with the genotypes in the database to ensure that the assay tray was

146 assembled without error. Discrepancy rates were calculated as above; these rates describe the
147 difference between original project data and QC data for all SNPs but are based on the same
148 DNA plate and so are incapable of detecting DNA plate errors.

149 The “New” QC method is our current QC method and consists of re-extracting 8% of project fish
150 and genotyping them for the same SNPs assayed in the original. Discrepancy rates were
151 described as above; these rates describe the difference between original project data and QC data
152 for all SNPs and are capable of identifying extraction, assay plate, and genotyping errors. This
153 QC method is the best representation of the error rate of our current genotype production.

154 For all QC methods, assuming that the discrepancies among analyses were due equally to errors
155 during the original genotyping and during quality control, error rates in the original genotyping
156 can be estimated as half the rate of discrepancies.

157 *Statistical Analysis*

158 *Data retrieval and quality control*

159 We retrieved genotypes from LOKI and imported them into *R* (R Development Core Team
160 2010). All subsequent analyses were performed in *R* unless otherwise noted. Prior to statistical
161 analysis, we performed three analyses to confirm the quality of the data used. First we identified
162 SNP markers that were invariant in all individuals. We excluded these markers from further
163 statistical analyses.

164 Second, we removed individuals that were missing substantial genotypic data from further
165 analyses. We used the 80% rule (Dann et al. 2009) to exclude individuals missing genotypes for
166 20% or more of loci because these individuals likely have poor-quality DNA. The inclusion of
167 individuals with poor-quality DNA might introduce genotyping errors into the baseline and
168 reduce the accuracies of mixed stock analyses.

169 The final data confirmation analysis identified individuals with duplicate genotypes and removed
170 them from further analyses. Duplicate genotypes can occur as a result of sampling or extracting
171 the same individual twice, and were defined as pairs of individuals sharing the same alleles in 5
172 of 6 loci screened (80 of the 96 SNPs in this study). One individual from each duplicate pair was
173 removed from further analyses.

174 *Hardy-Weinberg Equilibrium*

175 After calculating allelic frequencies for each locus, we tested observed genotype frequencies for
176 each baseline collection for conformance to Hardy-Weinberg expectations (HWE) at each locus
177 by Monte Carlo simulation with 10,000 iterations using the *adegenet* package (Jombart 2008).
178 We combined probabilities for each collection across loci using Fisher’s method (Sokal and
179 Rohlf 1995) and removed collections that departed significantly from HWE after correcting for
180 multiple tests with Bonferroni’s method ($\alpha = 0.05 / \#$ of collections) from subsequent analyses.

181 *Pooling collections into populations*

182 When appropriate we pooled collections to obtain better estimates of allele frequencies following
183 a step-wise protocol. First, we pooled collections from the same geographic location, sampled at

184 similar calendar dates but in different years, as suggested by Waples (1990b). We then tested for
185 differences in allele frequencies between pairs of geographically proximate collections that were
186 collected at similar calendar dates and might represent the same population. We used Fisher's
187 exact test (Sokal and Rohlf 1995) of allele frequency homogeneity and based our decisions on a
188 summary across loci using Fisher's method. When these tests indicated no difference between
189 collections ($P > 0.01$), we pooled them. When these pooled collections were near other
190 collections we followed the same protocol until we found significant differences between the
191 pairs of collections being tested. After this pooling protocol, we considered these final
192 collections to be populations. Finally, we tested populations for conformance to HWE following
193 the same protocol described above to ensure that our pooling was appropriate, and that tests for
194 linkage disequilibrium would not result in falsely positive results due to departure from HWE.

195 *Removal of collections from the baseline*

196 We removed collections from further analysis for different reasons. These reasons included
197 collections not meeting our desired minimum sample size of 75 individuals and not pooling with
198 others that were appropriate to pool with. Similarly, we removed collections that did not pool
199 with geographically close collections and lacked reliable metadata to discern their exact sample
200 date and location. We removed collections from the Yukon River sub-regional group following
201 the recommendation of the AP at the March 2011 meeting. We also removed collections that
202 were believed to be mixtures of multiple populations and not representative of single, spawning
203 populations. We removed hatchery broodstocks collections that were believed to not represent
204 either the hatchery or original population. Finally, we identified collections of escapement
205 samples that were previously used as baseline but were no longer needed to represent spawning
206 populations because more representative collections replaced them. We subsequently used these
207 collections as tests of the baseline.

208 *Linkage disequilibrium*

209 We tested for linkage disequilibrium between each pair of nuclear SNPs in each population to
210 ensure that subsequent baseline and mixed stock analyses would be based on independent
211 markers. We used the program *Genepop* version 4.0.11 (Rousset 2008) with 100 batches of
212 5,000 iterations for these tests. We summarized the frequency of significant linkage
213 disequilibrium between pairs of SNPs ($P < 0.05$), and further investigated pairs that exhibited
214 linkage in a substantial number of populations. We considered pairs to be linked if they
215 exhibited linkage in more than half of all populations. We also considered pairs to be linked if
216 they exhibited linkage in less than half of populations but in substantially more populations than
217 a majority of SNP pairs. We defined "substantially more" by examining a histogram of the
218 frequency of the number of populations that pairs were linked in. We also examined the
219 correlation coefficient r between the first alphabetical allele in each linked pair of SNPs in each
220 population to visualize the pattern of linkage across the geographic range of the baseline. We
221 used the *BRugs* package to estimate the error around these correlation coefficient estimates
222 (Thomas et al. 2006). [Note: This analysis was conducted before we included the collection of
223 river-spawning sockeye salmon from the Middle Fork of the Goodnews River. It is possible, but
224 unlikely, that the results of this analysis will change in the baseline reanalysis that includes this
225 collection.]

226 We either removed one of the linked SNPs or combined the pair into a composite, haploid
 227 marker in further analyses if the pattern of linkage provided information useful for mixed stock
 228 analysis. We used f_{ORCA} as our measure of information. f_{ORCA} assesses the rate of correct
 229 allocation of simulated individuals to defined reporting groups based upon the markers in
 230 question (Rosenberg 2005). Because combinations of alleles from two or more markers can
 231 exist in more forms than single markers (9 possible phenotypes vs. 4 alleles for a pair of SNPs),
 232 composite markers generally have higher f_{ORCA} values than the single markers that form them.
 233 Simple comparisons of these values would always suggest combining linked pairs into composite
 234 markers. However, there is a cost associated with composite markers as estimates of 8 phenotype
 235 frequencies are less precise than estimates of 1 allele frequency at 2 loci for a given sample size.

236 To account for this cost, and to ensure that we combined only SNP pairs that provided
 237 significantly more information than the single SNPs in question, we compared the difference
 238 between f_{ORCA} values of the composite marker and the single SNP with the greater f_{ORCA} value in
 239 the pair ($\Delta = f_{ORCA-pair} - \max(f_{ORCA-single1}, f_{ORCA-single2})$). This difference (Δ) was our test statistic.
 240 Since we did not know the distribution of Δ , we conducted a sampled randomization test (Sokal
 241 and Rohlf 2005). We randomly selected 1,000 SNP pairs, calculated Δ for each pair to
 242 empirically define the test statistic distribution, and set the 90th quantile of the distribution as a
 243 critical value (Δ_{90}). We then either combined linked SNPs into composite, haploid markers if Δ
 244 was greater than this critical value or dropped the SNP with the lower f_{ORCA} value if Δ was less
 245 than the critical value. [Note: The f_{ORCA} analysis was conducted before it was agreed upon by
 246 the AP to collapse the Aleutian Islands sub-regional group within the NW District-Black Hills
 247 sub-regional group, so it was based upon 25 groups instead of the final 24 groups. It is possible,
 248 but unlikely, that the results of this analysis will change in the baseline reanalysis that is based
 249 upon 24 groups.]

250 *Analysis of genetic structure*

251 *Analysis of temporal variance*

252 We examined the among-year temporal variation of allele frequencies with a hierarchical, three-
 253 level Analysis of Variance (ANOVA). We treated the temporal samples as sub-populations
 254 based on the method described in Weir (1996). This method allowed the quantification of the
 255 sources of total allelic variation and permitted the calculation of the between-collection
 256 component of variance and the assessment of its magnitude relative to the between-population
 257 component of variance. This analysis was conducted using the software package *GDA* (Lewis
 258 and Zaykin 2001).

259 *Visualization of genetic distances*

260 We visualized pairwise F_{ST} estimates among collections from the final set of independent
 261 markers estimated with the package *hierfstat* (Goudet 2006). We constructed 1,000 bootstrapped
 262 Neighbor-Joining (NJ) trees by resampling loci with replacement to assess the stability of tree
 263 nodes across markers. We plotted the consensus tree with the *ape* package (Paradis et al. 2004).
 264 These trees provided insight into the variability of the genetic structure of these populations.

265 *Baseline evaluation for MSA*

266 We evaluated the utility of the baseline for MSA by assessing the identifiability of our sub-
267 regional reporting groups. To do this we conducted two types of tests. The first were “100%
268 proof tests”, where we sampled 200 individuals without replacement from each sub-regional
269 reporting group and analyzed them as a mixture against the reduced baseline. These tests
270 provided an indication of the power of the baseline for MSA under the assumption that all the
271 populations from a reporting group were represented in the baseline. The second were
272 “escapement tests”, where we analyzed samples of the escapement to a river within a sub-
273 regional reporting group as an independent mixture against the full baseline. These tests
274 assumed that the fish sampled at escapement enumeration projects were destined to spawn
275 upstream from where they were sampled.

276 For both types of tests, we summarized the results following the dynamic reporting groups
277 protocol described in Technical Document 11. For example, for tests of a sub-regional reporting
278 group within the Bristol Bay region, we provided estimates to all 9 Bristol Bay sub-regional
279 groups as well as the 6 other regional groups, but for tests of sub-regional groups outside of
280 Bristol Bay only a regional Bristol Bay estimate was provided.

281 Stock compositions of these test mixtures were estimated with the program *BAYES* (Pella and
282 Masuda 2001). The Bayesian model implemented by *BAYES* places a Dirichlet distribution as the
283 prior distribution for the stock proportions, and the parameters for this distribution must be
284 specified. We defined prior parameters for each reporting group to be equal (i.e., a “flat” prior)
285 with the prior for each reporting group subsequently divided equally to populations within that
286 reporting group. We set the sum of all prior parameters to 1 (prior weight), which is equivalent
287 to adding one fish to each mixture (Pella and Masuda 2001). We ran five independent Markov
288 Chain Monte Carlo (MCMC) chains of 40,000 iterations with different starting values and
289 discarded the first 20,000 iterations to remove the influences of the initial start values. We
290 combined the second half of each chain to form the posterior distribution and tabulated mean
291 estimates and 90% credibility intervals from a total of 100,000 iterations. We also assessed the
292 among-chain convergence of these estimates using the Gelman-Rubin shrink factor, which
293 compares variation within a chain to the total variation among chains (Gelman and Rubin 1992).
294 If a shrink factor for any stock group estimate was greater than 1.2 we reanalyzed the mixture
295 with 80,000-iteration chains following the same protocol. We repeated this procedure for each
296 reporting group mixture. A critical level of 90% correct allocation was used to determine if the
297 reporting group was acceptably identifiable (Seeb et al. 2000). We presented these results as
298 barplots using the *gplots* package (Warnes 2010).

299 **Results**

300 *Tissue Sampling*

301 *Baseline collections*

302 We compiled a library of quality baseline tissues from 94,554 sockeye salmon in 849 collections.
303 These samples were collected from 1991 through 2010 and ranged from the Kamchatka
304 Peninsula in Russia to Washington State. We chose to restrict the area of our baseline to include
305 only those populations possibly present in WASSIP mixtures, so we chose a subset of collections

306 from this library ranging from Salmon Lake on the Seward Peninsula to Bering Lake near Cape
307 Suckling (Figures 1 and 2). These collections spanned the years 1991-2010 (Table 1) and totaled
308 41,406 sockeye salmon from 472 collections.

309 A majority of the baseline collections chosen for WASSIP were located in the WASSIP study
310 area (22,950 individuals; 55% of total). The Bristol Bay regional reporting group comprised the
311 majority of these individuals, with 12,654 chosen to be genotyped for 96 SNPs (31% of total).

312 *Escapement samples*

313 A total of 5,921 sockeye salmon from 30 collections of the escapement to 14 different rivers
314 within the WASSIP study area were successfully genotyped (Table 2). These samples were
315 collected between 2001 and 2010 and provided tests of 14 of the 24 sub-regional reporting
316 groups.

317 *Laboratory Analysis*

318 *Assaying genotypes*

319 We genotyped all individuals selected from baseline and escapement samples for 96 SNPs
320 (Tables 1, 2 and 3). A majority of these genotypes were produced on the Biomark platform. The
321 number of individuals genotyped from baseline collections ranged from 6 to 190 and averaged
322 88 individuals. Within the WASSIP baseline collections, the number of individuals genotyped
323 ranged from 30 to 190 and averaged 90 individuals. Baseline collections from East of WASSIP
324 ranged from 6 to 190 individuals with an average of 85 (Table 1).

325 *Quality control*

326 Quality control demonstrated a low overall discrepancy rate of 0.33% for WASSIP sockeye
327 salmon baseline collections (Table 4). A majority of discrepancies were between homozygotes
328 and heterozygotes, and very few homozygote-homozygote discrepancies were observed (total of
329 295 out of 1,333,444 genotypes compared). Discrepancy rates for “Old”, “New”, “39”, and
330 “Assay” QC method collections were 0.24%, 0.13%, 0.34%, and 0.14%, respectively. Baseline
331 collections of sockeye salmon were genotyped with a process that produced genotypes with an
332 error rate of 0.12% for “Old”, 0.07% for “New”, 0.18% for “39”, 0.07% for “Assay” QC method
333 collections and an overall rate of 0.17% if equal error rates in the original and QC genotyping
334 process are assumed.

335 *Statistical Analysis*

336 *Data retrieval and quality control*

337 All SNPs were variant for populations in the WASSIP study area. A total of 583 individuals
338 from WASSIP baseline collections were missing genotypes from greater than 20% of the loci (19
339 SNPs) and were removed from further analyses (Table 1). Three hundred and eighty-four of
340 these were from within the WASSIP study area and 199 were from East of WASSIP collections.
341 For baseline collections within the WASSIP area, 27 individuals were removed from the North
342 of Kuskokwim River reporting group (5.24%), 48 individuals from Kuskokwim Bay (1.43%),

343 273 individuals from Bristol Bay (2.17%), 17 individuals from North Peninsula (0.45%), 6
344 individuals from South Peninsula (0.71%), and 13 individuals from Chignik (0.72%).

345 There were 127 duplicate individuals identified in WASSIP baseline collections, 91 in the
346 WASSIP area collections and 36 from East of WASSIP collections. For baseline collections
347 within the WASSIP study area, 2 duplicate individuals were removed from North of Kuskokwim
348 River (0.38%), 19 individuals from Kuskokwim Bay (0.57%), 40 individuals from Bristol Bay
349 (0.32%), 25 individuals from North Peninsula (0.66%), 0 individuals from South Peninsula
350 (0.00%), and 4 individuals from Chignik (0.27%).

351 *Hardy-Weinberg Equilibrium*

352 Five baseline collections deviated from HWE and were removed from further analyses (Table 1).
353 There was no geographic pattern to the deviation from HWE. Similarly, we observed no pattern
354 in the deviation from HWE among loci.

355 *Pooling collections into populations*

356 After conducting chi-square tests of allele frequency homogeneity and identifying collections
357 that indicated no differences in allele frequency from one another, we pooled those collections
358 that were appropriate to pool. The geographic distance between pooled collections was often
359 small (< 5 km).

360 *Removal of collections from the baseline*

361 In our pooling tests, we discovered 7 collections that did not pool with others and were too small
362 to include in the baseline (2 of these lacked reliable metadata and were not pooled; Table 1).
363 Nine collections lacked reliable metadata or were believed to be mixtures of populations and
364 were not included in pooling tests. We removed 3 collections from the Yukon River sub-
365 regional reporting group following the AP decision at the March 2011 meeting. The Main Bay
366 collection (SMAN91; collection # 426) came from multiple broodstocks for a hatchery and was
367 thought to not represent either the hatchery or original population and was removed from further
368 analyses. We identified 8 collections of escapement samples that were previously used as
369 baseline but were no longer needed to represent spawning populations because more
370 representative collections replaced them. These collections were subsequently used as tests of
371 the baseline (Table 2), except for 1 small collection (SGOOD91, n=46; Table 1).

372 Of the 41,406 individuals from 472 collections selected to be genotyped, the final baseline was
373 composed of 38,193 individuals from 439 collections representing 290 populations. Average
374 population sample size was 131 individuals (range: 69-473; Table 1).

375 *Linkage disequilibrium*

376 Three SNP pairs were significantly linked in a majority of WASSIP-area sockeye salmon
377 populations in tests for LD. Two pairs were linked in greater than half of all populations
378 (*One_MHC2_190* & *One_MHC2_251*, $P < 0.05$ for 71% of populations; *One_GPDH-201* &
379 *One_GPDH2-1872*, 57% of populations), while one other pair was linked in substantially more
380 populations than most other pairs (*One_Tf_ex11-750* & *One_Tf_in3-182*, 41% of populations;
381 Figures 3 and 4). The 90% critical value of the f_{ORCA} difference distribution (Δ_{90}) was 0.0091,

382 which was greater than Δ for two of the linked pairs (*One_GPDH-201* & *One_GPDH2-1872*: Δ
 383 = 0.0081; *One_Tf_ex11-750* & *One_Tf_in3-182*: Δ = 0.0025; Table 5 and Figure 5). While Δ_{90}
 384 was slightly greater than Δ for the third pair (*One_MHC2_190* & *One_MHC2_251*: Δ = 0.0089),
 385 this difference was small (0.0002) and a further examination of the correlation coefficient r of
 386 alleles at the two SNPs suggested a useful pattern in the linkage across reporting groups (Figures
 387 6 and 7). So we dropped the SNP with the lowest f_{ORCA} value in two pairs (*One_GPDH-201* and
 388 *One_Tf_ex11-750*) and combined the MHC SNPs.

389 *Analysis of genetic structure*

390 *Analysis of temporal variance*

391 We included 125 collections belonging to 59 populations in the analysis of temporal variance
 392 using the three-level ANOVA. These ranged from the Necons River in the Kuskokwim River
 393 drainage to Kushtaka Lake near Cape Suckling, and included 50 collections from 24 populations
 394 in the WASSIP study area (Table 1). The ANOVA indicated that the variation between
 395 populations was 90 times greater than the amount of temporal variation between years within
 396 populations (between collections, $\sigma_S = 0.041$; between populations, $\sigma_P = 3.701$; ratio 90.155)

397 *Visualization of genetic distances*

398 The NJ tree of pairwise F_{ST} indicated that sockeye salmon from the East of WASSIP group
 399 exhibit the greatest diversity (Figure 8), and that substantial genetic structure existed within
 400 WASSIP-area groups for use in MSA (Figures 9 and 10). We observed high concordance among
 401 loci for many population groupings, in particular for populations spawning in lacustrine
 402 environments (e.g., Upper Kuskokwim River lakes, Lake Clark, Alagnak), but also for some
 403 riverine populations (e.g., Kuskokwim River populations; Figure 9). Interestingly, some
 404 population groupings were defined more by life history and habitat usage than by geography
 405 (riverine sockeye salmon from the Kuskokwim and Nushagak drainages; Figure 9). The tree of
 406 genetic distances indicated weak structuring among small populations of the North and South
 407 Peninsula (e.g., Nelson River, NW District-Black Hills, South Peninsula) but more defined
 408 structure for some of the larger populations of lake-type sockeye salmon from the North
 409 Peninsula (e.g., Bear).

410 *Baseline evaluation for MSA*

411 Correct allocations for proof tests averaged 0.96 and ranged from 0.82 to 0.99 (Tables 6-10;
 412 Figure 11). Twenty-two of the 24 proof tests met our goal of 90% correct allocation.

413 For Norton Sound, South Peninsula, and East of WASSIP reporting groups, correct allocations in
 414 the proof tests were 0.95, 0.99, and 0.99, respectively (Table 6; Figure 11). Within Kuskokwim
 415 Bay, correct allocations averaged 0.89 and ranged from 0.82 to 0.98 across the 3 tests (Table 7;
 416 Figure 11). One proof test did not reach the 90% correct allocation level: Goodnews (correct
 417 allocation=0.82, misallocation to Bristol Bay=0.15; Table 7). Within the Bristol Bay regional
 418 reporting group, proof test correct allocations averaged 0.96 and ranged from 0.82 to 0.99 across
 419 the 9 tests (Table 8). One proof test did not reach the 90% correct allocation level: Egegik
 420 (correct allocation=0.82, misallocation to Ugashik=0.18; Table 8). Within the North Peninsula,
 421 proof test correct allocations averaged 0.97 and ranged from 0.93 to 0.99 across the 7 tests

422 (Table 9). Within Chignik, correct allocations to the 2 Chignik regional reporting groups (Black
423 and Chignik lakes) were both 0.99 (Table 10).

424 Correct allocations for escapement tests averaged 0.93 and ranged from 0.49 to 0.99 (Tables 11,
425 12 and 13; Figure 12). Twenty-six of the 30 escapement tests met our goal of 90% correct
426 allocation.

427 Within Kuskokwim Bay, correct allocations in the escapement tests averaged 0.89 and ranged
428 from 0.49 to 0.99 with 6 of the 8 tests correctly allocating at 0.94 or greater (Table 11). Two
429 escapement tests did not meet the 90% correct allocation goal: Kwethluk River weir 2007
430 (correct allocation=0.49, misallocation to Nushagak River populations within the Bristol Bay
431 reporting group=0.48; Table 11), and Goodnews River weir 2007 (correct allocation=0.83,
432 misallocation to Togiak populations within the Bristol Bay reporting group=0.15; Table 11).

433 Within Bristol Bay, correct allocations in the escapement tests averaged 0.95 and ranged from
434 0.79 to 0.99 across the 20 tests (Table 12; Figure 12). Two escapement tests did not meet the
435 90% correct allocation goal: Togiak Subsistence 2008 (correct allocation=0.79, misallocation to
436 Goodnews River populations within the Kuskokwim Bay reporting group=0.21), and the
437 Nushagak Radio Telemetry 2005 test (correct allocation=0.84, misallocation to Kuskokwim
438 River populations within the Kuskokwim Bay reporting group=0.13; Table 12).

439 Within Chignik, correct allocations for the two escapement tests were 0.98 and 0.92 for the early
440 (6/14-21/2010) and late (7/23-30/2010) runs, respectively (Table 13).

441 **Discussion**

442 We set out to describe the methodology used to build the sockeye salmon baseline for WASSIP,
443 investigate the genetic structure among sockeye salmon populations in the WASSIP study area,
444 and test the performance of this baseline for use in MSA in WASSIP.

445 *Genetic variation among sockeye salmon in the WASSIP area*

446 The distribution of variation observed in this baseline (Figure 8) is concordant with that
447 previously observed using 45 SNPs (Technical Document 5, Habicht et al. 2010). Much of this
448 diversity is distributed in the East of WASSIP reporting group, but significant structure among
449 reporting groups within the WASSIP area suggest potential high power for MSA.

450 Similar to the previous sockeye salmon baseline, a consistent pattern of fewer heterozygotes than
451 expected was observed in the multi-locus genotypes (Table 3). This is either a signal that not all
452 the “populations” in our baseline consist of randomly interbreeding individuals (i.e., Wahlund
453 effect; Hedrick 2005) or that our genotyping methods are biased against heterozygotes. The lack
454 of deviations from HWE within putative populations indicate that most populations generally
455 follow expectations based on random mating. We observed a reduction of F_{IS} for the 32
456 uncombined, nuclear SNPs common to old (48-SNP chip) and new (96-SNP chip) genotyping
457 efforts from an average of 0.14 in the old baseline (376 populations; Technical Document 5) to
458 0.12 in the new baseline (290 populations), an average difference of 0.02 (data not shown). This
459 reduction indicates that at least some of the heterozygote deficiency is due to scoring bias but
460 that this bias is getting smaller. The 96-SNP chip generally produces more distinct homozygote

461 and heterozygote clusters, which likely explains the reduction in bias against scoring
462 heterozygote genotypes. The fact that the markers that are difficult to score also have the largest
463 F_{IS} values (e.g., *One_STC-410* new baseline 0.22, old 0.23; *One_Zp3b-49* new baseline 0.29, old
464 0.26; data not shown) provides additional support for this hypothesis.

465 *MSA performance*

466 *Conservative tests*

467 Both types of tests of MSA performance of the baseline (proof and escapement) were performed
468 with fewer fish than the sampling goal for WASSIP mixed fisheries strata (400 individuals).
469 These tests also used flat priors. As a result, estimates of correct allocations to reporting groups
470 may be conservative, especially estimates from tests with small sample sizes.

471 Proof tests were performed with 200, and in a few cases 100, individuals to avoid depopulating
472 the baseline for reporting groups represented by fewer individuals (the minimum target size for a
473 reporting group was set at 400 fish; Technical Document 11). Escapement tests were performed
474 with single-year collections to allow for the inspection of year-to-year variation in performance
475 within drainages. These tests always contained fewer than 400 fish, and sometimes as few as 95
476 fish. Results of escapement tests containing fewer than 190 fish should be interpreted with
477 caution.

478 Using a flat prior in baseline evaluation tests is also conservative as we anticipate using an
479 informative prior (sequential prior based on the posterior distribution of similar mixtures) in the
480 MSA of WASSIP mixtures. The use of a flat prior in both the proof and escapement tests is
481 likely to have the most negative impact on the correct assignments for reporting groups that have
482 populations with similar allele frequencies, such as riverine sockeye salmon (see below). We
483 anticipate that an informed prior, such as the sequential prior, will improve the performance of
484 the baseline.

485 *Effect of genetic similarity among riverine sockeye salmon on MSA performance*

486 The Kwethluk River weir test of the Kuskokwim River reporting group had the lowest correct
487 allocation of any of the baseline evaluation tests (49%; Table 11). This weir collection (#23;
488 Table 1) was included in previous baselines, but was removed from this baseline and used as an
489 escapement test because we received more representative baseline collections from spawning
490 aggregations higher in the drainage. However, the collection of riverine sockeye from the
491 Kwethluk River (Collection #21; Table 1) was removed due to small sample size (n=49). The
492 low correct allocation of this test is likely the result of the sample containing many riverine
493 sockeye salmon, the absence of riverine sockeye salmon from the Kwethluk River in the
494 baseline, and the genetic similarity among riverine sockeye salmon populations that has observed
495 elsewhere (e.g., Dann et al. 2009; Technical Document 12; Wood et al. 2008; McPhee et al.
496 2009). These findings are supported by the misallocation to Nushagak populations within the
497 Bristol Bay reporting group, similar misallocations between reporting groups with genetically
498 similar riverine sockeye salmon (e.g., Goodnews River and Togiak; Technical Document 12),
499 and a reciprocal misallocation from the Nushagak reporting group to the Kuskokwim Bay
500 reporting group in escapement test # 15 (Nushagak radio telemetry 2005 – correct
501 allocation=84%, misallocation to Kuskokwim=13%). However, the small contribution of the

502 Kwethluk River to the overall Kuskokwim River escapement (2000-2009 Kwethluk weir
503 average=3,235; Bavilla et al. 2010), combined with the high correct allocations of other tests of
504 this reporting group (98%, 99%, 94%; Table 11), and the small sample size of the test (n=141)
505 suggest that results of this test should be interpreted with caution.

506 The poor performance of two tests of the Goodnews and Togiak reporting groups was previously
507 described in Technical Document 12. This is likely the result of genetic similarity among
508 populations along the drainage divide and the absence of riverine sockeye from the Togiak
509 reporting group in the baseline. Sockeye salmon spawning in the mainstem of the Togiak River
510 are believed to contribute an average of 43.5% of the escapement to the drainage, as estimated by
511 aerial surveys (average of 1988-2007; Salomone et al. 2009). We have since collected riverine
512 sockeye from the mainstem of the Togiak River as well as sockeye from a tributary of the Togiak
513 River that contributes a large portion of the Togiak drainage escapement (9% to Pungokepuk
514 Lake; Salomone et al. 2009). We hope these collections will better represent the genetic
515 relationship between populations in these two reporting groups, and improve MSA performance,
516 but we may not be able to incorporate these collections into the WASSIP baseline.

517 *Baseline representation versus population size affects proof tests*

518 The poor performance of the Egegik proof test is likely due to an artifact of the sampling
519 procedure used in our proof tests, coupled with highly divergent population sizes and a genetic
520 outlier population. The proof test for the Egegik reporting group had a surprisingly low correct
521 allocation (82%; Table 8) given previously reported results and the results of the escapement
522 tests for Egegik (97% and 99%; Table 12). These previous results included proof tests based
523 upon an older baseline comprising fewer collections and 45 SNPs (correct allocation 96%; Dann
524 et al. 2009). The 96 SNPs used in the current baseline are the result of marker selection that
525 specifically included measures to improve MSA distinction between the Egegik and Ugashik
526 reporting groups (Technical Document 6). We believe the lower than expected correct allocation
527 is a result of sampling individuals from an Egegik population (South Becharof, Collection # 179,
528 Table 1) with very divergent MHC frequencies. The contribution to the proof test sample from
529 the population in question was greater than the contribution the population represents to the total
530 escapement to the Egegik River (P. Salomone, pers. comm.). This explains the discrepancy
531 between the low correct allocation of the proof test and high correct allocations of the
532 escapement tests (97% and 99%, Table 12) and reinforces the importance of adequately
533 representing population abundances in a baseline.

534 *Baseline is adequate for WASSIP objectives*

535 We believe that the baseline we have built for sockeye salmon for use in WASSIP meets the
536 goals of accurately describing the genetic structure among populations within the WASSIP area
537 as well as consistently meeting our goal of 90% correct allocation in MSA applications. We are
538 confident in the methods used to build the baseline as well as the product of those methods, and
539 believe that this baseline will provide accurate and precise estimates of stock composition in
540 WASSIP fisheries.

541

Questions for the Technical Committee

542

1. Do you think that this baseline is adequate for use in MSA to meet the goals of WASSIP?

543

544

2. Are there other methods we should apply to the baseline to bring it up to acceptable standards for use in MSA given the goals of WASSIP?

545

546

547

Technical Committee Review and Comments

548

1. Do you think that this baseline is adequate for use in MSA to meet the goals of WASSIP?

549

This is a very impressive baseline and panel of SNPs. Many of the populations are relatively strongly differentiated, which should lead to high precision in stock composition estimates.

550

551

‘Adequate’ is a value-laden word and interpretation depends on context. Most of the populations

552

met the 90% proof test criterion that has been adopted by the project as a key metric. In

553

addition, I agree with the authors that some aspects of the proof tests might be expected to

554

produce conservative results, and actual performance might be better. However, it still could be

555

the case that not all stocks can be resolved with the desired level of precision. As we noted in

556

our original comments (September 2008), some of the goals originally articulated for this project

557

(e.g., detecting very small stock contributions with a very high degree of certainty) were

558

unrealistic based on basic statistical considerations.

559

2. Are there other methods we should apply to the baseline to bring it up to acceptable

560

standards for use in MSA given the goals of WASSIP?

561

Again, ‘acceptable’ is a subjective term. It is probably fair to say that the sampling and

562

analytical efforts that have been expended in producing this baseline are impressive even by the

563

standards that have been set during 30 years of interagency cooperation to produce coastwide

564

genetic baselines for salmon. Although this would not necessarily affect the baseline per se, one

565

thing that could be done to improve the proof tests is to adopt the leave-one-out procedure

566

suggested by Anderson et al. (2008) when using simulations based on samples from real

567

populations. The key issue is to ensure proper cross-validation, and the method used here

568

(dropping some individuals from baseline collections so they can be used in the simulated

569

mixtures) causes a reduction in baseline sample size. In contrast, the Anderson et al. method

570

retains essentially the entire original sample size, which should produce more realistic results. I

571

believe this issue has been recognized and that the problem is that the method has not yet been

572

implemented by the core software. Doing so would not change the results of the actual mixed-

573

stock fishery analyses, but it should provide a more accurate indication of the precision to be

574

expected in those analyses. Another option would be to evaluate performance of the baseline

575 with more realistic simulated mixtures, where the mixture fractions are based on expected
576 contributions to the fishery, as was recently done for chum salmon.

577

578 Other comments:

579 The information regarding Hardy-Weinberg equilibrium in the baseline stocks is confusing and
580 incomplete. At line 352 you say that 5 baseline samples were excluded because they were out of
581 HWE, but no details are given. The implication is that the remaining populations are in
582 equilibrium. But Table 3 shows that there is a consistent deficit of heterozygotes at most loci,
583 with mean $F_{is} = 0.12$ (line 458). What is not clear is how these heterozygote deficiencies were
584 calculated. Was the H_e in Table 3 calculated as the mean of the H_e values for each
585 subpopulation? Or was it calculated using average allele frequencies across all populations in
586 the baseline? The former is relevant to HW evaluations, but I suspect it might be the latter that is
587 shown. Similarly, is the mean F_{is} (0.12) the mean of the F_{is} values from the individual baseline
588 populations? Or was it calculated from the entire dataset pretending it was a single population?
589 If the latter was done, the positive F_{is} is easily explained by the Wahlund effect.

590 Line 364: should include a brief rationale for why the 3 Yukon populations were dropped

591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625

References

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences*. 38:1507-1514.

Bavilla, J., D. Bue, H. Carroll, T. Elison, D. Taylor, J. Estensen and C. Brazil. 2010. 2009 Kuskokwim area management report. Alaska Department of Fish and Game, Fishery Management Report No. 10-56, Anchorage.

Bugaev, A.V., I.I.Glevov, E.V. Golub, K.W. Myers, J.E. Seeb, and M. Foster. 2008. Origin and distribution of sockeye salmon *Oncorhynchus nerka* local stocks in the western Bering Sea in August-October 2006. *Izv. TINRO* 153:88-108.

Dann, T.H., C. Habicht, J.R. Jasper, H.A. Hoyt, A.W. Barclay, W.D. Templin, T.T. Baker, F.W. West, and L.F. Fair. 2009. Genetic stock composition of the commercial harvest of sockeye salmon in Bristol Bay, Alaska, 2006-2008. Alaska Department of Fish and Game, Fishery Manuscript Series No. 09-06, Anchorage.

Eggers, D.M., and A.M. Carroll. 2011. Run forecasts and harvest projections for 2011 Alaska salmon fisheries and review of the 2010 season. Alaska Department of Fish and Game, Special Publication No. 11-03, Anchorage.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7:457–511.

Goudet, J. 2006. hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-4. <http://www.r-project.org>, <http://www.unil.ch/popgen/software/hierfstat.htm>

Habicht, C., L.W. Seeb, K.W. Myers, E.V. Farley, and J.E. Seeb. 2010. Summer-fall distribution of stocks of immature sockeye salmon in the Bering Sea as revealed by single-nucleotide polymorphisms. *Transactions of the American Fisheries Society* 139:1171-1191.

Hedrick, P.W. 2005. *Genetics of populations*, third edition. Jones and Bartlett. Sudbury, MA.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403-1405. doi: 10.1093/bioinformatics/btn129

Lewis, P. O., and Zaykin, D. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>

McPhee, M.V., T.H. Tappenbeck, D.C. Whited, and J.A. Stanford. 2009. Genetic diversity and population structure in the Kuskokwim River drainage support the recurrent evolution hypothesis for sockeye salmon life histories. *Transactions of the American Fisheries Society* 138:1481-1489.

Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.

- 626 Pella, J., and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic
627 characters. *Fishery Bulletin* 99:151–167. BAYES can be obtained from
628 <ftp://ftp.afsc.noaa.gov/sida/mixture-analysis/bayes/>
- 629 R Development Core Team. 2010. R: A language and environment for statistical computing. R
630 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL
631 <http://www.R-project.org/>.
- 632 Rosenberg, N.A. 2005. Algorithms for selecting informative marker panels for population
633 assignment. *Journal of Computational Biology* 12(9):1183-1201.
- 634 Rousset, F. GENEPOP '007: a complete re-implementation of the GENEPOP software for
635 Windows and Linux. 2008. *Molecular Ecology Resources* vol. 8:103-106.
- 636 Salomone, P., S. Morstad, T. Sands, and M. Jones. 2009. Salmon spawning ground surveys in
637 the Bristol Bay Area, Alaska, 2008. Alaska Department of Fish and Game, Fishery
638 Management Report No. 09-42, Anchorage.
- 639 Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R. Z. Davis, L. K. Brannian, and J. E.
640 Seeb. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application
641 to management of populations affected by the Exxon Valdez oil spill. *Transactions of the*
642 *American Fisheries Society* 129(6):1223-1249.
- 643 Sokal, R.R. and F.J. Rohlf. 1995. *Biometry*. 3rd Edition. Freeman, San Francisco, CA. Thomas,
644 A., O'Hara, B., Ligges, U., and Sturtz, S. 2006. Making BUGS Open. *R News* 6 (1), 12-17.
- 645 Thomas, A., B. O'Hara, U. Ligges, and S. Sturtz. 2006. Making BUGS Open. *R News* 6 (1), 12-
646 17.
- 647 Warnes, G. R. 2010. gplots: Various R programming tools for plotting data. R package version
648 2.8.0. <http://CRAN.R-project.org/package=gplots>
- 649 Waples, R. S. 1990a. Conservation genetics of Pacific salmon III. Estimating effective
650 population size. *Journal of Heredity* 81(4):277-289.
- 651 Waples, R. S. 1990b. Temporal changes of allele frequency in Pacific salmon - implications for
652 mixed-stock fishery analysis. *Canadian Journal of Fisheries and Aquatic Sciences* 47(5):968-
653 976.
- 654 Weir, B. 1996. *Genetic Data Analysis* (2nd edition). Sinauer Associates, Inc, Sunderland, MA.
655 445pp.
- 656 Wood, C.C., J.W. Bickham, R.J. Nelson, C.J. Foote, and J.C. Patton. 2008. Recurrent evolution
657 of life history ecotypes in sockeye salmon: implications for conservation and future
658 evolution. *Evolutionary Applications* 1:207-221.

659

Tables

660 Table 1. Regional and sub-regional reporting group, ADF&G collection code, location, collection and population number, collection date, and the
 661 numbers of sockeye salmon used to describe the genetic structure of sockeye salmon in the WASSIP study area and estimate the stock
 662 composition of WASSIP mixed fisheries. The number of individuals includes the number of individuals initially genotyped for the set of 96 SNPs
 663 (Initial), the numbers removed because of missing loci (Missing) and duplicate individuals (Duplicate), and the number of individuals incorporated
 664 into the baseline (Final). Footnotes associated with some ADF&G collection codes indicate the reason they were excluded from the baseline.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
North of Kuskokwim River	Seward Peninsula	SSALM01	Salmon Lake	1	1	8/3/2001	88	5	0	83
		SGLAC04	Glacial Lake	2	2	8/15/2004	190	3	0	187
		SUNA07	Unalakleet River	3	3	8/22/2007	95	5	0	90
Seward Peninsula Total							373	13	0	360
Yukon River	Yukon River	SANDRE05 ^A	Andreafsky River	4		7/12/2005	47	13	1	0
		SANDRE06 ^A		5		6/28/2006	48	1	0	0
		SANDRE08 ^A		6		7/19/2008	47	0	1	0
Yukon River Total							142	14	2	0
North of Kuskokwim River Total							515	27	2	360
Kuskokwim Bay	Kuskokwim River	SUTAK06 ^B	Upper Takotna River	7		2006	40	4	0	0
		SNECO06	Necons River	8	4	8/1/2006	55	0	0	55
		SNECO07		9	4	7/28/2007	95	2	0	93
		STELA03	Telaquana Lake	10	5	8/14/2003	96	0	0	96
		STELA05		11	6	10/4/2005	95	0	0	95
		SKOGR08	Kogruklu River	12	7	8/4/2008	71	0	0	71
		SKOGR01 ^C		13		7/6/2001	96	5	0	0
		SKOGR07 ^C		14		7/24/2007	95	0	0	0
		SCHUKO08	Chukowan River	15	7	8/7/2008	75	0	0	75
		SHOLI08	Holitna River	16	7	8/9/2008	75	0	0	75
		SSALMR06 ^D	Salmon River, Aniak Basin	17		8/2/2006	142	2	0	0
SATSAK09	Atsaksovluk Creek, Aniak Basin	18	8	8/6/2009	95	0	6	89		

665

666 Table 1: Page 2 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Kuskokwim Bay	Kuskokwim River	STULU08	Tuluksak River	19	9	7/4/2008	75	0	1	74
		SKWETL06	Kwethluk River	20	10	8/8/2006	68	14	3	51
		SKWETR06 ^B		21		8/8/2006	57	7	0	0
		SKWETR07		22	10	8/5/2007	50	1	0	49
		SKWET07 ^C		23		2007	142	0	1	0
				Kuskokwim River Total			1,422	35	11	823
Kanektok		SKAGF09	Kagati Lake tributary	24	11	8/9/2009	95	1	0	94
		SPEGF09	Pegati Lake tributary	25	11	8/8/2009	95	0	0	95
		SKAGB09	Kagati Lake beach	26	12	8/9/2009	95	0	0	95
		SPEGB09	Pegati Lake beach	27	12	8/8/2009	95	0	1	94
		SKAGO09	Kagati-Pegati Lake outlet	28	13	8/10/2009	95	0	0	95
		SKANE09	Kanektok River mainstem	29	14	8/12/2009	95	2	0	93
		SKANE02 ^C	Kanektok River	30		7/16/2002	95	0	0	0
		SKANE07 ^C	Kanektok River	31		7/10/2007	96	0	0	0
				Kanektok Total			761	3	1	566
Goodnews		SGOODSO10NF	Goodnews River - North Fork lake tributary	32	15	8/12/2010	95	0	3	92
		SGOODB10NF	Goodnews River - North Fork lake beach	33	16	8/12/2010	95	0	0	95
		SGOODO10NF	Goodnews River - North Fork lake outlet	34	16	8/12/2010	95	0	1	94
		SGOODR10NF	Goodnews River - North Fork	35	17	8/13/2010	95	0	0	95
		SGOOD02		36	17	7/23/2002	95	4	0	91
		SGOOD06		37	17	7/20/2006	48	2	1	45
		SGOODSO10MF	Goodnews River - Middle Fork lake tributary	38	18	8/8/2010	95	0	2	93
		SGOODB10MF	Goodnews River - Middle Fork lake beach	39	18	8/8/2010	95	0	0	95

667

668 Table 1: Page 3 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Kuskokwim Bay	Goodnews	SGOODO10MF	Goodnews River - Middle Fork lake outlet	40	18	8/8/2010	95	0	0	95
		SGOODR10MF	Goodnews River - Middle Fork	41	19	9/14/2010	69	0	0	69
		SGOOD91 ^C		42		8/1/1991	48	2	0	0
		SGOOD01 ^C		43		7/15/2001	96	0	0	0
		SGOOD07 ^C		44		2007	142	2	0	0
Goodnews Total							1,163	10	7	864
Kuskokwim Bay Total							3,346	48	19	2,253
Bristol Bay	Togiak	SSLUG10	Slug River	45	20	8/8/2010	108	1	0	107
		SOSVIAK10	Osviak River	46	20	8/8/2010	75	0	0	75
		STOGL00	Sunday Creek	47	21	8/21/2000	95	1	1	93
		STOGT06	Togiak Tower	48	21	7/27/2006	95	0	0	95
		SONGI06	Ongivinuk Lake	49	22	8/24/2006	95	0	0	95
		SNENE06	Nenevok Lake	50	23	8/24/2006	95	1	0	94
		SGECH00	Gechiak Lake	51	24	8/21/2000	96	1	0	95
		SKULU06	Kulukak River Lake	52	25	8/24/2006	95	0	1	94
Togiak Total							754	4	2	748
	Igushik	SUALI03	Ualik Lake	53	26	8/14/2003	99	1	0	98
		SUALI03f		54	26	8/14/2003	30	0	0	30
		SONGU07	Ongoke River - Upper	55	27	8/27/2007	95	6	1	88
		SONGL07	Ongoke River - Lower	56	27	8/28/2007	95	0	0	95
		SAMAN03	Amanka Lake	57	28	8/14/2003	100	1	2	97
		SAMAN03f		58	28	8/14/2003	57	1	0	56
		SSNAKLKB10	Snake Lake beach	59	29	8/11/2010	89	1	0	88
		SSNAKLKO10	Snake Lake outlet	60	29	8/11/2010	83	0	0	83
Igushik Total							648	10	3	635

669

670 Table 1: Page 4 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Bristol Bay	Wood	SLKUL07	Lake Kulik beaches	61	30	9/10/2007	95	4	0	91
		SGRANT07	Grant River	62	31	8/22/2007	95	9	3	83
		SKULIK01	Lake Kulik	63	32	8/1/2001	96	2	0	94
		SMIKCH09	Mikchalk Lake	64	33	9/10/2009	95	1	0	94
		SSILVH07	Silver Horn beaches	65	34	9/10/2007	95	0	0	95
		SHARDL07	Hardluck Bay beaches	66	35	9/10/2007	95	0	0	95
		SMOOSCK09	Moose Creek	67	36	8/17/2009	95	1	1	93
		SAGULU01	Agulukpak River	68	37	8/21/2001	94	2	0	92
		SANVI06	Anvil Bay Beach	69	38	8/20/2006	95	0	1	94
		SSIXCK08	Sixth Creek	70	39	2008	95	1	0	94
		SN4BE06	N4 Beach	71	40	8/11/2006	95	0	1	94
		SABEA04	A Beach - Little Togiak Lake	72	41	8/8/2004	65	0	0	65
		SABEA05		73	41	8/10/2005	30	2	0	28
		SLTOG08	Little Togiak River	74	42	2008	95	13	0	82
		SPICK01	Pick Creek	75	43	8/3/2001	95	1	2	92
		SPICK08		76	43	7/22/2008	93	1	3	89
		SLYNXLK09	Lynx Lake	77	44	9/9/2009	95	2	1	92
		SLYNX06	Lynx Beach	78	45	8/11/2006	95	0	0	95
		SLYNXCKT09	Lynx Creek - Cold Tributary	79	46	8/12/2009	81	2	0	79
		SLYNX01	Lynx Creek	80	47	8/22/2001	95	1	0	94
SLYNXCK09		81	47	8/21/2009	109	1	1	107		
SAGULO01	Agulowok River	82	48	8/22/2001	95	0	0	95		
SICEL07	Ice Creek	83	49	8/9/2007	95	6	0	89		
SHAPP01	Happy Creek	84	50	7/30/2001	95	0	0	95		
SHANS04	Hansen Creek	85	51	8/4/2004	95	0	0	95		
SBEAR01	Bear Creek	86	52	8/2/2001	96	2	0	94		
SEAGL07	Eagle Creek	87	53	8/12/2007	93	1	0	92		

671 Table 1: Page 5 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Bristol Bay	Wood	SYAKOB06	Yako Beach	88	54	8/19/2006	95	0	0	95
		SMISS98	Mission Creek	89	55	1998	94	2	1	91
		SWOOD09	Wood River	90	56	9/5/2009	95	3	0	92
Wood River Total							2,751	57	14	2,680
Nushagak		SFISHT10	Fish Trap Lake	91	57	9/4/2010	80	1	0	79
		SMULC01B	Mulchatna River	92	58	8/27/2001	65	0	0	65
		SMULC01A		93	58	8/27/2001	95	8	0	87
		SKOKT00	Koktuli River	94	59	8/13/2000	96	3	0	93
		SSTUY00	Stuyahok River	95	60	8/14/2000	96	2	0	94
		SUPNK01	Klutapuk Creek	96	61	8/18/2001	95	0	0	95
		SKING01	King Salmon River	97	61	8/18/2001	48	0	0	48
		SCHAU01	Chauekuktuli Lake beach	98	62	8/22/2001	96	0	0	96
		SALLE00	Allen River beach	99	63	8/17/2000	96	4	1	91
		SALLE01	Allen River	100	64	8/22/2001	95	1	0	94
		SNUYL00	Nuyakuk Lake	101	65	8/16/2000	95	4	0	91
		SNUYA01	Nuyakuk Lake - south beach	102	65	8/23/2001	94	0	0	94
		STIKC01	Tikchik River	103	66	8/18/2001	95	2	0	93
		STIKC00	Tikchik Lake	104	67	8/18/2000	95	1	0	94
Nushagak Total							1,241	26	1	1,214
Kvichak		STLGF99	Tlikakila River - Glacier Fork	105	68	10/6/1999	47	0	0	47
		SUTLIK01	Upper Tlikakila River	106	68	9/24/2001	96	0	0	96
		LLCL99	Little Lake Clark	107	69	10/9/1999	95	0	0	95
		SKIJ01	Kijik River	108	70	9/19/2001	96	9	0	87
		SLKIJ01	Lower Kijik River	109	71	9/18/2001	96	1	0	95
		SCHLB99	Chulitna Lodge beach	110	71	10/5/1999	96	1	0	95
		SCHLP99	Chulitna Lodge ponds	111	72	10/1/1999	48	0	1	47
		SSUCK07	Sucker Bay Lake	112	73	9/14/2007	95	0	0	95

672 Table 1: Page 6 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Bristol Bay	Kvichak	STAZI01	Tazimina River	113	74	8/29/2001	95	0	0	95
		SNHAL02	Newhalen River	114	74	9/3/2002	96	0	0	96
		STOMK00	Tomkok Creek	115	75	8/24/2000	95	0	0	95
		STOMK02		116	75	8/28/2002	48	8	0	40
		SKNUT00	Knutson Bay	117	76	8/27/2000	96	13	0	83
		SKNUT99L		118	76	10/16/1999	95	0	0	95
		SPEDR99	Pedro Ponds	119	77	1999	47	0	0	47
		SBEAR99L ^B	Pedro Ponds - Bear Pond late	120		10/17/1999	47	0	0	0
		SGRAS99L	Pedro Ponds - Grass Pond late	121	77	10/15/1999	44	0	0	44
		SCHIN00	Chinkelyes Creek	122	78	8/28/2000	96	1	0	95
		SILIA04B	Iliamna River	123	78	8/21/2004	95	0	3	92
		SILIA99L	Iliamna River - late	124	79	10/17/1999	96	10	0	86
		SFING00	Finger Beach 1	125	80	8/24/2000	84	1	0	83
		SSOUT99	Southeast Creek beach	126	80	8/26/1999	95	0	0	95
		SPORC99	Porcupine Island - Painted Rock	127	81	1999	48	0	0	48
		SFUEL00	Fuel Dump Island	128	81	8/28/2000	96	4	0	92
		SWOOD01	Woody Island - West Beach	129	81	8/19/2001	96	1	0	95
		STRIA00	Triangle Island 2	130	81	8/16/2000	96	1	0	95
		STOMM00	Tommy River	131	82	8/24/2000	96	4	0	92
		STOMM02		132	82	8/19/2002	48	0	1	47
SCOPP00	Copper River	133	83	8/28/2000	96	0	0	96		
SCOPP99		134	83	8/23/1999	47	0	0	47		
SNICK00	Nick N Creek	135	84	8/25/2000	96	4	0	92		
SSECK00	Southeast Creek	136	85	8/26/2000	96	2	0	94		
SDREA01	Dream Creek	137	85	8/22/2001	95	1	0	94		

673

676 Table 1: Page 8 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Bristol Bay	Naknek	SHEAD01	Headwater Creek	162	98	7/22/2001	93	19	0	74
		SBRLK00	Brooks Lake	163	98	8/22/2000	96	0	0	96
		SCHARL06 ^B	Charlene Creek	164		9/11/2006	47	0	0	0
		SLQTIP06	Lower Q-Tip Lake	165	99	9/12/2006	86	0	0	86
		SIDAV00	Idavain Creek	166	100	8/23/2000	95	2	0	93
		SIDAV06		167	100	8/29/2006	48	0	0	48
		SDUMP306	Dumpling Creek beach	168	101	9/17/2006	83	0	0	83
Naknek Total							1,145	35	4	938
Egegik		SCABI00	Cabin Creek	169	102	8/15/2000	96	1	0	95
		SRUTH00 ^D	Ruth Lake outlet	170		8/12/2000	96	7	0	0
		SSALCR06	Salmon Creek	171	102	8/16/2006	95	7	0	88
		SBURL06	Burls Creek	172	102	8/16/2006	95	1	2	92
		SCLEO01	Cleo Creek	173	102	8/16/2001	48	0	0	48
		SFEAT01	Featherly Creek	174	102	8/16/2001	48	0	0	48
		SBECH00	Becharof Creek	175	102	8/11/2000	94	3	0	91
		SKEJU00	Upper Kejulik River	176	103	8/8/2000	96	0	2	94
		SKEJU01	Kejulik River	177	103	8/17/2001	96	0	0	96
		SBECH08NT	Becharof Lake north	178	104	8/11/2008	95	1	1	93
SBECH08SB	Becharof Lake south	179	105	8/11/2008	95	1	1	93		
Egegik Total							954	21	6	838
Ugashik		SUGAS01	Ugashik Creek	180	106	7/21/2001	96	7	0	89
		SCROCK05	Crooked Creek	181	107	8/24/2005	95	0	0	95
		SDEER01	Deer Creek	182	107	7/20/2001	96	0	0	96
		SUGAS00	Ugashik Narrows	183	108	8/24/2000	96	0	0	96
		SBLACKU05	Black Creek	184	109	8/24/2005	95	1	0	94
		SECRE05	E Creek	185	109	8/8/2005	95	0	0	95

678 Table 1: Page 9 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals				
Regional	Sub-regional						Initial	Missing	Duplicate	Final	
Bristol Bay	Ugashik	SOUTL00	Outlet Stream	186	110	8/26/2000	95	4	0	91	
		SFIGU05	Figure 8 Creek	187	111	8/22/2005	95	4	1	90	
		SOLDH05	Old Ham Creek	188	112	8/22/2005	95	1	0	94	
Ugashik Total							858	17	1	840	
Bristol Bay Total							12,654	273	40	11,848	
North Peninsula	Cinder	SWIGGC05	Wiggly Creek	189	113	7/29/2005	90	0	10	80	
		SMAINC05	Mainstem Cinder River	190	113	7/29/2005	95	1	0	94	
		SLAVA04	Lava Creek	191	114	7/23/2004	95	0	3	92	
		SMUDA05	Mud Creek A	192	115	7/30/2005	95	0	0	95	
	Cinder Total							375	1	13	361
	Meshik	SMESLK05	Meshik Lake beach	193	116	7/30/2005	95	0	0	95	
		SMESLKO05	Meshik Lake outlet	194	116	7/30/2005	95	0	0	95	
		SMESHL05	L Creek - Meshik River	195	117	7/30/2005	95	1	0	94	
		SMESH202	Blue Violet Creek - Meshik River	196	117	7/29/2002	93	0	1	92	
		SMESH102	Landlock Creek - Meshik River	197	118	7/29/2002	96	0	0	96	
SREDBC05		Red Bluff Creek	198	119	7/30/2005	95	0	0	95		
Meshik Total							569	1	1	567	
Ilnik	SNPEN01	Willie Creek	199	120	8/27/2001	81	1	0	80		
	SOCEA01	Ocean River	200	120	2001	96	1	0	95		
	SILNIK07	Ilnik River	201	121	7/7/2007	190	1	1	188		
	SWILD05	Wildman Lake	202	122	7/30/2005	95	0	1	94		
Ilnik Total							462	3	2	457	
Sandy	SSAND00	Sandy Lake	203	123	6/30/2000	95	0	0	95		
	SSANDR07	Sandy River	204	123	7/8/2007	190	0	0	190		
Sandy Total							285	0	0	285	

679

680 Table 1: Page 10 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
North Peninsula	Bear	SBEAR00E	Bear River - early	205	124	6/30/2000	95	0	0	95
		SBEARR07		206	125	7/7/2007	95	0	0	95
		SCUB04	Cub Creek	207	126	8/15/2004	95	0	0	95
		SREDC04	Red Creek	208	126	8/15/2004	95	0	0	95
		SBEARS05	Bear Lake beach	209	127	8/29/2005	95	1	0	94
		SBEARO05	Bear Lake outlet	210	128	8/29/2005	95	0	0	95
		SBEAR00L	Bear River	211	129	8/18/2000	96	2	0	94
Bear Total							666	3	0	663
Nelson River		SHOOD01	Hoodoo Lake	212	130	7/31/2001	95	1	0	94
		SHOOD05	Hoodoo Lake beach	213	130	7/31/2005	95	1	0	94
		SNELSR07	Nelson River	214	130	7/1/2007	47	0	0	47
		SDAVI05	Davids River	215	131	7/31/2005	95	0	0	95
		SHOOD00 ^E	Nelson River	216		7/5/2000	96	0	0	0
Nelson River Total							428	2	0	330
NW District-Black Hills		SNCREK07	North Creek	217	132	7/25/2007	95	1	4	90
		SMOF09	Moffett Creek	218	133	8/18/2009	95	0	3	92
		SMOFF02	Paul Hansen tributary	219	134	7/30/2002	95	2	0	93
		SOUTE04	Outer Marker Lake	220	135	9/9/2004	95	2	0	93
		SBLUE04	Blue Bill Lake	221	135	9/7/2004	95	0	1	94
		SSWANL08	Swansons Lagoon	222	136	8/25/2008	95	1	0	94
		SPETELA05	Peterson Lagoon	223	137	8/2/2005	95	1	0	94
		SWHAL02	Whaleback Mountain Creek	224	138	7/30/2002	96	0	0	96
		SMCLE04	McLees Lake	225	139	6/4/2004	143	0	1	142
SSUMM99	Summer Bay Lake	226	140	8/25/1999	96	0	0	96		
NW District-Black Hills Total							1,000	7	9	984
North Peninsula Total							3,785	17	25	3,647

681

682 Table 1: Page 11 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals				
Regional	Sub-regional						Initial	Missing	Duplicate	Final	
South Peninsula	South Peninsula	SSANA08	Sanak Island	227	141	8/24/2008	86	0	0	86	
		SHANLK05	Hansen Lake	228	142	8/2/2005	95	0	0	95	
		SMIDL04	Middle Lagoon - Morzhovoi Bay	229	143	7/28/2004	95	2	0	93	
		STHIN05	Thin Point Lagoon	230	144	8/1/2005	95	1	0	94	
		SMORT04	Mortensen's Lagoon	231	145	8/2/2004	95	0	0	95	
		SLONGJ05	Long John Lagoon	232	146	8/1/2005	95	0	0	95	
		SCANBR08	Canoe Bay River	233	147	8/26/2008	95	1	0	94	
		SARCH05	Archeredin Lake	234	148	8/3/2005	95	1	0	94	
		SORZI00	Orzinski	235	149	7/1/2000	94	1	0	93	
South Peninsula Total							845	6	0	839	
Chignik	Black Lake	SBROAD97	Broad Creek	236	150	9/1/1997	96	1	1	94	
		SBSPR97	Big Spring	237	150	1997	95	2	0	93	
		SBOUL97	Boulevard Creek	238	150	9/1/1997	95	0	0	95	
		SFAN97	Fan Creek	239	150	1997	95	0	0	95	
		SALEC97	Alec River	240	150	9/1/1997	96	0	0	96	
	Black Lake Total							477	3	1	473
	Chignik Lake		SCHIA08	Chiaktuak Creek	241	151	8/29/2008	95	2	0	93
			SCHIA97E		242	151	1997	95	0	1	94
			SCHIA97M		243	151	9/18/1997	94	1	0	93
			SWESTF97 ^E	West Fork Chignik River	244		1997	95	0	0	0
SWESTF08				245	152	8/28/2008	95	1	0	94	
		SCUCU08	Cucumber Creek	246	153	8/29/2008	95	0	1	94	
		SHAT08E	Hatchery Beach	247	153	8/29/2008	95	1	2	92	
		SHAT96		248	154	10/18/1996	95	0	0	95	
		SHAT97E		249	153	9/15/1997	94	0	0	94	

683

684 Table 1: Page 12 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
Chignik	Chignik Lake	SCLARK08	Clark River	250	155	8/28/2008	94	3	0	91
		SCLARK96		251	154	10/19/1996	95	0	0	95
		SCLRK97E		252	155	9/16/1997	96	1	0	95
	SCHIG08	Chignik River	253	156	8/30/2008	95	1	0	94	
	SCHIG98		254	156	8/22/1998	95	0	0	95	
Chignik Lake Total							1,328	10	4	1,219
Chignik Total							1,805	13	5	1,692
East of WASSIP	East of WASSIP	SSURPL08	Surprise Lake	255	157	8/22/2008	95	0	0	95
		SOCEAB06	Ocean Beach	256	158	8/29/2006	95	0	0	95
		SHORS05	Horse Marine Lake	257	159	9/2/2005	95	0	0	95
		SPINNM08	Frazer Lake - Pinnell Creek	258	160	8/21/2008	78	0	0	78
		SSTUM08	Frazer Lake - Stumble Creek	259	160	8/21/2008	95	1	0	94
		SCOUR08	Frazer Lake - Courts Shoreline	260	161	8/21/2008	95	7	0	88
		SMIDWM08	Frazer Lake - Midway Creek	261	160	8/21/2008	93	1	0	92
		SMIDWS08	Frazer Lake - Midway beach	262	161	8/21/2008	95	4	0	91
		SLINDM08	Frazer Lake - Linda Creek	263	160	8/22/2008	95	5	0	90
		SHOLFS08	Frazer Lake - Hollow Fox beach	264	161	8/22/2008	95	1	0	94
		SVALA08	Frazer Lake - Valarian Creek	265	162	8/21/2008	95	0	0	95
		SOUTS08	Frazer Lake - Outlet beach	266	163	8/20/2008	95	10	0	85
		SDOGSC08	Frazer Lake - Dog Salmon Creek	267	164	8/22/2008	95	3	0	92
		SAKAL05L	Akalura Lagoon	268	165	9/2/2005	95	0	0	95
		SUPS00E	Upper Station	269	166	6/15/2000	95	0	0	95
		SUPUP93	Upper Station - Upper	270	167	9/1/1993	95	0	0	95
		SLUPS93	Upper Station - Lower	271	168	1993	95	1	0	94
SAYAK00	Ayakulik River	272	169	7/26/2000	96	1	2	93		
SAYAK08L		273	169	8/14/2008	95	3	1	91		

685

686 Table 1: Page 13 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SOMALL99	Karluk Lake - O'Malley River	274	170	9/30/1999	95	1	2	92
		SUTHU00E	Karluk Lake - Upper Thumb Lake	275	171	7/24/2000	95	0	0	95
		SLTHUM99	Karluk Lake - Lower Thumb River	276	170	9/30/1999	95	19	0	76
		SLRIV97	Little River Lake	277	172	7/15/1997	96	1	0	95
		SUGAN97	Uganik Lake	278	173	7/15/1997	95	0	0	95
		SBUSK05	Buskin Lake	279	174	6/26/2005	95	1	0	94
		SBUSKL10		280	174	6/13/2010	95	0	1	94
		SLKLOU05	Lake Louise - Buskin River	281	175	8/3/2005	95	0	0	95
		SLKLOU10		282	175	7/19/2010	95	0	2	93
		SPASA05	Pasagshak Lake	283	176	7/15/2005	95	0	0	95
		SLMIA05	Lake Miam	284	177	9/2/2005	95	0	1	94
		SSALT94	Saltery Lake	285	178	1994	95	2	0	93
		SSALT99		286	178	8/26/1999	95	1	0	94
		SAFOG93	Afognak Lake	287	179	8/15/1993	79	0	1	78
		SMALI93	Malina	288	180	8/15/1993	80	1	0	79
		STHOR06	Thorsheim Lake	289	181	8/23/2006	83	0	0	83
		SPORT98	Portage Lake	290	182	1998	96	0	0	96
		SLKIT93	Little Kitoi	291	183	9/10/1993	95	0	0	95
		SKAFL08	Kaflia Lake	292	184	8/27/2008	95	1	0	94
		SWACK09	Wackton Creek - Lake Fork Crescent River	293	185	8/13/2009	95	2	0	93
		SPYRAM09	Pyramid Creek - Crescent Lake	294	186	8/13/2009	95	0	0	95
		SCRES941 ^E	Crescent Lake	295		1994	48	0	0	0
		SCREE942 ^E		296		1994	47	0	0	0
		SCRESL09 ^F	Crescent Lake outlet	297		8/12/2009	95	0	0	0
SLJACK06	Little Jack Creek	298	187	9/6/2006	95	1	0	94		
SPACK92	Packers Lake	299	188	7/1/1992	95	0	0	95		

687

688 Table 1: Page 14 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SSFBIG07	South Fork Big River	300	189	8/14/2007	123	0	0	123
		SSFBIGF09	South Fork Big River Falls	301	189	7/7/2009	48	0	0	48
		SWOLV93	Wolverine Creek - Big River	302	190	7/5/1993	95	1	0	94
		SBLACSC07	Black Sand Creek	303	191	8/13/2007	95	0	0	95
		SFARR07	Farro Lake Creek	304	192	8/13/2007	95	0	0	95
		SMCCA93	McArthur River	305	193	1993	95	0	1	94
		SCHIL92	Chilligan River	306	194	1992	48	0	0	48
		SCHIL94		307	194	1994	48	0	0	48
		SCHAK08	Chakachatna Slough	308	195	8/27/2008	95	0	0	95
		SCOAS09	Coal Creek Spring	309	196	8/21/2009	48	0	0	48
		SCOAW09	West Fork Coal Creek	310	196	8/21/2009	47	0	0	47
		SMOOSE07	Moose Creek	311	197	8/27/2007	95	0	0	95
		SPUNT06	Puntilla Lake	312	198	9/6/2006	95	0	0	95
		SREDSA06	Red Salmon Lake	313	199	9/7/2006	95	2	0	93
		STRIM107 ^{B,E}	Trimble River	314		9/17/2007	61	0	0	0
		STRIM109 ^{B,E}		315		9/1/2009	18	1	0	0
		STRIM207		316	200	9/17/2007	47	0	0	47
		STRIM209		317	200	9/1/2009	48	0	0	48
		SHAYT08	Hayes River tributary	318	201	9/2/2008	48	0	1	47
		SHAYT09		319	201	8/28/2009	47	0	0	47
		SSKWEN07	Skwentna River	320	202	9/20/2007	108	0	0	108
		SCANYC07	Canyon Creek	321	202	9/20/2007	65	0	0	65
		SJUDD06	Judd Lake	322	203	7/26/2006	94	2	0	92
		SJUDD09		323	203	2009	95	2	0	93
SJUDD93		324	203	8/23/1993	96	0	0	96		
STRIN09	Trinity Lake - inlet	325	204	8/22/2009	95	0	0	95		

689

690 Table 1: Page 15 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	STRIN92	Trinity Lake	326	204	8/1/1992	48	0	1	47
		SSHEL06	Shell Lake	327	205	7/24/2006	95	1	0	94
		SSHEL09		328	205	2009	95	2	0	93
		SSHEL93		329	205	9/3/1993	48	0	0	48
		SWHISK06	Whiskey Lake Outlet	330	206	9/2/2006	58	0	0	58
		SWHISK09		331	206	9/1/2009	47	0	0	47
		SHEWI06	Hewitt Lake	332	206	8/2/2006	65	4	0	61
		SHEWI92 ^E		333		8/1/1992	47	0	0	0
		SJOHNCK09	Johnson Creek	334	207	8/28/2009	95	0	2	93
		SKICH107	Kichatna River	335	207	8/27/2007	95	0	0	95
		SYENW92	West Fork Yentna River	336	208	9/1/1992	96	8	0	88
		SYENW93		337	208	9/10/1993	100	3	1	96
		SCHEL06	Chelatna Lake	338	209	7/27/2006	95	0	0	95
		SCHEL09		339	209	8/7/2009	95	0	0	95
		SCHEL93		340	209	8/28/1993	95	0	0	95
		SBYER07	Byers Lake	341	210	8/13/2007	95	3	0	92
		SBYER93		342	210	1993	48	1	0	47
		SSPINK08	Spink Creek	343	211	8/30/2008	95	2	0	93
		SSWALK06	Swan Lake	344	212	9/2/2006	95	0	0	95
		SSWALK09		345	212	9/8/2009	48	0	1	47
		SSWLK07		346	212	8/15/2007	47	4	0	43
		SSUS9511	Susitna River sloughs	347	213	1995	50	2	0	48
		SSUS9611		348	213	9/5/1996	6	0	0	6
		SSUS97		349	213	9/5/1997	94	0	0	94
		SSTEP07	Stephan Lake	350	214	7/28/2007	95	0	0	95
		SSTEP93		351	214	9/2/1993	48	0	0	48

692 Table 1: Page 16 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SLARS06	Larson Lake	352	215	7/23/2006	95	1	0	94
		SLARS93		353	216	9/1/1993	95	0	0	95
		SMAMA97	Mama and Papa Bear Lakes	354	217	9/3/1997	50	0	0	50
		SPAPA07		355	217	8/28/2007	54	0	1	53
		STALK97	Talkeetna River sloughs	356	217	9/4/1997	79	11	0	68
		SBIRC07	Birch Creek	357	218	8/28/2007	95	0	1	94
		SBIRC93 ^E		358		1993	67	2	0	0
		SSHEEP08	Sheep River	359	219	8/30/2008	95	0	0	95
		SNANC10	Nancy Lake	360	220	9/3/2010	95	0	0	95
		SNANC93		361	220	8/27/1993	95	0	0	95
		SLMEAD09	Little Meadow Creek	362	221	8/8/2009	142	0	0	142
		SFISH93	Fish Creek	363	222	1993	95	0	0	95
		SFISH94		364	221	8/15/1994	94	0	0	94
		SBIGL92	Big Lake	365	222	8/1/1992	95	0	0	95
		SCOTT93	Cottonwood Creek	366	223	1993	95	0	0	95
		SWASI98	Wasilla Creek	367	223	1998	71	5	0	66
		SESKA06	Eska Creek	368	224	9/5/2006	95	0	0	95
		SJIM97	Jim Creek	369	225	9/2/1997	95	1	0	94
		SBODE06	Bodenburg Creek	370	226	8/30/2006	95	1	0	94
		SSIXM08	Sixmile Creek	371	227	7/30/2008	95	0	1	94
		SCARMLK10	Carmen Lake	372	228	8/23/2010	95	0	0	95
		SWILLIW06	Williwaw Creek	373	229	9/7/2006	39	0	0	39
		SWILLIW07		374	229	8/23/2007	69	0	0	69
		SCHICK10	Chickaloon River	375	230	7/13/2010	95	0	0	95
SSWAN97	Swanson River	376	231	8/21/1997	95	0	0	95		
SBISH93	Bishop Creek	377	232	1993	95	0	0	95		

694 Table 1: Page 17 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SDANI93	Daniels Lake	378	233	1993	95	2	0	93
		SRAIL97	Railroad Creek	379	234	8/13/1997	48	0	0	48
		SJOHN97	Johnson Creek	380	234	8/12/1997	88	1	0	87
		SMOOK93	Moose Creek	381	235	7/27/1993	47	0	0	47
		SMOOK94		382	235	1994	95	0	0	95
		SPTAR92	Ptarmigan Creek	383	236	8/1/1992	47	1	0	46
		SPTAR93		384	236	1993	95	1	0	94
		STERN92	Tern Lake	385	237	9/1/1992	48	0	1	47
		STERN93		386	237	1993	48	0	0	48
		SQUAR93	Quartz Creek	387	238	8/6/1993	94	1	0	93
		SURGOAT09E	Upper Russian Lake - Goat Creek	388	239	7/20/2009	95	2	0	93
		SURGOATM09		389	240	9/3/2009	95	1	0	94
		SURUS97		390	239	8/19/1997	95	0	0	95
		SRBEAR09	Upper Russian Lake - Bear Creek	391	241	9/3/2009	95	1	1	93
		SURSHOAL09	Upper Russian Lake beach	392	242	9/4/2009	95	0	0	95
		SUPRUS99	Upper Russian Lake South beach	393	242	9/16/1999	95	0	1	94
		SURUSA99	Upper Russian Lake outlet	394	243	9/17/1999	95	1	0	94
		SUROUT09		395	243	9/2/2009	95	0	0	95
		SRUSA92E	Russian River above falls	396	239	7/1/1992	96	0	0	96
		SRUSA93L ^F		397		8/2/1993	95	0	0	0
SRUSB93	Russian River below falls	398	244	8/2/1993	95	0	1	94		
SSKK194L	Kenai River	399	244	8/22/1994	47	0	0	47		
SSKK294L		400	244	8/22/1994	48	0	0	48		
SSKK494L		401	244	8/22/1994	48	0	0	48		
SSKK394E		402	244	1994	96	1	0	95		
SSKK394L		403	244	8/22/1994	47	0	0	47		

696 Table 1: Page 18 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SSKKE93E	Kenai River	404	245	8/18/1993	48	1	0	47
		SSKKE93L		405	245	9/11/1993	47	1	0	46
		SSKK594L		406	245	9/9/1994	95	0	0	95
		SHIDDN08	Hidden Lake - North shore	407	246	9/23/2008	95	2	0	93
		SHIDD93		408	246	7/29/1993	95	0	0	95
		SSKIL95	Skilak Lake	409	247	1995	48	0	0	48
		SSKIL92	Skilak Lake - outlet	410	247	8/1/1992	96	0	0	96
		SSKIL94E		411	247	1994	45	2	0	43
		SSKIL94L		412	247	1994	50	3	0	47
		STUST941	Tustumena Lake	413	248	1994	48	0	0	48
		STUST942		414	248	1994	48	0	0	48
		SSEEP94	Seepage Creek	415	248	1994	95	0	0	95
		SGLAC94	Glacier Flats Creek	416	249	1994	95	0	0	95
		SMOOT92	Moose Creek	417	249	8/1/1992	96	2	0	94
		SBEAR92	Bear Creek	418	250	8/1/1992	95	0	0	95
		SNIKO92	Nikolai Creek	419	250	7/1/1992	95	0	0	95
		SENG92E	English Bay	420	251	6/1/1992	95	9	0	86
		SENG92L		421	252	10/1/1992	95	1	0	94
		SBEARLK10	Bear Lake	422	253	8/9/2010	190	0	1	189
		SBAIN10	Bainbridge Lake	423	254	8/6/2010	95	0	0	95
SESHA91	Eshamy Lake	424	255	10/1/1991	96	6	0	90		
SESHAR08	Eshamy Creek	425	255	8/3/2008	95	0	0	95		
SMAIN91 ^G	Main Bay	426		7/13/1991	96	0	0	0		
SCOG92HL	Coghill Lake tributary	427	256	8/27/1992	96	3	0	93		
SCOG92ES		428	256	8/27/1992	96	1	0	95		

697

698 Table 1: Page 19 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SCOGH91	Coghill Lake	429	257	9/1/1991	96	1	0	95
		SCOGH10		430	257	7/7/2010	95	0	0	95
		SMINE09	Miners Lake	431	258	7/9/2009	95	0	0	95
		SMINE91		432	258	8/9/1991	96	0	0	96
		SEYAK10	Eyak Lake - Hatchery Creek	433	259	7/24/2010	95	0	0	95
		SEYAM07	Eyak Lake - Middle Arm	434	260	8/2/2007	95	0	0	95
		SEYASB07	Eyak Lake - South beaches	435	261	8/22/2007	95	7	1	87
		SMCKI07	McKinley Lake	436	262	8/20/2007	95	0	0	95
		SMCKI08		437	263	7/29/2008	95	0	0	95
		SMCKI91	McKinley Lake - Salmon Creek	438	264	7/1/1991	95	0	0	95
		SMCKSC07		439	264	7/25/2007	95	2	0	93
		STANAS09	Tanada Lake beach	440	265	9/9/2009	95	2	0	93
		STANAO09	Tanada Lake outlet	441	266	9/9/2009	95	0	0	95
		STANA05	Tanada Creek	442	267	8/21/2005	95	0	1	94
		SMENT08	Mentasta Lake	443	268	7/15/2008	95	0	0	95
		SFISHC08	Fish Creek - East Fork Gulkana River	444	269	8/1/2008	95	0	0	95
		SGULK08EF	East Fork Gulkana River	445	270	8/1/2008	75	0	0	75
		SSWEDE08	Swede Lake	446	271	8/13/2008	95	0	0	95
		SPAXSO09	Paxson Lake outlet	447	272	8/21/2009	77	0	2	75
		SMEND08	Mendeltna Creek	448	273	8/22/2008	95	0	1	94
		SMEND09		449	273	8/12/2009	94	0	0	94
		SBANA08	Banana Lake - Klutina drainage	450	274	8/18/2008	82	2	0	80
		SBEARH08	Bear Hole - Klutina tributary	451	275	8/14/2008	95	1	0	94
		SKLUTI08	Klutina Lake inlet	452	276	8/21/2008	44	0	0	44
		SKLUTI09		453	276	8/13/2009	51	0	0	51

699

700 Table 1: Page 20 of 20.

Reporting Group		ADF&G code	Location	Collection	Population	Date	No. of individuals			
Regional	Sub-regional						Initial	Missing	Duplicate	Final
East of WASSIP	East of WASSIP	SSANN05	St. Anne Creek	454	277	7/15/2005	95	0	1	94
		SSTACR08		455	277	7/22/2008	95	0	3	92
		SMAHL08	Mahlo River	456	278	7/22/2008	95	0	1	94
		SKLUT08	Klutina River mainstem	457	279	8/21/2008	95	0	0	95
		STONSL09	Tonsina Lake	458	280	8/8/2009	95	0	1	94
		SLONGLK05	Long Lake	459	281	9/7/2005	95	0	0	95
		STEBA08	Tebay Outlet	460	282	8/18/2008	94	1	0	93
		SSALMC08	Salmon Creek - Bremner drainage	461	283	8/17/2008	95	2	0	93
		SSTEAM08	Steamboat Lake	462	284	8/17/2008	95	0	0	95
		SCLEAR07	Clear Creek at 40 Mile	463	285	8/24/2007	95	8	0	87
		STOKUN08	Tokun Lake	464	286	6/19/2008	95	0	0	95
		STOKUN09		465	286	6/25/2009	94	0	0	94
		SMART07	Martin Lake	466	287	7/26/2007	95	2	0	93
		SMART08		467	287	7/21/2008	95	1	0	94
		SERB91 ^E		468		7/28/1991	96	0	0	0
		SMARTR08	Martin River Slough	469	288	7/11/2008	95	0	0	95
		SKUSH07	Kushtaka Lake	470	289	8/9/2007	95	1	0	94
		SKUSH08		471	289	8/8/2008	95	0	0	95
		SBERI91	Bering Lake	472	290	7/10/1991	95	0	0	95
		East of WASSIP Total							18,456	199
Cape Suckling to Cape Prince of Wales Total							41,406	583	127	38,193

701 ^A These collections belong to a reporting group that has too few fish to be independent, lacks genetic distinction from riverine sockeye salmon from the
 702 Kuskokwim River group and were recommended for exclusion by the WASSIP Advisory Panel.

703 ^B These collections have a sample size less than our desired minimum cut-off of 75 individuals and do not pool with other collections.

704 ^C These collections are samples taken from weirs that are not needed in the baseline. They have been used as escapement tests of the baseline.

705 ^D These collections failed to conform to Hardy-Weinberg expectations.

706 ^E These collections appear as outliers based upon measures of genetic distance, lack reliable collection metadata and were removed from the baseline.

707 ^F These collections are likely a mixture of fish from multiple populations.

708 ^G This collection is a sample from a hatchery and is not needed to represent its broodstock population.

709 Table 2. Escapement test number, reporting group of origin, river of origin, type of collection, year
 710 collected and final sample size for samples of escapement that served as tests of the WASSIP sockeye
 711 salmon baseline.

Number	Reporting Group	River	Type	Date	N
1	Kuskokwim River	Kuskokwim	Bethel Test Fishery	6/8-7/29/2010	373
2 ^A	Kuskokwim River	Kogrukluuk	Weir	7/6/2001	91
3 ^A	Kuskokwim River	Kogrukluuk	Weir	7/24-8/15/2007	95
4 ^A	Kuskokwim River	Kwethluk	Weir	2007	141
5 ^A	Kanektok	Kanektok	Weir	7/16/2002	95
6 ^A	Kanektok	Kanektok	Weir	7/10-19/2007	96
7 ^A	Goodnews	Goodnews	Weir	7/15/2001	96
8 ^A	Goodnews	Goodnews	Weir	6/1-7/31/2007	140
9	Togiak	Togiak	Subsistence	7/11-8/1/2008	473
10	Igushik	Igushik	Tower	6/26-7/16/2005	190
11 ^A	Igushik	Igushik	Tower	6/26-7/19/2007	186
12 ^A	Wood	Wood	Tower	7/5-7/2003	174
13	Wood	Wood	Tower	6/19-7/13/2004	191
14	Wood	Wood	Tower	6/24-7/13/2007	190
15 ^A	Nushagak	Nushagak	Sonar	6/19-7/18/2006	185
16	Nushagak	Nushagak	Radio telemetry	2005	190
17 ^A	Nushagak	Nushagak	Radio telemetry	2006	164
18 ^A	Nushagak	Nuyakuk	Tower	6/27-7/16/2004	189
19	Kvichak	Kvichak	Tower	7/6/2005	190
20	Kvichak	Kvichak	Tower	7/7-9/2006	190
21	Alagnak	Alagnak	Tower	7/13/2004	192
22 ^A	Alagnak	Alagnak	Tower	7/1-12/2008	188
23	Naknek	Naknek	Tower	6/28/2002	190
24	Naknek	Naknek	Tower	6/29-7/9/2008	190
25	Egegik	Egegik	Tower	6/19-7/9/2004	192
26	Egegik	Egegik	Tower	6/23-7/16/2007	190
27	Ugashik	Ugashik	Tower	7/7-21/2004	192
28	Ugashik	Ugashik	Tower	7/4-6/2008	190
29	Black Lake	Chignik	Weir	6/14-21/2010	379
30	Chignik Lake	Chignik	Weir	7/23-30/2010	379

712 ^A These tests contain fewer than 190 individuals and results should be interpreted with caution.

713 Table 3. Source, expected and observed heterozygosity, and F_{ST} for the 89 unlinked single nucleotide
 714 polymorphisms (SNPs) of the 96 used to analyze the population genetic structure of sockeye salmon in
 715 the WASSIP study area. F_{ST} estimates are also provided for the two sets of linked loci combined as
 716 haplotypes. These summary statistics are based on the 290 populations included in the WASSIP sockeye
 717 salmon baseline.

Assay	Source ¹	H _E	H _O	F _{ST}	Assay	Source ¹	H _E	H _O	F _{ST}
<i>One_ACBP-79</i>	A	0.461	0.401	0.110	<i>One_Ots208-234</i>	C	0.159	0.141	0.109
<i>One_agt-132</i>	B	0.422	0.386	0.084	<i>One_Ots213-181</i>	A	0.261	0.230	0.105
<i>One_aldB-152</i>	C	0.392	0.352	0.101	<i>One_p53-534</i>	A	0.049	0.042	0.145
<i>One_apoe-83</i>	B	0.372	0.324	0.123	<i>One_pax7-248</i>	C	0.250	0.229	0.080
<i>One_c3-98</i>	B	0.121	0.129	0.058	<i>One_PIP</i>	D	0.473	0.432	0.081
<i>One_CD9-269</i>	B	0.378	0.342	0.091	<i>One_Prl2</i>	A	0.499	0.454	0.094
<i>One_cetn1-167</i>	B	0.500	0.437	0.113	<i>One_rab1a-76</i>	B	0.247	0.220	0.095
<i>One_CFP1</i>	D	0.235	0.218	0.065	<i>One_RAG1-103</i>	A	0.048	0.043	0.115
<i>One_cin-177</i>	C	0.492	0.461	0.061	<i>One_RAG3-93</i>	A	0.132	0.123	0.062
<i>One_CO1²</i>	A	N/A	N/A	N/A	<i>One_redd1-414</i>	C	0.497	0.421	0.138
<i>One_ctgf-301</i>	A	0.037	0.036	0.033	<i>One_RFC2-102</i>	A	0.327	0.295	0.088
<i>One_Cytb₁₇²</i>	A	N/A	N/A	N/A	<i>One_RFC2-285</i>	A	0.082	0.075	0.082
<i>One_Cytb₂₆²</i>	A	N/A	N/A	N/A	<i>One_rpo2j-261</i>	C	0.327	0.304	0.064
<i>One_E2-65</i>	A	0.362	0.332	0.091	<i>One_sast-211</i>	C	0.074	0.069	0.033
<i>One_gdh-212</i>	C	0.463	0.425	0.075	<i>One_spf30-207</i>	C	0.368	0.336	0.089
<i>One_GHII-2165</i>	A	0.245	0.185	0.241	<i>One_srp09-127</i>	C	0.029	0.028	0.039
<i>One_ghsR-66</i>	C	0.447	0.383	0.142	<i>One_ssrD-135</i>	C	0.495	0.464	0.074
<i>One_GPDH-201³</i>	A	N/A	N/A	N/A	<i>One_STC-410</i>	A	0.478	0.372	0.209
<i>One_GPDH2-187</i>	A	0.152	0.134	0.107	<i>One_STR07</i>	A	0.439	0.384	0.130
<i>One_GPH-414</i>	A	0.452	0.404	0.107	<i>One_SUMO1-6</i>	C	0.336	0.313	0.068
<i>One_HGFA-49</i>	A	0.276	0.254	0.065	<i>One_sys1-230</i>	C	0.496	0.435	0.121
<i>One_HpaI-71</i>	A	0.448	0.401	0.101	<i>One_taf12-248</i>	C	0.022	0.020	0.087
<i>One_HpaI-99</i>	A	0.177	0.140	0.203	<i>One_Tf_ex11-750³</i>	A	N/A	N/A	N/A
<i>One_hsc71-220</i>	A	0.332	0.301	0.092	<i>One_Tf_in3-182</i>	A	0.162	0.114	0.290
<i>One_Hsp47</i>	D	0.320	0.283	0.113	<i>One_tshB-92</i>	C	0.140	0.124	0.115
<i>One_IL8r-362</i>	A	0.116	0.107	0.092	<i>One_txnip-401</i>	C	0.003	0.003	0.038
<i>One_KCT1-453</i>	B	0.212	0.194	0.083	<i>One_U1003-75</i>	B	0.283	0.228	0.184
<i>One_KPNA-422</i>	A	0.370	0.333	0.096	<i>One_U1004-183</i>	B	0.498	0.350	0.302
<i>One_LEI-87</i>	A	0.487	0.429	0.113	<i>One_U1009-91</i>	B	0.303	0.261	0.138
<i>One_lpp1-44</i>	B	0.455	0.377	0.168	<i>One_U1010-81</i>	B	0.062	0.058	0.057
<i>One_metA-253</i>	C	0.072	0.044	0.389	<i>One_U1012-68</i>	B	0.257	0.225	0.112
<i>One_MHC2₁₉₀²</i>	A	N/A	N/A	N/A	<i>One_U1013-108</i>	B	0.261	0.239	0.068
<i>One_MHC2₂₅₁²</i>	A	N/A	N/A	N/A	<i>One_U1014-74</i>	B	0.279	0.258	0.073
<i>One_Mkpro-129</i>	C	0.500	0.442	0.113	<i>One_U1016-115</i>	B	0.465	0.411	0.103
<i>One_ODC1-196</i>	B	0.478	0.428	0.106	<i>One_U1024-197</i>	B	0.178	0.167	0.060

719 Table 3: Page 2 of 2.

Assay	Source ¹	H _E	H _O	F _{ST}
<i>One_U1101</i>	B	0.342	0.319	0.063
<i>One_U1103</i>	B	0.049	0.042	0.127
<i>One_U1105</i>	B	0.393	0.322	0.162
<i>One_U1201-492</i>	B	0.452	0.428	0.057
<i>One_U1202-1052</i>	B	0.455	0.411	0.075
<i>One_U1203-175</i>	B	0.456	0.409	0.098
<i>One_U1204-53</i>	B	0.355	0.326	0.071
<i>One_U1205-57</i>	B	0.073	0.061	0.144
<i>One_U1206-108</i>	B	0.323	0.304	0.055
<i>One_U1208-67</i>	B	0.431	0.405	0.068
<i>One_U1209-111</i>	B	0.249	0.220	0.107
<i>One_U1210-173</i>	B	0.192	0.181	0.051
<i>One_U1212-106</i>	B	0.487	0.421	0.134
<i>One_U1214-107</i>	B	0.091	0.083	0.081
<i>One_U1216-230</i>	B	0.458	0.400	0.113
<i>One_U301-92</i>	A	0.285	0.258	0.093
<i>One_U401-224</i>	A	0.480	0.443	0.083
<i>One_U404-229</i>	A	0.079	0.068	0.122
<i>One_U502-167</i>	A	0.046	0.045	0.040
<i>One_U503-170</i>	A	0.256	0.235	0.087
<i>One_U504-141</i>	A	0.385	0.356	0.068
<i>One_vamp5-255</i>	C	0.374	0.343	0.083
<i>One_vatf-214</i>	C	0.067	0.059	0.119
<i>One_VIM-569</i>	A	0.204	0.184	0.081
<i>One_ZNF-61</i>	A	0.433	0.371	0.145
<i>One_Zp3b-49</i>	A	0.197	0.139	0.293
<i>One_CO1_Cytb17_26</i> ²	A			0.263
<i>One_MHC2_190_251</i> ²				0.243
<i>Overall</i>				0.116

720

721 ¹ A) Gene Conservation Laboratory of the Alaska Department of Fish and Game; B) International Program for
722 Salmon Ecological Genetics at the University of Washington; C) Hagerman Genetics Laboratory of the Columbia
723 River Inter-Tribal Fish Commission; and D) Molecular Genetics Laboratory at the Canadian Department of
724 Fisheries and Oceans.

725 ² These SNPs were combined into haplotypes and treated together as single loci, *One_CO1_Cytb17_26* and
726 *One_MHC2-190_251*.

727 ³ These SNPs were dropped due to linkage.

728 Table 4. Quality control (QC) results including the number of genotypes compared, discrepancy rates and
 729 estimated error rates of the collections genotyped for the WASSIP sockeye salmon baseline for the four
 730 methods used: “Old”, “New”, “39”, and “Assay”. See text for descriptions of methods and QC details.
 731 Discrepancy rates include the rate due to differences of alternate homozygote genotypes (Homo-homo),
 732 of homozygote and heterozygote genotypes (Homo-het) and the total discrepancy rate. Error rate assumes
 733 that differences are the result of errors that are equally likely to have occurred in the production and QC
 734 genotyping process.

QC Method	Genotypes compared	Discrepancy rate			Error Rate
		Homo-homo	Homo-het	Overall	
Old	8,448	0.00%	0.24%	0.24%	0.12%
New	62,400	0.01%	0.12%	0.13%	0.07%
39	1,172,836	0.02%	0.34%	0.36%	0.18%
Assay	89,760	0.00%	0.13%	0.14%	0.07%
Total	1,333,444	0.02%	0.31%	0.33%	0.17%

735

736 Table 5. Pairs of single nucleotide polymorphisms (SNPs) that exhibited significant ($P < 0.01$) linkage
 737 disequilibrium in 289 populations of sockeye salmon in the WASSIP study area, f_{ORCA} values for each
 738 locus separate as well as for combined loci, and decision for handling linkage for each locus pair based
 739 upon the Δ_{90} of 0.0091 (see text for details).

Locus	Linkage pair	f_{ORCA}	Decision
<i>One_GPDH-201</i>	1	0.042	Drop
<i>One_GPDH2-1872</i>	1	0.056	Keep
<i>One_GPDH-201_GPDH2-1872</i>	1	0.064	Do not combine
<i>One_MHC2_190</i>	2	0.037	Drop
<i>One_MHC2_251</i>	2	0.035	Drop
<i>One_MHC2_190_251</i>	2	0.046	Combine
<i>One_Tf_ex11-750</i>	3	0.039	Drop
<i>One_Tf_in3-182</i>	3	0.041	Keep
<i>One_Tf_ex11-750_in3-182</i>	3	0.043	Do not combine

740 Table 6. Estimates of stock composition, upper and lower 90% credibility interval bounds, and standard deviations for mixtures of known-origin
 741 fish removed from the WASSIP baseline populations of sockeye salmon that comprise the regional reporting groups that are not sub-divided into
 742 smaller sub-regional reporting groups (Norton Sound, South Peninsula and East of WASSIP; i.e., 100% proof tests) using the program BAYES
 743 with a flat prior. One hundred fish were removed from the Norton Sound group while 200 fish were removed from the South Peninsula and East of
 744 WASSIP groups. Correct allocations are in bold.

Reporting Group	Norton Sound				South Peninsula				East of WASSIP			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.95	0.90	0.99	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.01	0.00	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Bristol Bay	0.03	0.00	0.08	0.02	0.01	0.00	0.02	0.01	0.01	0.00	0.02	0.01
North Peninsula	0.00	0.00	0.02	0.01	0.00	0.00	0.02	0.01	0.00	0.00	0.02	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.99	0.97	1.00	0.01	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.97	1.00	0.01

745

746 Table 7. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for mixtures of 200 known-origin fish
 747 removed from the WASSIP baseline populations of sockeye salmon that comprise the Kuskokwim Bay reporting groups (i.e., 100% proof tests)
 748 using the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Kuskokwim River				Kanektok				Goodnews			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim River	0.93	0.88	0.97	0.03	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.01
Kanektok	0.01	0.00	0.04	0.01	0.98	0.95	1.00	0.02	0.02	0.00	0.05	0.02
Goodnews	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.01	0.82	0.75	0.89	0.04
Bristol Bay	0.06	0.00	0.11	0.03	0.01	0.00	0.03	0.01	0.15	0.09	0.22	0.04
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

749

750 Table 8. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for mixtures of 200 known-origin fish
 751 removed from the WASSIP baseline populations of sockeye salmon that comprise the Bristol Bay reporting groups (i.e., 100% proof tests) using
 752 the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Togiak				Igushik				Wood			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.03	0.00	0.11	0.04	0.01	0.00	0.03	0.01	0.00	0.00	0.01	0.00
Togiak	0.97	0.88	1.00	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Igushik	0.00	0.00	0.00	0.00	0.98	0.93	1.00	0.02	0.00	0.00	0.01	0.01
Wood	0.00	0.00	0.00	0.00	0.01	0.00	0.05	0.02	0.99	0.97	1.00	0.01
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

753

754 Table 8: Page 2 of 3.

Reporting Group	Nushagak				Kvichak				Alagnak			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.01	0.00	0.04	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.01	0.00	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nushagak	0.97	0.92	1.00	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.98	0.96	1.00	0.01	0.00	0.00	0.00	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.98	1.00	0.01
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.02	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00

755

756 Table 8: Page 3 of 3.

Reporting Group	Naknek				Egegik				Ugashik			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.99	0.98	1.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.82	0.70	0.93	0.07	0.00	0.00	0.03	0.02
Ugashik	0.00	0.00	0.00	0.00	0.18	0.06	0.29	0.07	0.98	0.95	1.00	0.02
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

757

WASSIP Technical Document 14: Sockeye salmon baseline based upon 96 SNPs

758 Table 9. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for mixtures of known-origin fish
 759 removed from the WASSIP baseline populations of sockeye salmon that comprise the North Peninsula reporting groups (i.e., 100% proof tests)
 760 using the program BAYES with a flat prior. One hundred fish were removed from the Cinder, Sandy and Nelson River groups while 200 fish
 761 were removed from the Meshik, Ilnik, Bear, and NW District-Black Hills groups. Correct allocations are in bold.

Reporting Group	Cinder				Meshik				Ilnik			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Bristol Bay	0.01	0.00	0.03	0.01	0.00	0.00	0.02	0.01	0.00	0.00	0.01	0.00
Cinder	0.98	0.92	1.00	0.03	0.01	0.00	0.07	0.02	0.00	0.00	0.03	0.01
Meshik	0.01	0.00	0.06	0.02	0.98	0.92	1.00	0.03	0.00	0.00	0.01	0.00
Ilnik	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.99	0.96	1.00	0.01
Sandy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bear	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nelson River	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NW District-Black Hills	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
South Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

762

763 Table 9 cont'd: Page 2 of 3.

Reporting Group	Sandy				Bear				Nelson River			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Bristol Bay	0.04	0.01	0.08	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.01
Cinder	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Meshik	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ilnik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sandy	0.93	0.88	0.97	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bear	0.01	0.00	0.04	0.01	0.99	0.98	1.00	0.01	0.03	0.00	0.06	0.02
Nelson River	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.96	0.92	0.99	0.02
NW District-Black Hills	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

764

765 Table 9 cont'd: Page 3 of 3.

Reporting Group	NW District-Black Hills			
	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00
Bristol Bay	0.00	0.00	0.01	0.00
Cinder	0.00	0.00	0.00	0.00
Meshik	0.00	0.00	0.00	0.00
Ilnik	0.00	0.00	0.00	0.00
Sandy	0.00	0.00	0.00	0.00
Bear	0.00	0.00	0.00	0.00
Nelson River	0.00	0.00	0.00	0.00
NW District-Black Hills	0.99	0.97	1.00	0.01
South Peninsula	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00
East of WASSIP	0.01	0.00	0.02	0.01

766

767 Table 10. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for mixtures of 200 known-origin
 768 fish removed from the WASSIP baseline populations of sockeye salmon that comprise the Chignik reporting groups (i.e., 100% proof tests) using
 769 the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Black Lake				Chignik Lake			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bristol Bay	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Black Lake	0.99	0.98	1.00	0.01	0.00	0.00	0.01	0.01
Chignik Lake	0.00	0.00	0.01	0.01	0.99	0.98	1.00	0.01
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

770 Table 11. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for samples of escapement to rivers
 771 within the Kuskokwim Bay reporting groups (i.e., escapement tests) using the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Bethel Test Fish 2010				KogrukluK Weir 2001				KogrukluK Weir 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim River	0.98	0.97	1.00	0.01	0.99	0.97	1.00	0.01	0.94	0.88	0.98	0.03
Kanektok	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Goodnews	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bristol Bay	0.01	0.00	0.02	0.01	0.00	0.00	0.02	0.01	0.06	0.02	0.11	0.03
North Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

772

773 Table 11 cont'd: Page 2 of 3.

Reporting Group	Kwethluk Weir 2007				Kanektok Weir 2002				Kanektok Weir 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01
Kuskokwim River	0.49	0.39	0.62	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kanektok	0.00	0.00	0.01	0.01	0.98	0.95	1.00	0.02	0.98	0.94	1.00	0.02
Goodnews	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01
Bristol Bay	0.48	0.35	0.58	0.07	0.01	0.00	0.04	0.01	0.01	0.00	0.04	0.01
North Peninsula	0.02	0.00	0.05	0.02	0.00	0.00	0.01	0.01	0.00	0.00	0.02	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

774

775 Table 11 cont'd: Page 3 of 3.

Reporting Group	Goodnews Weir 2001				Goodnews Weir 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim River	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kanektok	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01
Goodnews	0.95	0.85	1.00	0.05	0.83	0.75	0.91	0.05
Bristol Bay	0.03	0.00	0.13	0.05	0.15	0.08	0.23	0.04
North Peninsula	0.01	0.00	0.03	0.01	0.01	0.00	0.04	0.01
South Peninsula	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00

776

777 Table 12. Estimates of stock composition, upper and lower 90% credibility intervals, and standard deviations for samples of escapement to rivers
 778 within the Bristol Bay reporting groups (i.e., escapement tests) using the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Togiak 2008 Subsistence				Igushik Tower 2005				Igushik Tower 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.21	0.08	0.32	0.07	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.01
Togiak	0.79	0.67	0.92	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.00	0.00	0.91	0.80	1.00	0.06	0.91	0.76	1.00	0.08
Wood	0.00	0.00	0.00	0.00	0.09	0.00	0.19	0.06	0.08	0.00	0.23	0.08
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.01
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

779

780 Table 12 cont'd: Page 2 of 7.

Reporting Group	Wood Tower 2003				Wood Tower 2004				Wood Tower 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.02	0.00	0.10	0.03	0.01	0.00	0.04	0.02	0.01	0.00	0.06	0.02
Wood	0.93	0.85	0.97	0.04	0.98	0.92	1.00	0.03	0.98	0.93	1.00	0.02
Nushagak	0.05	0.02	0.08	0.02	0.01	0.00	0.05	0.02	0.00	0.00	0.01	0.01
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

781

782 Table 12 cont'd: Page 3 of 7.

Reporting Group	Nushagak Sonar 2006				Nushagak Radio Telemetry 2005				Nushagak Radio Telemetry 2006			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.02	0.01	0.13	0.05	0.21	0.05	0.00	0.00	0.02	0.01
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.03	0.01	0.02	0.00	0.05	0.02	0.01	0.00	0.05	0.02
Wood	0.02	0.00	0.04	0.01	0.01	0.00	0.03	0.01	0.00	0.00	0.03	0.01
Nushagak	0.97	0.93	0.99	0.02	0.84	0.75	0.92	0.05	0.97	0.93	1.00	0.02
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00

783

784 Table 12 cont'd: Page 4 of 7

Reporting Group	Nuyakuk Tower 2004				Kvichak Tower 2005				Kvichak Tower 2006			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.07	0.02	0.13	0.03	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nushagak	0.92	0.85	0.97	0.03	0.04	0.01	0.07	0.02	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.95	0.91	0.98	0.02	0.99	0.98	1.00	0.01
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

785

786 Table 12 cont'd: Page 5 of 7.

Reporting Group	Alagnak Tower 2004				Alagnak Tower 2008				Naknek Tower 2002			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Alagnak	0.99	0.98	1.00	0.01	0.99	0.97	1.00	0.01	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.99	0.98	1.00	0.01
Egegik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Ugashik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

787

788 Table 12 cont'd: Page 6 of 7.

Reporting Group	Naknek Tower 2008				Egegik Tower 2004				Egegik Tower 2007			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.00
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Alagnak	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.99	0.97	1.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.01	0.00	0.97	0.91	1.00	0.03	0.99	0.97	1.00	0.01
Ugashik	0.00	0.00	0.00	0.00	0.02	0.00	0.08	0.03	0.00	0.00	0.01	0.00
North Peninsula	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.01
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

789

790 Table 12 cont'd: Page 7 of 7.

Reporting Group	Ugashik Tower 2004				Ugashik Tower 2008			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Togiak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Igushik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Wood	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nushagak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kvichak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Alagnak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Naknek	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Egegik	0.00	0.00	0.03	0.02	0.06	0.00	0.21	0.08
Ugashik	0.99	0.95	1.00	0.02	0.93	0.78	1.00	0.08
North Peninsula	0.01	0.00	0.02	0.01	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Chignik	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

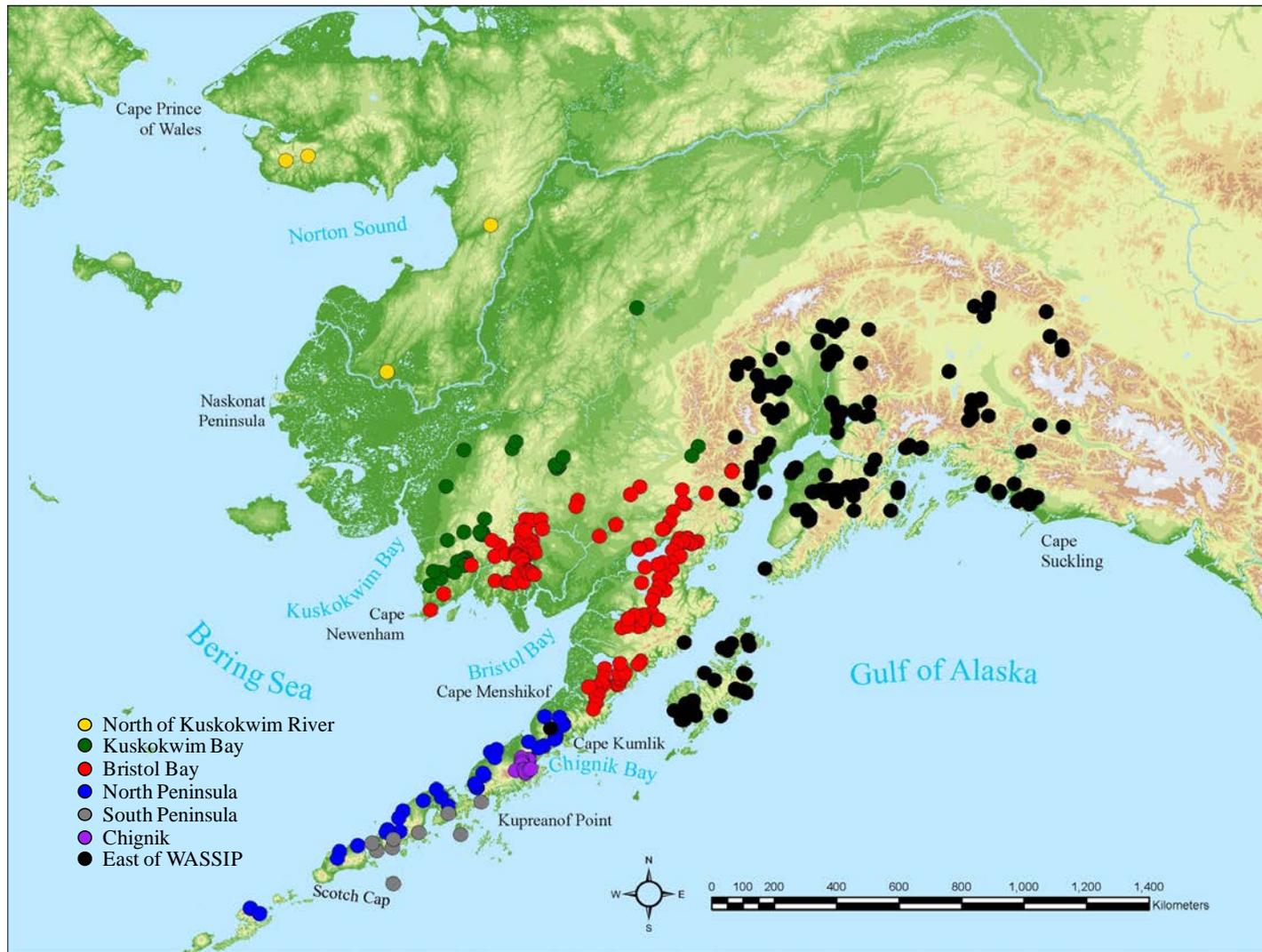
791

792 Table 13. Estimates of stock composition, upper and lower 90% credibility intervals, and standard
 793 deviations for samples of escapement to rivers within the Chignik reporting groups (i.e., escapement
 794 tests) using the program BAYES with a flat prior. Correct allocations are in bold.

Reporting Group	Chignik 2010 Early Run				Chignik 2010 Late Run			
	Proportion	Lower	Upper	SD	Proportion	Lower	Upper	SD
Norton Sound	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kuskokwim Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bristol Bay	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
North Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
South Peninsula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Black Lake	0.98	0.94	1.00	0.02	0.07	0.03	0.11	0.02
Chignik Lake	0.02	0.00	0.06	0.02	0.92	0.88	0.96	0.02
East of WASSIP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

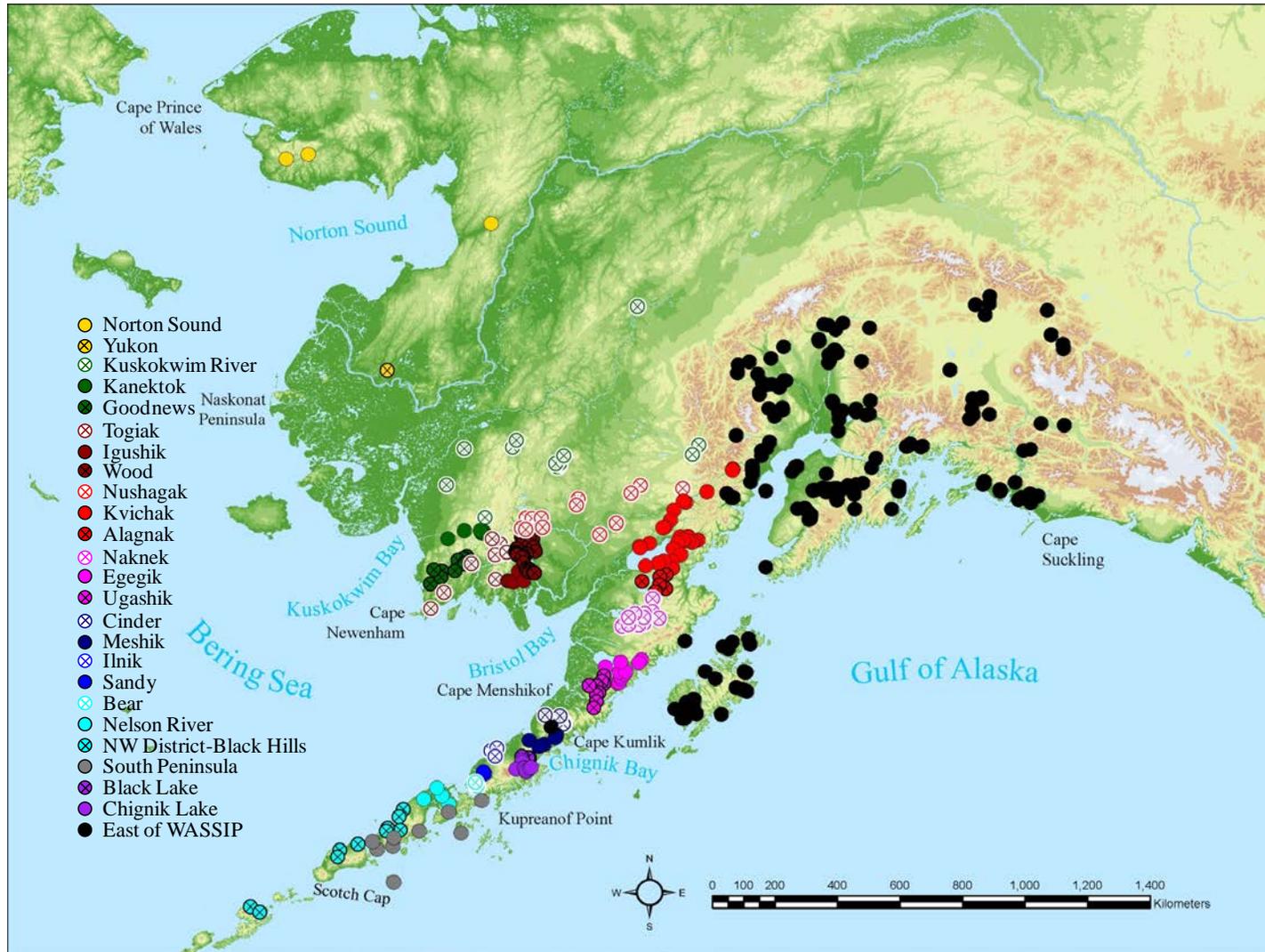
795

Figures



796

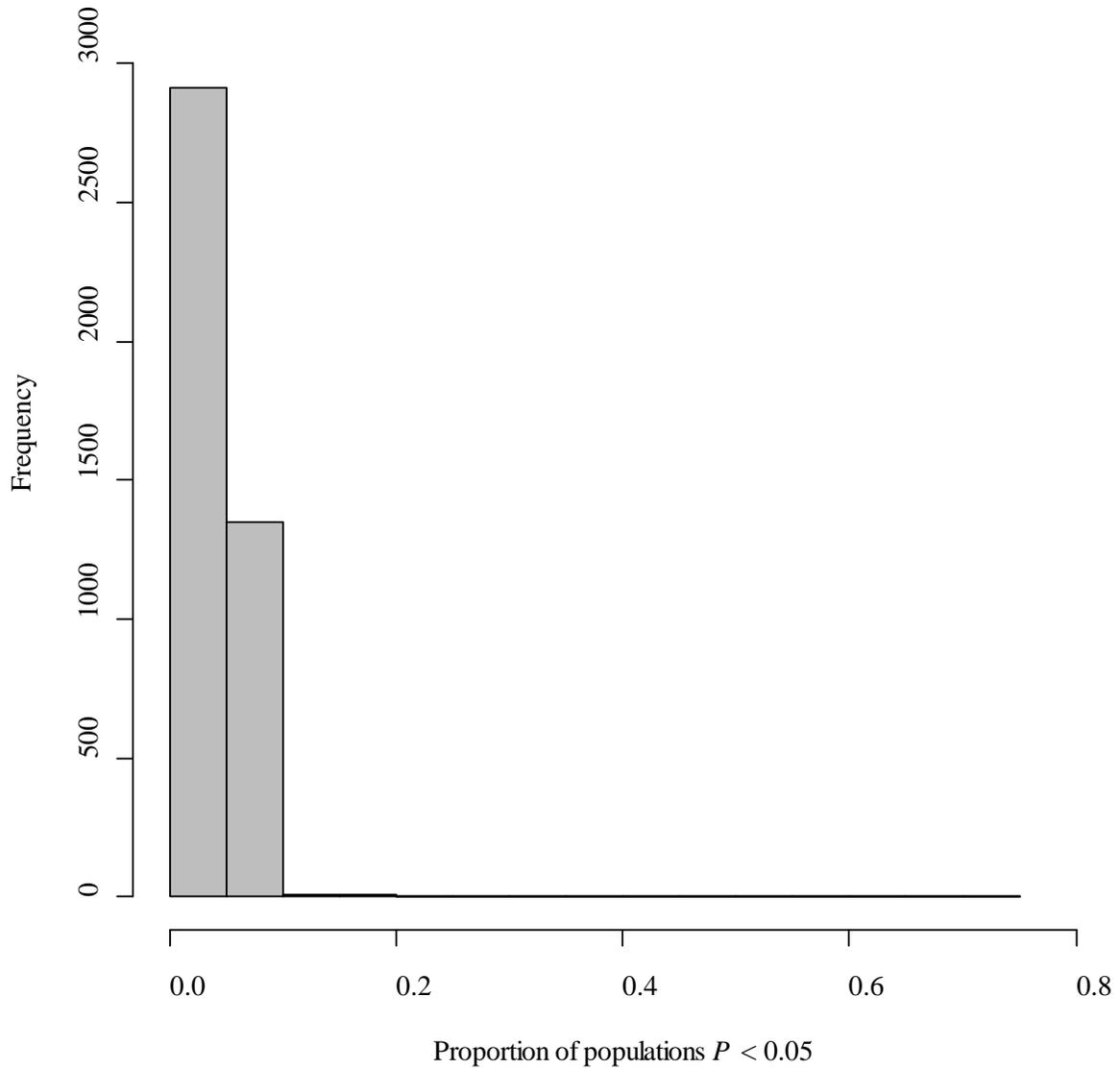
797 Figure 1. The location and regional reporting group affiliation of 472 collections of sockeye salmon initially included in baseline analyses for
798 WASSIP.



799

800 Figure 2. The location and sub-regional reporting group affiliation of 472 collections of sockeye salmon initially included in baseline analyses for
 801 WASSIP. Note that the Yukon reporting group was removed from final analyses.

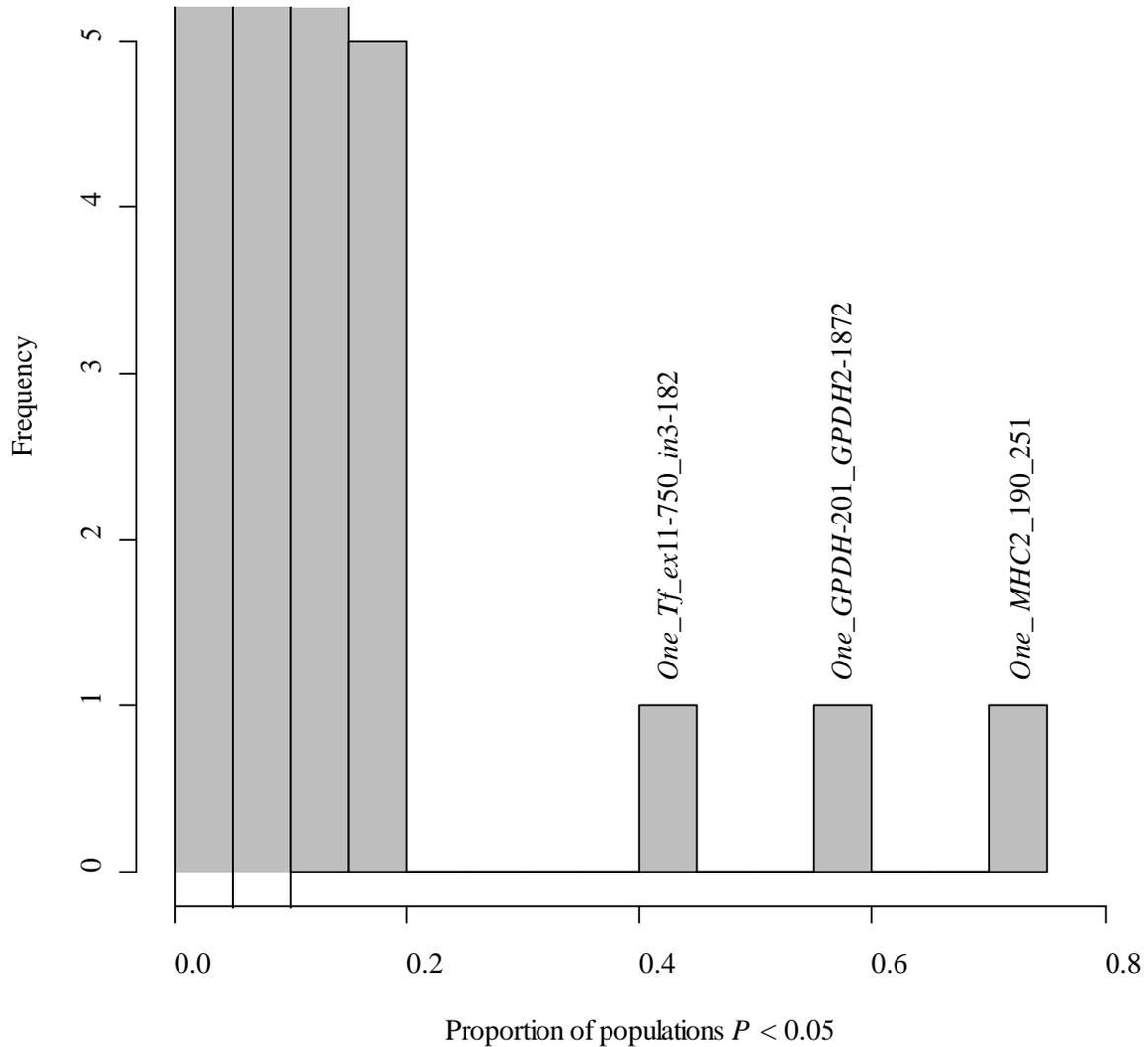
802



803

804 Figure 3. Histogram of the proportion of populations with significant ($P < 0.05$) linkage disequilibrium
805 between the 4,278 pairs of the 96 SNPs tested in 289 WASSIP area populations.

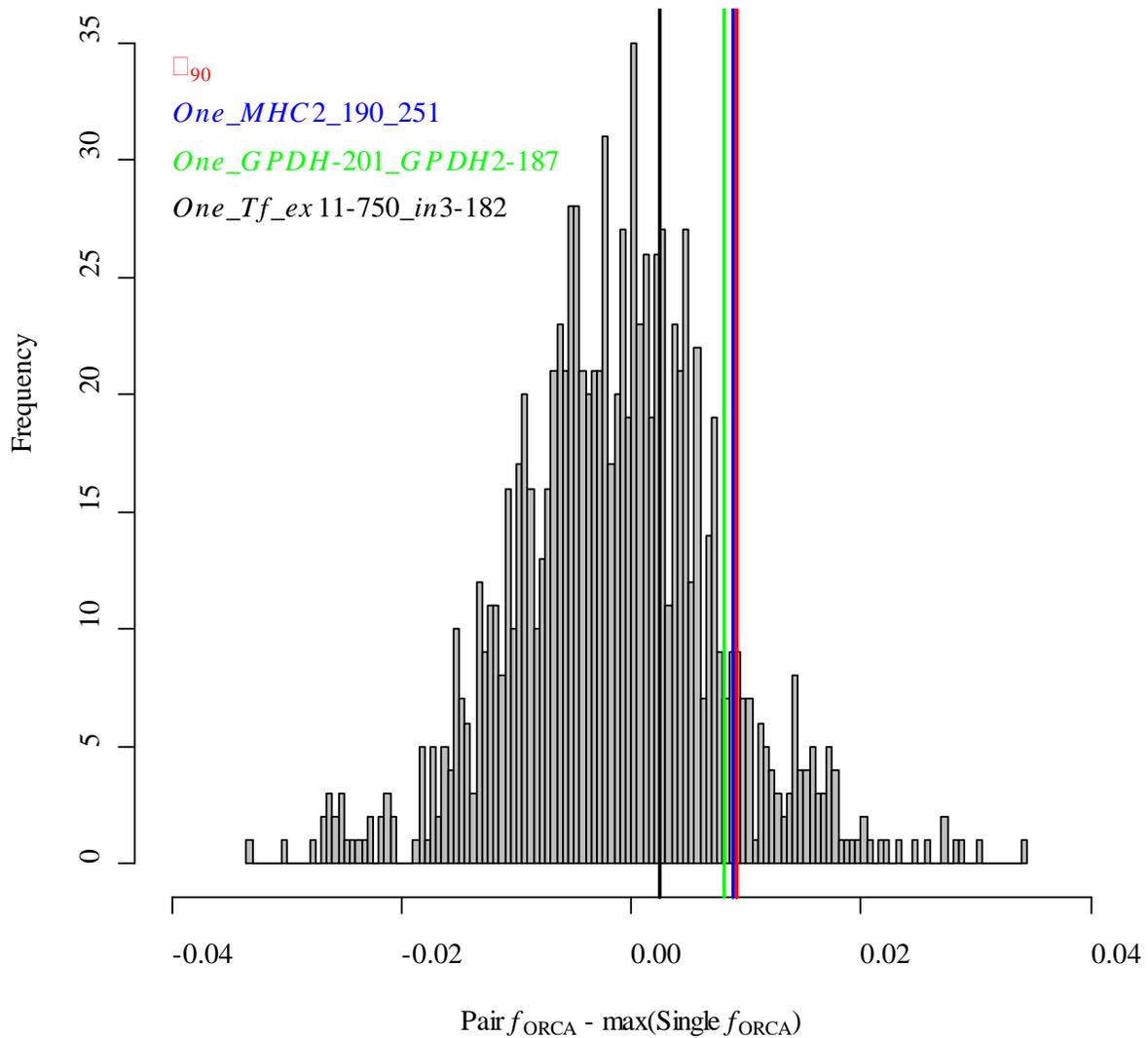
806



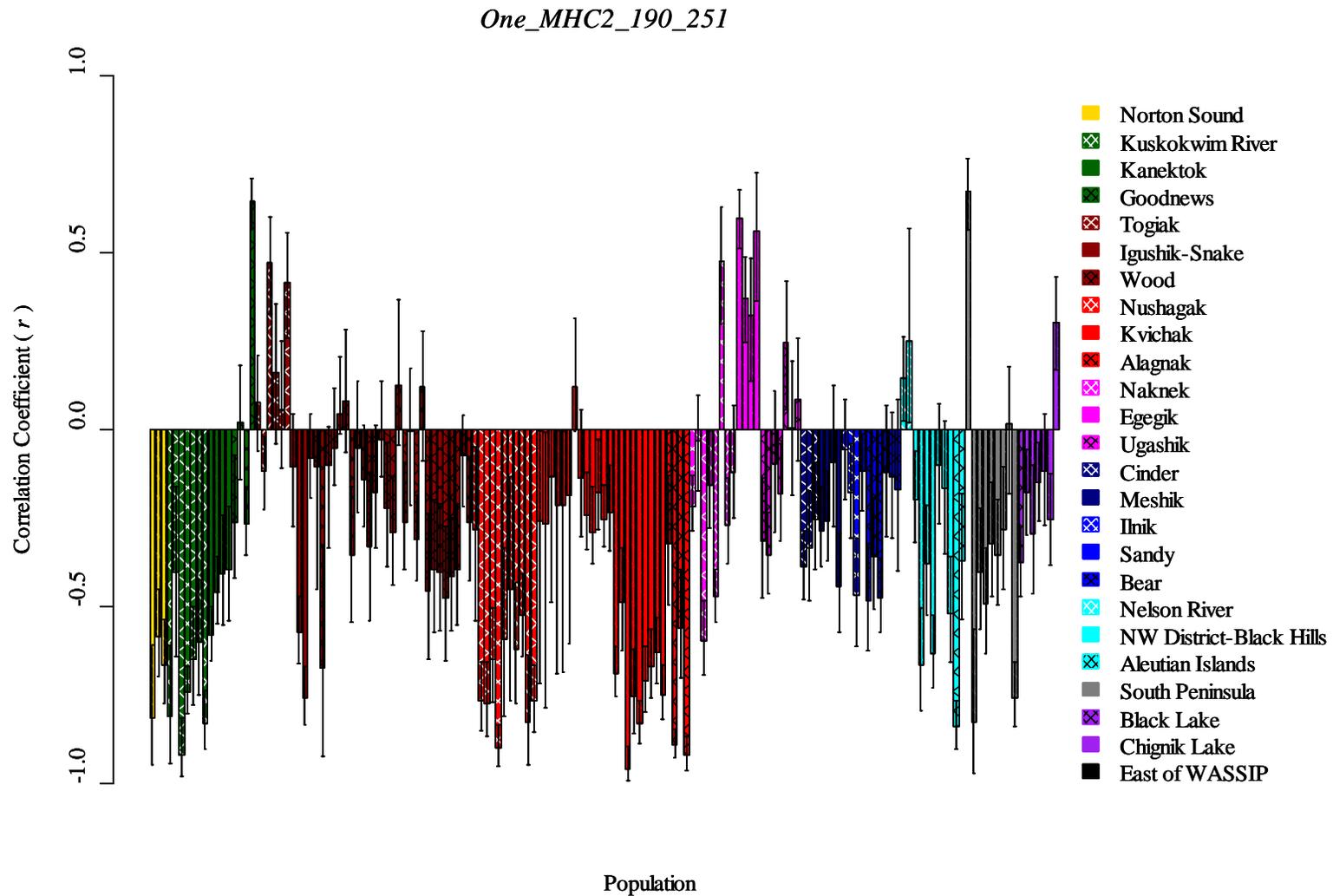
807

808 Figure 4. Detailed histogram of the proportion of populations with significant ($P < 0.05$) linkage
 809 disequilibrium between the 4,278 pairs of the 96 SNPs tested in 289 WASSIP area populations. Note
 810 three SNP pairs exhibit significant linkage disequilibrium in substantially more populations than most
 811 SNP pairs: *One_Tf_ex11-750* & *One_Tf_in3-182*, *One_GPDH-201* & *One_GPDH2-1872*, and
 812 *One_MHC2_190* & *One_MHC2_251*.

813

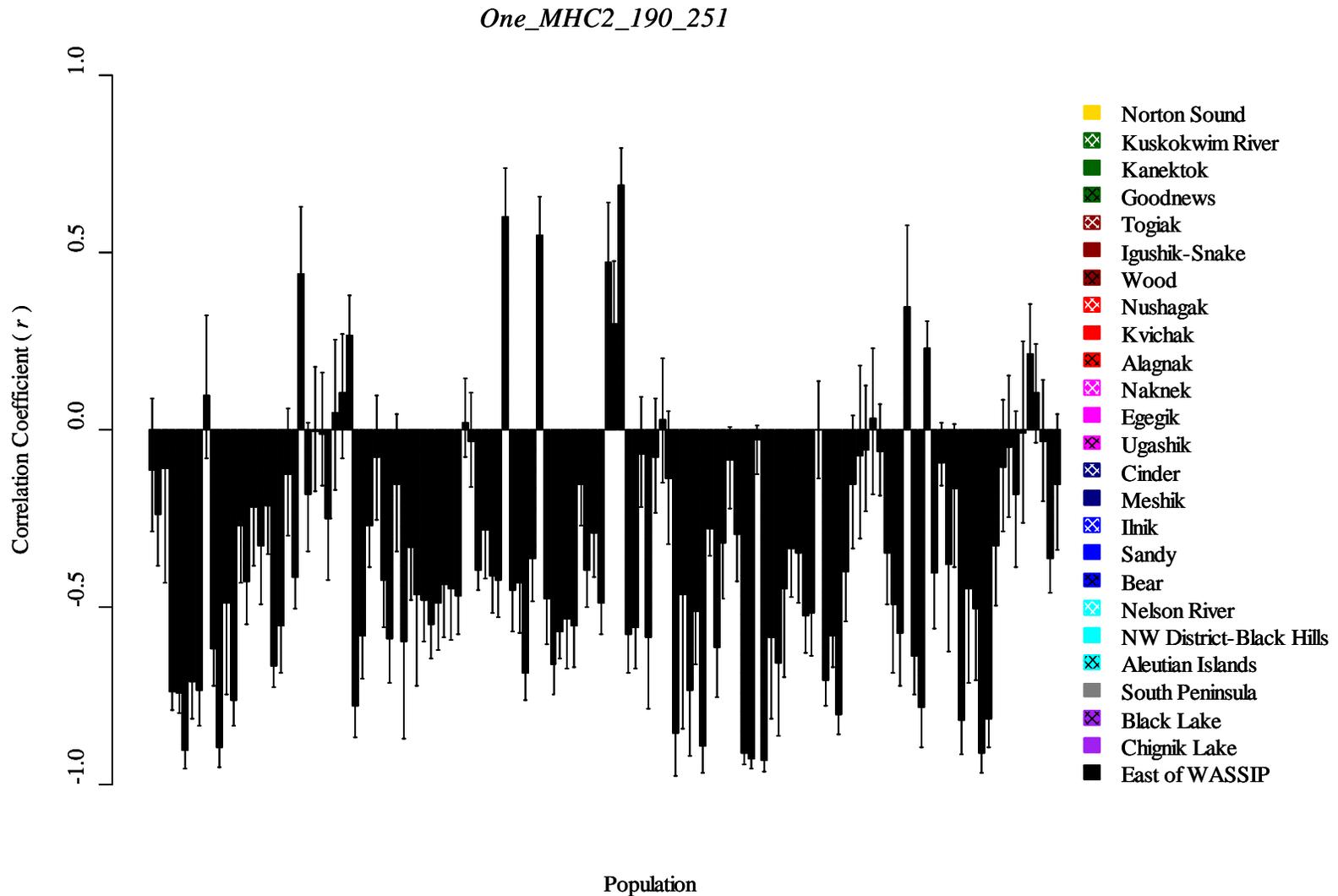


814
 815 Figure 5. The distribution of Δ for 1,000 random SNP pairs with Δ_{90} in red and the Δ values for
 816 *One_MHC2_190_251* in blue, *One_GPDH-201_GPDH2-187* in green, and *One_Tf_ex11-750_in3-182* in
 817 black. See text for details.



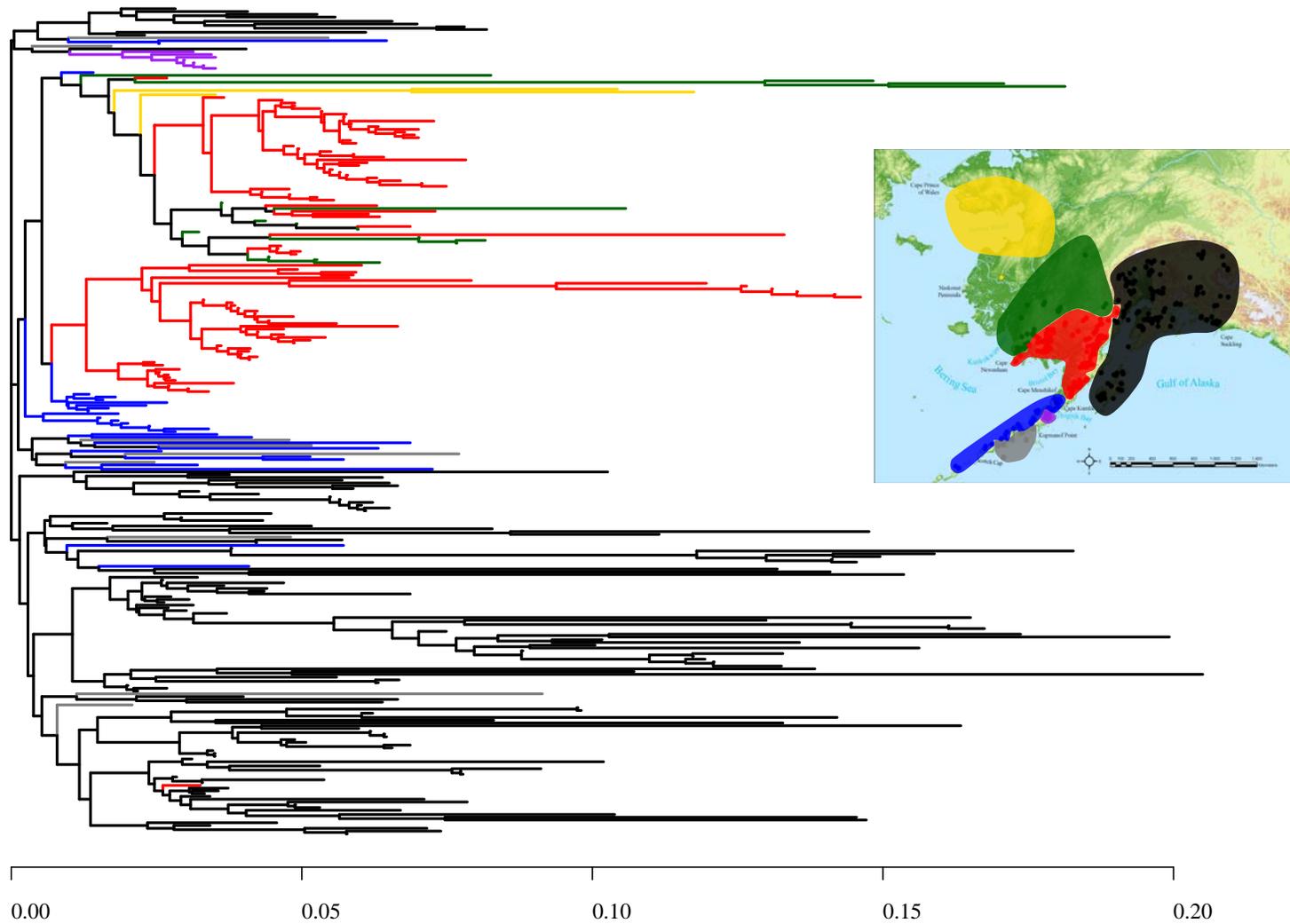
818

819 Figure 6. Correlation coefficient r between the first alphabetical allele in the pair of MHC SNPs in each population within the WASSIP study area
 820 ranging from Salmon Lake in the north (left) to the Chignik River to the south (right).



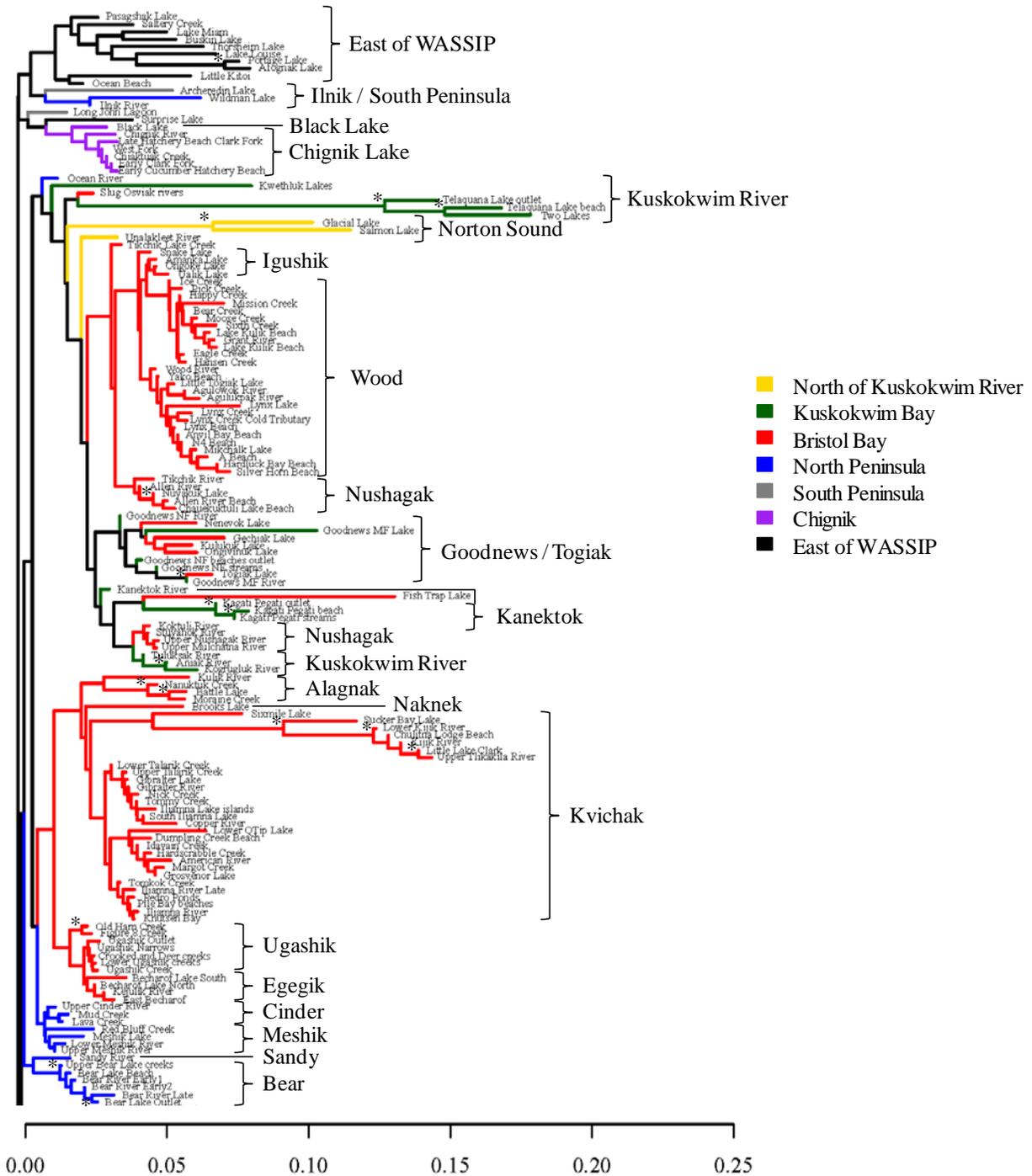
821

822 Figure 7. Correlation coefficient r between the first alphabetical allele in the pair of MHC SNPs in each population in the East of WASSIP
 823 reporting group ranging from Surprise Lake in the west (left) to Bering Lake to the east (right).



824

825 Figure 8. Neighbor-joining tree based upon pairwise F_{ST} between populations of sockeye salmon included in the WASSIP baseline and map
826 denoting the regional reporting group colors represented on tree branches.



827

828 Figure 9. Page one of two of a consensus neighbor-joining tree based upon pairwise F_{ST} between
 829 populations of sockeye salmon included in the WASSIP baseline. Tree branch colors denote regional
 830 reporting group memberships, text brackets denote general population groupings by sub-regional
 831 reporting group and asterisks indicate nodes where bootstrap consensus > 90%.

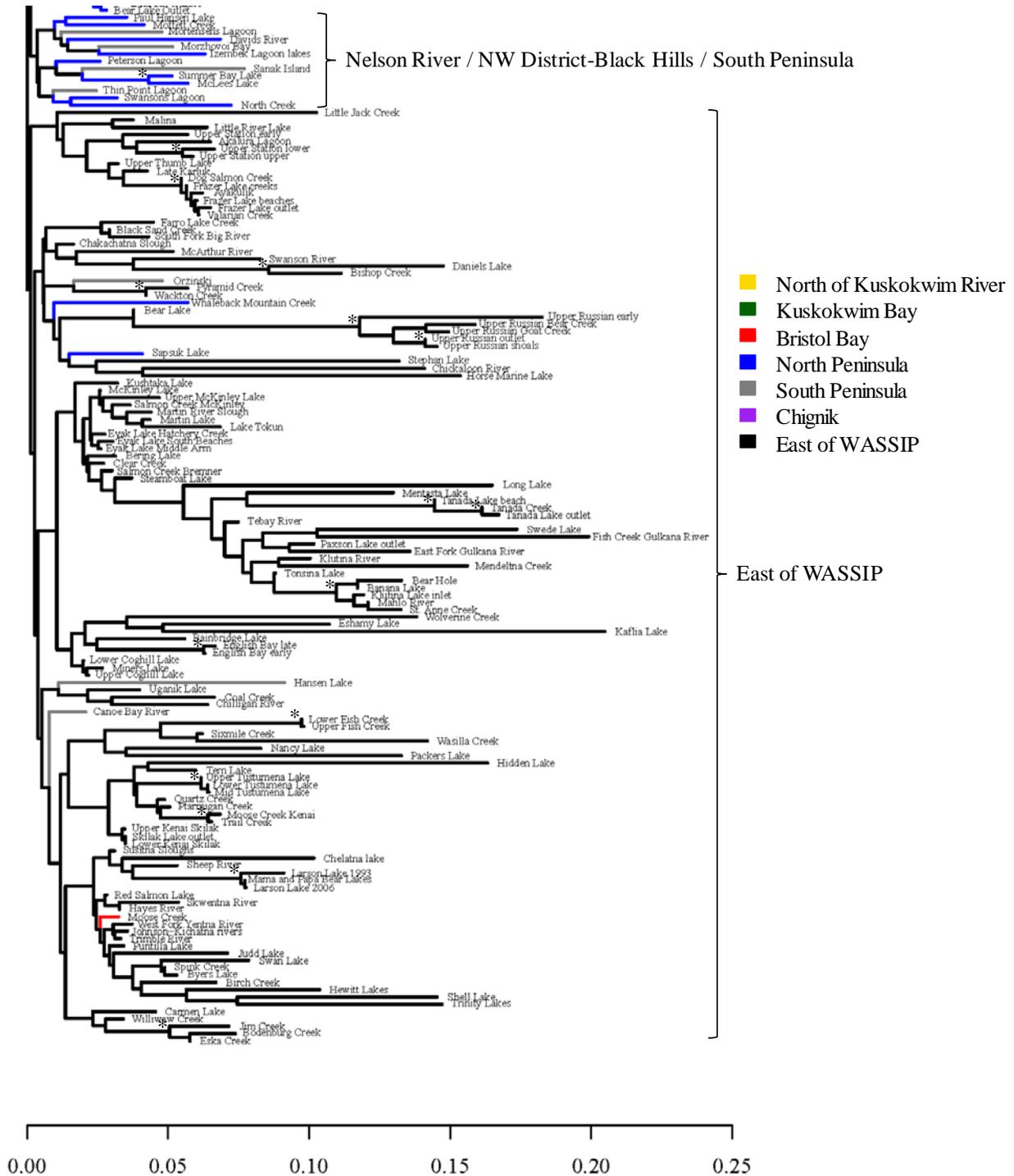
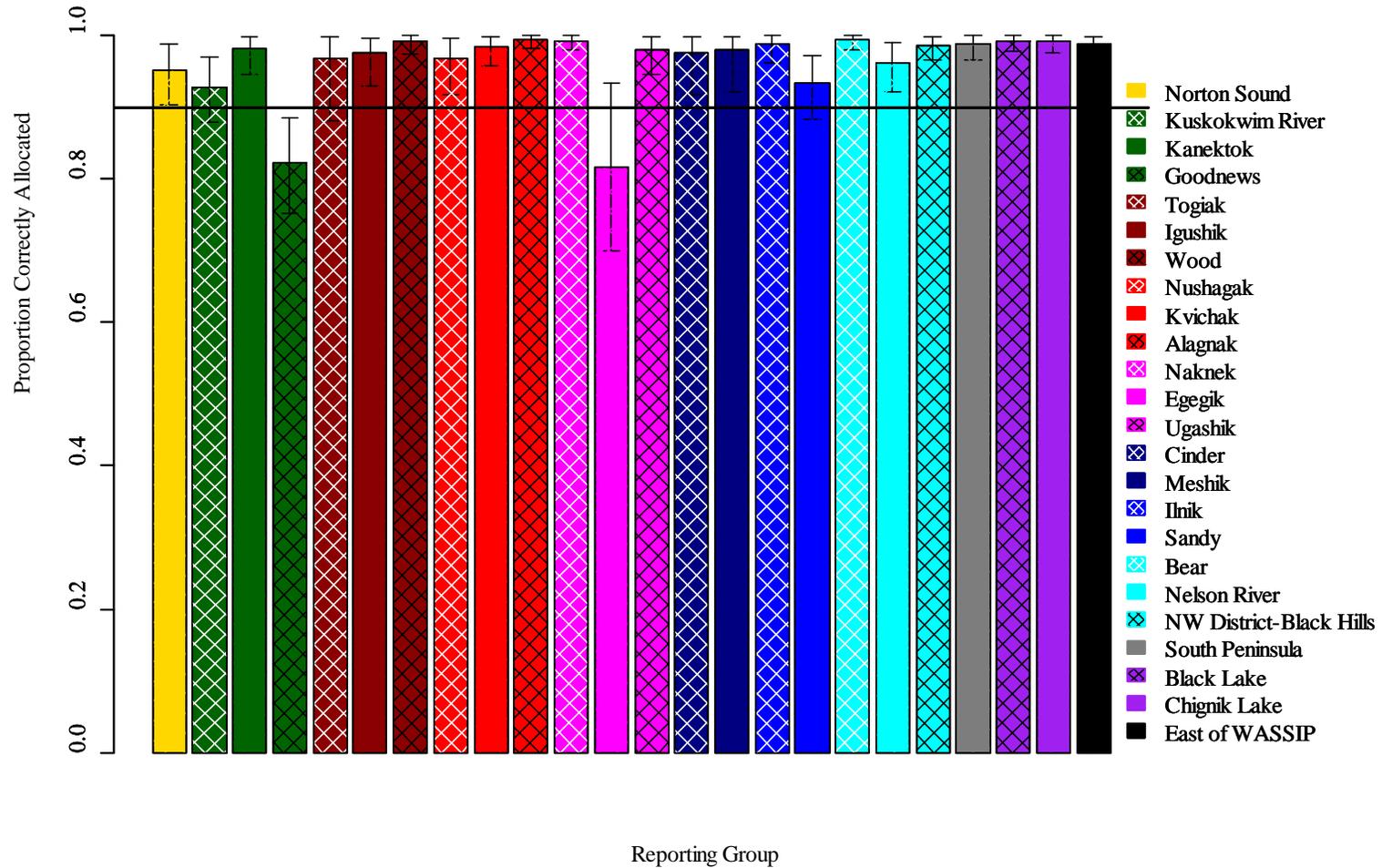
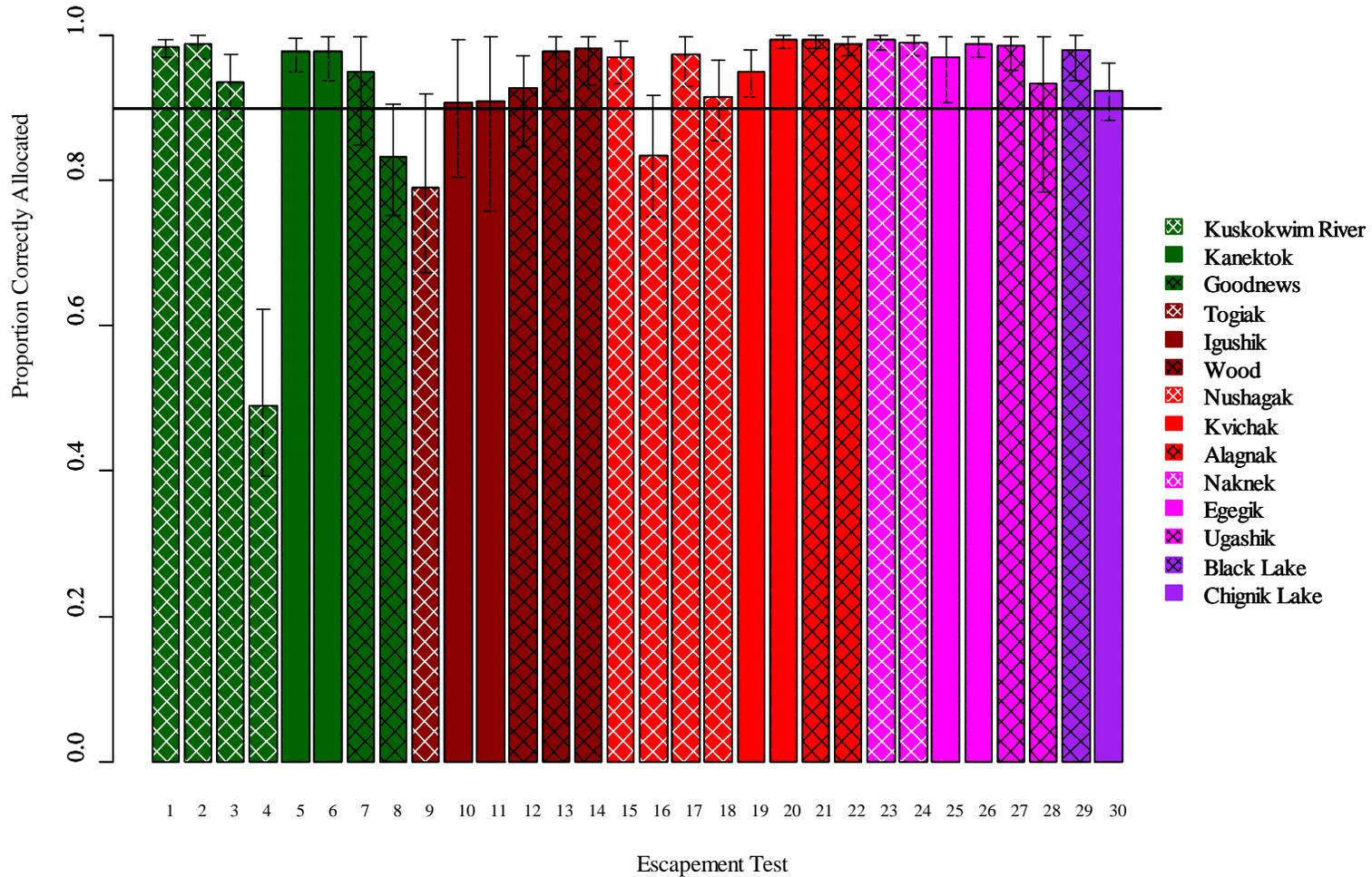


Figure 10. Page two of two of a consensus neighbor-joining tree based upon pairwise F_{ST} between populations of sockeye salmon included in the WASSIP baseline. Tree branch colors denote regional reporting group memberships, text brackets indicate general population groupings by sub-regional reporting group and asterisks indicate nodes where bootstrap consensus > 90%.



837

838 Figure 11. Proportion of fish correctly allocated back to sub-regional reporting group of origin and 90% credibility intervals for mixtures of known
 839 individuals removed from the baseline population that comprise each reporting group (100% proof tests) using the program *BAYES* with a flat
 840 prior. One hundred individuals were removed from the Norton Sound, Cinder, Sandy and Nelson River groups, while 200 individuals were
 841 removed from all others.



842

843 Figure 12. Proportion of fish allocated to sub-regional reporting group of origin and 90% credibility intervals for samples of the escapement of
 844 sockeye salmon to rivers within 14 WASSIP area reporting groups using the program *BAYES* with a flat prior. Escapement test numbers refer to
 845 tests detailed in Table 2.