A Comparison of Stock and Individual Identification for Chinook Salmon in British Columbia Provided by Microsatellites and Single-Nucleotide Polymorphisms

Authors: Beacham, Terry D., Jonsen, K., and Wallace, C.

ARTICLE

A Comparison of Stock and Individual Identification for Chinook Salmon in British Columbia Provided by Microsatellites and Single-Nucleotide Polymorphisms

Terry D. Beacham,* K. Jonsen, and C. Wallace

Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada

Abstract

The following questions were addressed in this study: (1) If a suite of 12–15 microsatellites were used in the genetic stock identification (GSI) of Chinook salmon *Oncorhynchus tshawytscha*, which microsatellites should be in the suite? (2) How many microsatellites are required to provide stock identification resolution equivalent to that of 72 single-nucleotide polymorphisms (SNPs)? (3) How many SNPs are required to replace the current microsatellite baselines used in GSI applications? (4) If additional GSI power is required for microsatellite baselines, what is the incremental increase provided by SNPs and microsatellites? The variation at 29 microsatellite loci and 73 SNP loci was surveyed in 60 populations of Chinook salmon in 16 regions in British Columbia. Microsatellites with more observed alleles provided more accurate estimates of stock composition than those with fewer alleles. The options available for improving the accuracy and precision of stock composition estimates for a 12-locus Fisheries and Oceans Canada (DFO) microsatellite suite range include adding either 4 microsatellites or 25 SNPs to the existing suite to achieve an overall population-specific accuracy of 86% across 60 populations. For the 13-locus Genetic Analysis of Pacific Salmon (GAPS) microsatellites, either 2 microsatellites or 20–25 SNPs can be added to the existing suite to achieve approximately 86% population-specific accuracy in estimated stock composition. The enhanced DFO (16 loci) and GAPS (15 loci) microsatellite baselines were projected to require 179 and 166 SNPs, respectively, for equivalent precision of the population-specific estimates. The level of regional accuracy of individual assignment available from the enhanced DFO and GAPS suites of microsatellites was projected to require 90 and 82 SNPs, respectively. The level of individual assignment to specific populations available from the enhanced DFO and GAPS suites of microsatellites was projected to require 137 and 121 SNPs, respectively.

One key aspect of the management of Pacific salmon *Oncorhynchus* spp. fisheries is estimating the stock composition of mixed-stock fishery samples with enough resolution for effective management decisions, with the constraint that the estimates be timely and cost-effective. Prior to 2002, microsatellites had been successfully applied in sockeye salmon *O. nerka* fishery management in British Columbia, where the twin management objectives were to restrict exploitation of populations of conservation concern while enabling the harvest of abundant populations (Beacham et al. 2004). In 2002, this was precisely the dilemma confronting Canadian fishery managers in the management of the Chinook salmon *O. tshawytscha* fisheries off the Queen Charlotte Islands in northern British Columbia and the west coast of Vancouver Island in southern British Columbia, as only a small portion of the quota available to Canadian fishermen from 1995 to 2001 was harvested due to conservation concerns about specific Chinook salmon stocks. Microsatellite variation had been surveyed previously for stocks likely to be present in these two fisheries (Beacham et al. 2003, 2006). Beginning in 2002, with managers' knowledge of the timing and locations

1

of specific stocks of Chinook salmon derived from previous or in-season microsatellite-based stock composition analysis, fisheries were managed with the intent of allowing Canadian fishermen to harvest the quota to which they were entitled under the Pacific Salmon Treaty (PST) while providing protection to stocks of conservation concern in Canada (Beacham et al. 2008c). This change in the management of Canadian fisheries led to increased emphasis on microsatellites for the stock identification of Chinook salmon. Subsequently, staff from U.S. agencies and Fisheries and Oceans Canada (DFO) formed a research group known as the Genetic Analysis of Pacific Salmon (GAPS), which agreed to develop a 13-locus microsatellite baseline to be used for the stock composition estimation of Chinook salmon in fisheries subject to the PST (Seeb et al. 2007). There was no formal evaluation of the genetic stock identification (GSI) power of the 13 microsatellites prior to their inclusion in the baseline. The set of microsatellites used by Beacham et al. (2006a) for previous stock composition estimation became known as the DFO loci; four of the loci in the DFO suite were also in the GAPS suite. Other than the comparison outlined by Beacham et al. (2008b) for 19 populations of Yukon River Chinook salmon, there has been no formal comparison of the relative power of the microsatellites in the GAPS and DFO baselines, nor has there been any comparison of the power of adding microsatellites currently not in either baseline.

While the GAPS microsatellite baseline was being developed, single-nucleotide polymorphisms (SNPs) were also being developed and promoted by some fishery management agencies as an alternative to microsatellites for salmon stock identification. The benefits of using SNPs were suggested to be ease of data standardization among laboratories, high throughput, high among-population diversity, lower genotyping errors, and lower cost of analysis per individual (Smith et al. 2005b, 2005d). However, there was no consensus among agency laboratories as to the preferred technique to use for GSI in Chinook salmon. In an evaluation of the two techniques with respect to Chinook salmon stock identification, it was concluded that comparisons between microsatellites and SNPs were necessary before conclusions could be drawn as to their relative efficacy (PSC 2008). Comparisons between microsatellites and SNPs have been conducted for other species and applications (Beacham et al. 2010; Glover et al. 2010; Haasl and Payseur 2010) and would be useful for Chinook salmon. The focus of any evaluation for Chinook salmon would be the resolution of the stock composition and individual identification estimates provided by the two techniques and the cost per individual required to obtain the observed resolution. An initial comparison of stock identification resolution was conducted for 19 Yukon River populations incorporating 30 microsatellites and 9 SNPs (Beacham et al. 2008b). In comparisons of population-specific estimates, a 9-SNP baseline was approximately equivalent to a single microsatellite locus with 17–22 alleles. In a subsequent study comparing the ability of the 13 GAPS microsatellites and 37 SNPs to differentiate 29 broadly distributed populations, closely related populations were better

differentiated by microsatellites than by SNPs but a combination of microsatellites and SNPs provided the most effective suite of loci for individual assignment to population (Narum et al. 2008). More comprehensive baselines, both in terms of the numbers of microsatellites and SNPs and the number of populations included in the analysis, need to be developed before definitive conclusions can be drawn about the resolution of stock composition estimates derived from the two techniques for Chinook salmon.

If a single approach to Chinook salmon GSI is to be implemented, the key question is how many SNPs must be used to provide stock composition and individual identification estimates of equivalent quality (in terms of both accuracy and precision) to that of the estimates obtained from a high-resolution microsatellite baseline. Kalinowski (2002, 2004) had previously suggested that equivalency in stock identification estimates could be obtained by using either a limited number of loci with many alleles or more loci with fewer alleles. Empirical evidence from microsatellites has indicated that loci with greater numbers of alleles generally provided more accurate and precise estimates of stock composition than loci with fewer alleles (Beacham et al. 2005, 2006, 2008b). Single-nucleotide polymorphisms generally display only two alleles, and thus individual SNPs will generally be less powerful than individual microsatellites in stock identification applications. The lesser power of individual SNPs can be compensated for by simply adding more SNPs to a GSI application, so that equivalency in the accuracy and precision of the estimated stock compositions is obtained. Once the number of SNPs is determined, evaluations of which technique produces the more cost-effective method of stock identification can be conducted within individual laboratories.

The genetic structure of Chinook salmon is generally regionally based, with populations in the same geographic area being more similar to each other than to populations in more distant areas (Waples et al. 2004; Beacham et al. 2006b). A regional genetic structure is the basis for defining reporting groups in GSI applications. In more complex applications, such as when fishery management actions are to be directed toward specific populations, differentiation among populations within a reporting group may be required (Parken et al. 2008). The final level of resolution required is the identification of individuals to specific populations in a reporting group or to specific reporting groups, and this is the most demanding aspect of GSI applications. Thus, the effectiveness of GSI techniques needs to be evaluated with respect to the accuracy and precision of their identification of regional reporting groups, specific populations, and individuals.

In the current study, 29 microsatellites were surveyed in 60 populations of Chinook salmon in British Columbia. These microsatellites included the 13 loci of the GAPS baseline, the 12 loci of the DFO baseline (4 of which are also in the GAPS baseline), and 8 additional microsatellites. We also evaluated the accuracy and precision of estimates of stock composition derived from a suite of 73 SNPs from the same 60-population baseline. The key questions investigated were as follows: (1) If

a suite of 12–15 microsatellites were used in Chinook salmon GSI applications, which microsatellites should be in the suite? (2) How many microsatellites are required to provide stock identification resolution equivalent to that of 72 SNPs? (3) How many SNPs are required to replace the current GAPS and DFO microsatellite baselines used in GSI applications? (4) If additional GSI power is required for either the GAPS or DFO microsatellite baselines, what is the increase in power provided by adding SNPs and microsatellites?

METHODS

Collection of DNA samples and laboratory analysis.—The sampling sites or populations surveyed in each geographic region are outlined in Table 1. The geographic locations of the populations listed in Table 1 are indicated in Figure 1. Tissue samples were collected from mature Chinook salmon in these populations, preserved in 95% ethanol, and sent to the Molecular Genetics Laboratory at the Pacific Biological Station. DNA was

extracted from the tissue samples using a variety of methods, including the chelex resin protocol outlined by Small et al. (1998), the Qiagen 96-well Dneasy procedure, and the Promega Wizard SV96 Genomic DNA Purification system. Once extracted DNA was available, surveys of the variation at the following 29 microsatellite loci were conducted: *Ots100, Ots101, Ots104, Ots107* (Nelson and Beacham 1999), *Ssa197* (O'Reilly et al. 1996), *Ogo2, Ogo4* (Olsen et al. 1998), *Oke4* (Buchholz et al. 2001), *Omy325* (O'Connell et al. 1997), *Oki100* (Beacham et al. 2008a), *Omm1009, Omm1037* (Rexroad et al. 2002), *Omm1080* (Rexroad et al. 2001), *Ots201b, Ots208b, Ots211, Ots212, Ots213* (Grieg et al. 2003), *Omm1276, Bhms417* (Danzmann et al. 2005), *Ots2, Ots9* (Banks et al. 1999), *Ots3M* (Grieg and Banks 1999), *Oki10* (Smith et al. 1998), *OtsG474, OtsG68* (Williamson et al. 2002), *OmyRGT3TUF, OmyRGT30TUF* (Sakamoto et al. 2000), and *Ssa408* (Cairney et al. 2000).

In general, polymerase chain reaction (PCR) DNA amplifications were conducted using DNA Engine Cycler Tetrad2 (BioRad, Hercules, California) in 6-μL volumes consisting of
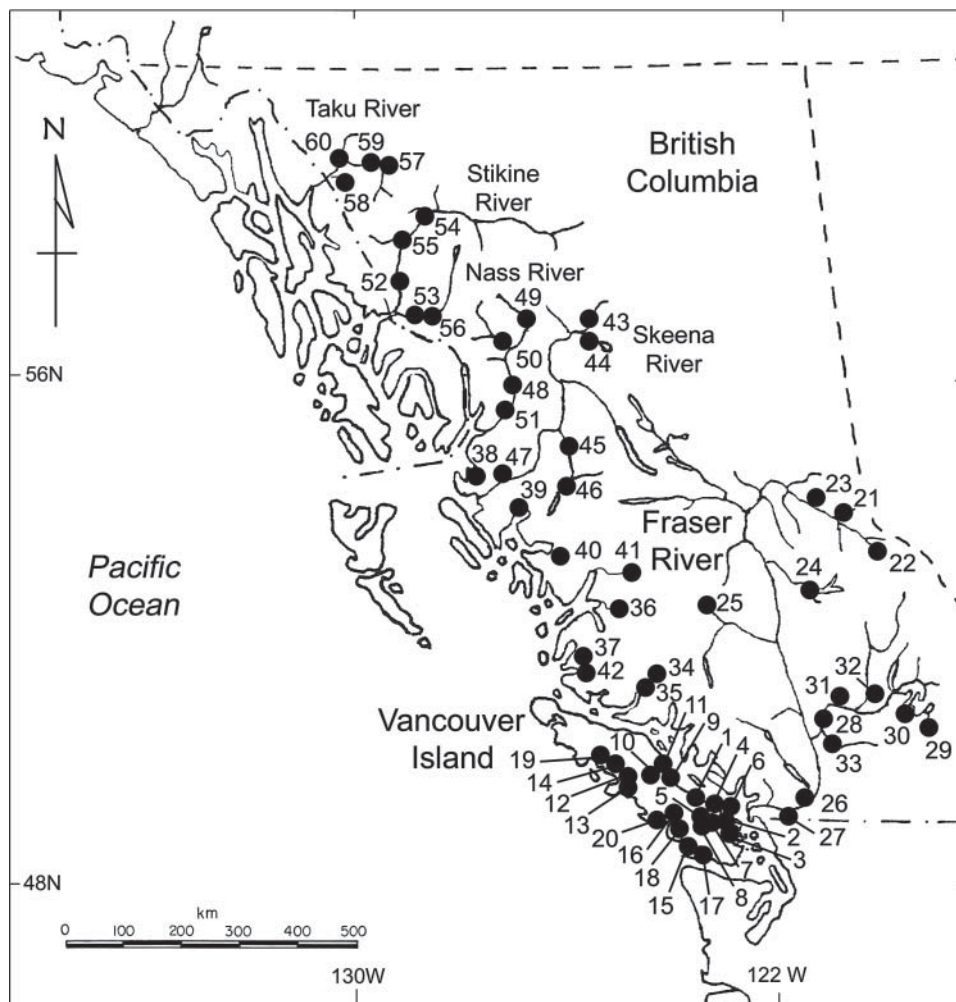

FIGURE 1. Map indicating the locations of the 60 Chinook salmon populations surveyed for the study. Population and region names are given in Table 1.

TABLE 1. Regions, populations within regions, and sample sizes available in the survey of single-nucleotide polymorphisms (SNPs) and microsatellites for 60 Chinook salmon populations in British Columbia. The locations of the populations are shown in Figure 1. Region codes are as follows: East Coast Vancouver Island (ECVI), West Coast Vancouver Island (WCVI), upper Fraser River (UPFR), middle Fraser River (MUFR), lower Fraser River (LWFR), South Thompson River (SOTH), lower Thompson River (LWTH), southern British Columbia mainland (SOMN), and northern British Columbia mainland (NOMN).

Region	Population	SNPs	Microsatellites
ECVI	1. Big Qualicum	95	200
	2. Chemainus	95	176
	3. Cowichan	95	200
	4. Little Qualicum	190	166
	5. Nanaimo, fall	97	200
	6. Nanaimo, spring	95	89
	7. Nanaimo, summer	185	184
	8. Nanaimo, upper	95	109
	9. Puntledge, summer	95	200
	10. Puntledge, fall	108	200
	11. Quinsam	95	200
WCVI	12. Gold	95	151
	13. Burman	95	200
	14. Conuma	95	200
	15. Nitinat	95	173
	16. Robertson	95	169
	17. San Juan	95	160
	18. Sarita	99	195
	19. Tahsis	95	199
	20. Toquart	88	74
UPFR	21. Morkill	95	150
	22. Swift	95	192
	23. Torpy	95	141
MUFR	24. Quesnel	95	200
	25. Upper Chilcotin	95	193
LWFR	26. Harrison	95	200
	27. Chilliwack	95	200
SOTH	28. Lower Thompson	95	157
	29. Middle Shuswap	95	196
	30. Salmon	95	129
LWTH	31. Deadman	95	200
	32. Louis	95	200
	33. Nicola	95	179
SOMN	34. Devereux	95	167
	35. Klinaklini	95	200
NOMN	36. Atnarko	95	164
	37. Chuckwalla	95	125
	38. Kateen	95	123
	39. Kitimat	105	200
	40. Kitlope	95	170
	41. Upper Dean	95	163
	42. Wannock	95	200
Upper Skeena River	43. Bear	80	110
	44. Sustut	95	200
Skeena River, Bulkley	45. Bulkley	95	200
	46. Morice	95	132
Lower Skeena River	47. Kitsumkalum	95	200

TABLE 1. Continued.

Region	Population	SNPs	Microsatellites
Nass River	48. Cranberry	79	122
	49. Damdochax	95	105
	50. Kwinageese	95	139
	51. Tseax	95	171
Stikine River	52. Christina	95	177
	53. Craig	95	99
	54. Little Tahltan	95	200
	55. Shakes Creek	95	167
	56. Verrett	94	200
Taku River	57. Dudidontu	95	189
	58. Little Tatsamenie	144	200
	59. Nahlin	95	200
	60. Nakina	95	200
Mean		99	172

0.15 units of *Taq* polymerase, 1 μL of extracted DNA, 1 × PCR buffer (Qiagen, Mississauga, Ontario), 60 μM of each nucleotide, 0.40 μM of each primer, and deionized H_2O. The thermal cycling profile involved one cycle of 15 min at 95°C followed by 30–40 cycles of 20 s at 94°C, 30–60 s at 47–65°C and 30–60 s at 68–72°C (depending on the locus). The PCR conditions for particular loci could vary from this general outline. The PCR fragments were initially size-fractionated in denaturing polyacrylamide gels using an ABI 377 automated DNA sequencer, and genotypes were scored by Genotyper 2.5 software (Applied Biosystems, Foster City, California) using an internal lane sizing standard. Later in the study, microsatellites were size-fractionated in an ABI 3730 capillary DNA sequencer, and genotypes were scored by GeneMapper software 3.0 (Applied Biosystems) using an internal lane sizing standard. Allele identification between the two sequencers was standardized by analyzing approximately 600 individuals on both platforms and converting the sizing in the gel-based data set so that it would match that obtained from the capillary-based set for the 12 original DFO loci and 13 GAPS loci. The repeatability of genotyping was evaluated by repeat PCR analysis and scoring of the genotypes of individual fish. Discrepancies in scoring were observed in 8 of the 15,360 genotypes scored across all loci, for a genotyping error rate of 0.05%.

Variation was analyzed at 72 nuclear and 1 mitochondrial SNP with the primer and probe sequences outlined by Smith et al. (2005a, 2005b, 2005c, 2007), Campbell and Narum (2008), and Miller et al. (2008). A listing of all the SNPs surveyed is given in Table 2. After PCR amplification, the plates (384 well) were read on an ABI Prism 7900HT Sequence Detection System by one individual using sequence detection software from ABI. One SNP (*Ots_CI_A1*) required analysis on the automated sequencer, as it was an insertion or deletion and as such a size variant. The repeatability of genotyping was evaluated by repeat

PCR analysis and scoring of genotypes. Discrepancies in scoring were observed in 3 of the 21,408 genotypes scored across all loci, for a genotyping error rate of 0.01%.

Estimating stock composition in single-population samples.—Two software packages were utilized in estimating the stock composition of single-population mixtures: the Statistical Package for the Analysis of Mixtures (SPAM version 3.7; Debevec et al. 2000) and ONCOR (Kalinowski et al. 2007). A mitochondrial DNA SNP (mtSNP) was analyzed in the survey (*Ots_C3N3*), and as ONCOR is currently unable to analyze variations incorporating mitochondrial haplotypes, SPAM was used exclusively for those analyses. Genotypic frequencies were determined for each locus in each population and were used to estimate the stock composition of simulated single-population samples. The Rannala and Mountain (1997) correction to baseline allele frequencies was used in SPAM analyses to avoid the occurrence of fish in the mixed sample from a specific population with an allele not observed in the baseline samples from that population. This correction incorporated Bayesian modeling of baseline allele frequency distributions. All loci were considered to be in Hardy–Weinberg equilibrium (Beacham et al. 2006b); the expected genotypic frequencies were determined from the observed allele frequencies. The reported stock compositions for the simulated single-population samples are the bootstrap mean estimates for each mixture of 200 fish analyzed, where the mean and variance estimates are derived from 1,000 bootstrap simulations. Each baseline population and simulated single-population sample was sampled with replacement in order to simulate the random variation involved in the collection of the baseline and fishery samples. When ONCOR was used to estimate stock composition, the Rannala and Mountain (1997) correction to baseline allele frequencies was again implemented, the precision of the stock compositions being calculated by bootstrapping (100 simulations) over the observed baseline population sample sizes

TABLE 2. Ranking of 102 markers (29 microsatellites and 73 single-nucleotide polymorphisms [SNPs]) in terms of the average estimated composition (population and regional accuracy determined with SPAM) of single-population samples over 60 populations of Chinook salmon, as well as the number of alleles observed at the locus and F_{st} value. Types are as follows: M = microsatellites, S = SNPs, and SM = mtDNA SNPs. H_e is expected heterozygosity.

Marker	Type	Alleles	F_{st}	H_e	Population (%)	SD (%)	Region (%)	SD (%)
1. Ots107	M	47	0.048	0.91	81.1	7.6	91.2	4.1
2. Ots201b	M	52	0.048	0.90	80.8	7.7	89.6	4.8
3. Ots213	M	55	0.037	0.93	80.1	7.2	88.6	4.6
4. Omm1080	M	62	0.021	0.95	79.3	7.1	86.9	5.3
5. Ots212	M	34	0.049	0.87	79.3	8.7	88.0	5.9
6. Ots100	M	54	0.019	0.94	78.5	7.6	85.2	6.1
7. Ots211	M	37	0.042	0.92	77.5	8.2	86.9	5.3
8. Ots208b	M	54	0.029	0.93	77.5	7.8	86.4	5.5
9. Omy325	M	43	0.110	0.76	77.3	11.7	91.3	5.1
10. Oki100	M	42	0.028	0.93	77.2	8.0	85.1	5.9
11. Ssa197	M	45	0.028	0.93	76.8	8.2	85.5	5.7
12. Oki10	M	57	0.032	0.94	76.3	8.0	86.3	5.3
13. Ots101	M	48	0.031	0.92	75.9	8.5	85.2	5.9
14. Ssa408	M	31	0.052	0.88	75.9	9.8	87.7	5.6
15. Ots104	M	41	0.028	0.93	75.6	8.5	84.5	5.9
16. Ogo4	M	20	0.102	0.78	75.1	12.2	88.4	6.2
17. OtsG68	M	49	0.021	0.93	72.7	8.8	83.0	6.3
18. Ots2	M	26	0.077	0.68	72.2	14.4	84.9	9.4
19. OmyRGT3TUF	M	22	0.074	0.79	71.6	12.1	84.3	7.4
20. Ogo2	M	26	0.059	0.73	70.9	14.0	83.0	9.4
21. Omm1276	M	38	0.049	0.87	70.2	10.8	81.7	7.2
22. Ots3M	M	14	0.101	0.73	69.1	15.5	84.1	9.1
23. OtsG474	M	16	0.141	0.51	59.6	21.0	75.6	15.2
24. Oke4	M	14	0.097	0.65	51.4	21.2	69.0	15.7
25. Bhms417	M	9	0.092	0.72	45.4	22.1	65.3	16.1
26. Omm1009	M	11	0.093	0.67	44.6	22.7	65.0	18.4
27. Ots9	M	12	0.062	0.53	41.3	23.7	58.2	20.1
28. Omm1037	M	10	0.091	0.57	41.1	24.5	62.3	18.1
29. OmyRGT30TUF	M	7	0.124	0.49	34.2	22.2	56.5	17.8
30. Ots_FARSLA-220	S	2	0.373	0.32	13.9	11.2	39.7	13.2
31. Ots_MHC2	S	2	0.455	0.18	13.3	9.3	32.1	12.1
32. Ots_RFC2-558	S	2	0.311	0.31	12.4	10.2	34.4	11.0
33. Ots_u211-85	S	2	0.235	0.37	11.9	9.6	29.2	11.6
34. Ots_hnRNPL-533	S	2	0.247	0.38	11.8	9.4	35.6	12.0
35. Ots_U07-25-325	S	2	0.206	0.25	11.2	7.6	28.0	10.0
36. Ots_NRAMP-321	S	2	0.237	0.36	11.1	9.3	31.1	12.2
37. Ots_PGK-54	S	2	0.175	0.31	11.0	7.6	25.0	10.4
38. Ots_RAG3	S	2	0.314	0.35	10.7	9.7	32.2	14.3
39. Ots_P450	S	2	0.266	0.31	10.6	8.7	35.9	13.0
40. Ots_HSP90B-100	S	2	0.192	0.40	10.6	8.8	27.9	12.1
41. Ots_SL	S	2	0.167	0.41	10.5	8.4	30.5	10.3
42. Ots_S7-1	S	2	0.145	0.43	10.5	8.3	26.7	11.1
43. Ots_CI-GCSH-R-R	S	2	0.214	0.35	10.5	8.4	28.3	10.6
44. Ots_MHC1	S	2	0.172	0.41	10.3	8.0	29.7	11.1
45. Ots_CI-THIO-1	S	2	0.108	0.44	10.0	7.0	28.7	8.3
46. Ots_P53	S	2	0.144	0.42	10.0	8.4	28.3	10.2
47. Ots_u202-161	S	2	0.191	0.39	9.6	8.5	25.2	10.6
48. Ots_NOD1	S	2	0.190	0.31	9.6	8.6	27.2	10.3
49. Ots_Prl2	S	2	0.111	0.42	9.6	6.7	26.5	10.9
50. Ots_u07-57-120	S	2	0.175	0.35	9.6	7.9	25.8	11.7

TABLE 2. Continued.

Marker	Type	Alleles	F_{st}	H_e	Population (%)	SD (%)	Region (%)	SD (%)
51. *Ots_CI-PEMT-Y*	S	2	0.116	0.43	9.6	7.3	22.5	10.1
52. *Ots_IL8R-CS*	S	2	0.139	0.36	9.5	7.9	29.1	9.7
53. *Ots_IL-11*	S	2	0.157	0.42	9.4	8.5	30.9	10.9
54. *Ots_NKEF-192*	S	2	0.150	0.26	9.3	8.0	31.7	10.8
55. *Ots_ETIF1*	S	2	0.167	0.41	9.1	8.1	25.1	11.0
56. *Ots_U07-53-133*	S	2	0.094	0.37	9.0	7.3	24.6	9.5
57. *Ots_ASPAT-196*	S	2	0.193	0.14	8.9	7.5	28.0	10.9
58. *Ots_TGFb1*	S	2	0.091	0.42	8.8	7.0	25.0	9.0
59. *Ots_u4-92*	S	2	0.196	0.17	8.7	8.3	31.3	10.6
60. *Ots_CI-CIRPA-2-2*	S	2	0.127	0.43	8.7	7.3	29.0	10.4
61. *Ots_MYOD-364*	S	2	0.140	0.34	8.7	8.0	25.8	11.4
62. *Ots_TLR3*	S	2	0.079	0.43	8.7	6.6	23.1	8.3
63. *Ots_FGF6A*	S	2	0.094	0.28	8.6	6.3	25.7	9.9
64. *Ots_GTH2B-550*	S	2	0.103	0.40	8.5	7.0	28.0	12.3
65. *Ots_U07-18-378*	S	2	0.016	0.23	8.5	8.2	28.2	12.5
66. *Ots_HSP90B-385*	S	2	0.096	0.30	8.5	6.8	24.3	9.6
67. *Ots_SClkF2R2-135*	S	2	0.108	0.43	8.4	7.7	22.7	10.7
68. *Ots_TAPBP*	S	2	0.088	0.22	8.4	6.6	19.5	8.1
69. *Ots_Tnsf*	S	2	0.135	0.44	8.3	7.4	25.9	11.3
70. *Ots_GPH-318*	S	2	0.129	0.19	8.3	6.8	30.7	11.5
71. *Ots_TCL1*	S	2	0.087	0.45	8.1	6.3	20.7	9.4
72. *Ots_MYO1-384*	S	2	0.134	0.15	8.1	7.0	23.3	8.0
73. *Ots_U07-49-290*	S	2	0.094	0.33	8.1	7.0	25.2	9.2
74. *Ots_E2-275*	S	2	0.095	0.46	8.0	6.7	20.2	10.2
75. *Ots_CI-OSTM1-1*	S	2	0.086	0.45	7.9	6.9	21.4	10.5
76. *Ots_U07-07-161*	S	2	0.066	0.40	7.8	6.5	21.1	10.4
77. *Ots_CD63*	S	2	0.079	0.45	7.4	6.1	20.9	9.2
78. *Ots_cox1-241*	S	2	0.078	0.43	7.4	7.0	25.5	9.4
79. *Ots_SWS1op-182*	S	2	0.068	0.47	7.3	6.1	21.9	8.4
80. *Ots_IGF-I.1-76*	S	2	0.157	0.19	7.2	7.5	25.8	12.5
81. *Ots_GPDH-338*	S	2	0.110	0.12	7.2	5.5	17.9	8.6
82. *Ots_CI-PHOS-R*	S	2	0.102	0.25	6.9	7.3	20.7	11.1
83. *Ots_u6-75*	S	2	0.091	0.18	6.9	6.7	20.3	10.0
84. *Ots_CI-1363-W*	S	2	0.083	0.32	6.9	6.9	18.5	9.7
85. *Ots_U212-158*	S	2	0.052	0.17	6.9	5.6	17.3	8.3
86. *Ots_CI_A1*	S	2	0.053	0.45	6.8	5.8	16.7	8.6
87. *Ots_CI-1740-R*	S	2	0.050	0.47	6.8	5.4	21.0	8.6
88. *Ots_C3N3*	SM	2			6.8	1.2	20.6	2.1
89. *Ots_AsnRS-60*	S	2	0.045	0.37	6.7	6.2	19.3	9.8
90. *Ots_unkn526*	S	2	0.071	0.21	6.7	6.0	21.2	10.2
91. *Ots_U07-17-373*	S	2	0.078	0.09	6.6	6.1	22.2	7.9
92. *Ots_LWSop-638*	S	2	0.117	0.09	6.5	6.5	22.3	8.7
93. *Ots_GnRH-271*	S	2	0.073	0.14	6.4	6.4	19.6	10.5
94. *Ots_CI-1803-K*	S	2	0.056	0.33	6.1	6.1	15.3	8.7
95. *Ots_CI-ISOT-Y*	S	2	0.064	0.09	5.7	6.0	20.2	10.3
96. *Ots_OTS311-101X*	S	2	0.034	0.15	5.7	5.6	18.1	8.8
97. *Ots_ZNF330-181*	S	2	0.189	0.02	5.3	3.5	16.2	7.2
98. *Ots_arf-188*	S	2	0.057	0.02	4.8	4.5	16.1	7.9
99. *Ots_Ikaros-250*	S	2	0.078	0.01	4.3	4.2	14.1	7.1
100. *Ots_HGFA-446*	S	2	0.014	0.01	2.9	3.8	13.1	7.5
101. *Ots_E9-BAC*	S	2	0.006	0.00	2.4	3.0	11.0	5.7
102. *Ots_CYP17*	S	2	0.006	0.00	2.4	3.4	10.8	6.2

and a mixture size of 200 fish. For both SPAM and ONCOR, the allocations to individual baseline populations were summed to provide estimates of stock composition for regional stock groups (Table 1). Additionally, ONCOR was used to provide estimates of the accuracy of identification of individuals to specific populations or regional stock groups through leave-one-out assignment testing.

The sample sizes for the microsatellites were variable among populations and loci (Table 1). To control for the effect of varying sample size, construction of the microsatellite baseline for the analysis proceeded on the basis of capping population sample size at 200 individuals. The population sample size for the microsatellite analysis ranged from 74 to 200 individuals. The population sample size for the SNP analysis ranged from 79 to 190 individuals but was typically set at 95 individuals. Allele frequencies for all of the populations surveyed in this study are available at the Molecular Genetics Laboratory's Web site (http://www.pac.dfo-mpo.gc.ca/science/facilities-installations/pbs-sbp/mgl-lgm/data-donnees/index-eng.htm).

Relative ranking of the loci.—The power of individual loci for stock composition estimation was initially evaluated by incorporating only a single locus in the estimation of the stock composition of simulated single-population samples. As an mt-SNP was included in the analysis, only SPAM was used to provide estimates of stock composition for all 60 single-population samples. Mean accuracy was determined as the average estimate across all 60 populations, with loci then being ordered from the most accurate to the least accurate relative to population-specific accuracy (Table 2). Heterozygosity for each locus over all populations was calculated with FSTAT version 2.9.3.2 (Goudet 1995).

How many microsatellites for SNP equivalency?—ONCOR was used exclusively for this analysis, with the proviso that the mtSNP was eliminated from the analysis. The microsatellite locus with the highest average population accuracy was initially incorporated into the analysis; lower-accuracy microsatellites were then added sequentially until the average population accuracy and precision of the stock composition estimates provided by the suite of microsatellites matched that provided by the SNPs. Single-population samples were analyzed with the best 3–14 microsatellites. Accuracy and standard deviation were determined for the population- and region-specific estimates of stock composition and then averaged over the 60 single-population simulations for each set of loci. The percent correct assignment of individuals to specific populations and regions was also determined for all 60 populations and then averaged over all populations for each set of microsatellites evaluated.

Comparing DFO and GAPS microsatellites.—For each run (injection) of the ABI 3730 sequencer, analysis of microsatellite variation can accommodate loci marked with up to four different tags in addition to the size standard necessary for estimating allele size. As the number of injections increases, the cost of the analysis rises, and thus efficient laboratory operation requires that the number of injections be minimized in surveys of microsatellite variation. The DFO suite of microsatellites

TABLE 3. Numbers of injections on the automated sequencer, primer tags employed, and loci surveyed for the Fisheries and Oceans Canada (DFO) and Genetic Analysis of Pacific Salmon (GAPS) surveys of microsatellite variation.

Injections	Tag	Loci
DFO		
1	VIC	*Ssa197*
	6FAM	*Oki100*, *Omy325*
	NED	*Ots2*, *Ots9*, *Oke4*
	PET	*Ogo2*, *Ogo4*
2	VIC	*Ots101*
	6FAM	*Ots100*
	NED	*Ots104*
	PET	*Ots107*
GAPS		
1	VIC	*Oki100*
	6FAM	*Omm1080*
	PET	*Ogo2*, *Ogo4*
2	VIC	*Ots213*
	6FAM	*Ots201b*
	NED	*Ots9*, *Ssa408*
3	VIC	*Ots208b*
	6FAM	*Ots212*
	NED	*Ots3*, *Ots211*
	PET	*OtsG474*

is comprised of 12 loci surveyed with two injections on the automated DNA sequencer, and the GAPs suite is comprised of 13 loci surveyed with three injections (two dyes unutilized) (Table 3). If additional accuracy is required for either set of loci, either microsatellites or SNPs can be added to the existing baselines to provide increased stock identification power. If microsatellites are added to either suite and the number of injections on the automated sequencer is capped at three, four microsatellites can be added to the DFO suite (4 dyes are available for each injection) and two microsatellites can be added to the GAPS suite (2 unutilized dyes). The microsatellites added to the DFO suite would include the top-ranked non-DFO microsatellites *Ots201b*, *Ots213*, *Omm1080*, and *Ots212*. Similarly, the microsatellites added to the GAPS suite would include *Ots107* and *Ots100*. Single-population samples were analyzed with the DFO (regular and enhanced suites) and GAPS microsatellites (regular and enhanced suites). Additionally, the original suites of DFO and GAPS microsatellites were enhanced with the addition of SNPs to the baselines used for stock composition analysis, the number of SNPs being added in increments of five, starting with the highest-rated ones. The addition of groups of SNPs continued until the average estimated population accuracy derived from the enhanced DFO and GAPS microsatellite baselines was achieved. Accuracy and standard deviation for population- and region-specific estimates of stock composition were determined for each group of genetic markers examined and then averaged over the 60 single-population simulations. The percent correct

assignment of individuals to specific populations and regions was also determined for all 60 populations and then averaged over all populations for each set of microsatellites evaluated.

How many SNPs for microsatellite equivalency?— Projections of the number of SNPs required for equivalency with the current microsatellite baseline were made by ranking the SNPs according to the average accuracy observed in estimating stock composition for single-population samples over the 60 populations surveyed. The mtSNP was eliminated in these analyses. Subsequent analyses were conducted exclusively with ONCOR. Single-population samples were analyzed with 10–72 SNPs in increments of 5 SNPs. The SNPs with the highest average accuracy were initially incorporated into the analyses of the single-population mixtures, with the less accurate SNPs being added sequentially. Additionally, the SNPs with the lowest average accuracy values were initially incorporated into the analyses, with progressively more accurate SNPs being added sequentially, again with average accuracy and precision being recorded. The overall mean accuracy and precision of each specified number of SNPs were determined by averaging the results from both processes; this was considered indicative of the average trend in estimating accuracy and precision when the number of SNPs employed in the analysis was increased. The average regional, population, and individual accuracy and precision over all 60 populations were recorded for each set of SNPs. A hyperbola function of the form $Y = a/X + b$ was fitted with Labfit curve-fitting software (Pereira da Silva and Pereira da Silva 2007), in which Y is the mean observed accuracy for the population and regional estimates and X is the number of SNPs incorporated into the analysis. Estimates of standard deviations were derived from the power function $Y = aX^b$, where Y and X are defined as before. Individual assignment accuracy for populations and regions was fitted with the modified geometric function $Y = aX^{b/X}$. Projections were then made with these regression models to estimate the number of SNPs required to provide estimates of comparable resolution to that provided by the microsatellites with respect to estimated stock composition at both the regional and population levels as well as individual identification at the regional and population levels.

RESULTS

Relative Rankings of the Markers

The number of alleles observed at a microsatellite locus was important in determining the value of the locus for stock identification applications. The number of alleles observed at a locus varied from 7 to 62 for the populations and loci surveyed in our study (Table 2). The number of alleles observed at a microsatellite locus was related to the accuracy of the estimated stock composition of the simulated single-population samples (Figure 2a). As the number of alleles increased up to about 30 per locus, the accuracy of population-specific estimated stock composition increased. However, larger numbers of alleles provided only a minor increase in the accuracy of es-
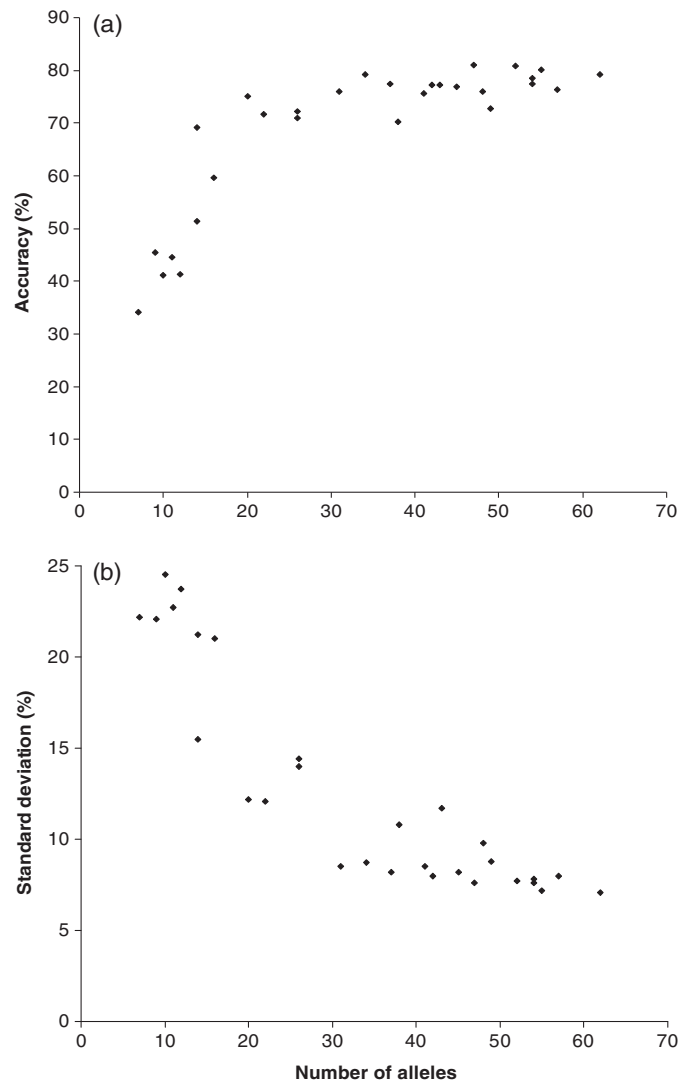

FIGURE 2. Relationships between the number of alleles observed at a microsatellite locus and **(a)** the average percentage accuracy of stock composition and **(b)** the standard deviation obtained for simulated single-population samples using only a single locus and the 60-population baseline of British Columbia Chinook salmon.

timated stock composition. The accuracy of the mean estimated stock compositions of the single-population samples (correct = 100%) were 40.2% for simulations with single loci having 10 or fewer alleles, 56.9% for loci with 11–20 alleles, 71.6% for loci with 21–30 alleles, 75.7% for loci with 31–40 alleles, 76.7% for loci with 41–50 alleles, 78.6% for loci with 51–60 alleles, and 79.3% for a locus with more than 60 alleles (Table 2). The precision of the estimated stock composition for a locus was influenced by the number of alleles observed at a microsatellite locus, with more precise estimates being derived from loci with larger numbers of alleles (Figure 2b). The mean standard deviations of the estimated stock compositions were 22.9% for loci with 10 or fewer alleles, 19.4% for loci with 11–20 alleles, 13.5% for loci with 21–30 alleles, 9.4% for loci with 31–40

alleles, 8.8% for loci with 41–50 alleles, 7.7% for loci with 51–60 alleles, and 7.1% for a locus with more than 60 alleles (Table 2). In general, microsatellites with more alleles provided more accurate and precise estimates of the stock compositions of the single-population samples than did loci with fewer alleles, but the results varied depending on the specific loci.

Average population-specific accuracy for the 102 markers evaluated ranged from 2.4% to 81.1%, with a clear break between SNPs and microsatellites. The average accuracy in the stock identification analysis ranged from 2.4% to 13.9% for individual SNPs and from 34.2% to 81.1% for individual microsatellites. The top 29 of the 102 markers evaluated were microsatellites, with the accuracy of the estimated population-specific stock compositions produced by incorporating the lowest-ranked microsatellite (*OmyRGT30TUF*, which has seven alleles) being 2.5 times as high as that of the highest-ranked SNP (*Ots_FARSLA-220*) (34.2% versus 13.9%; Table 2). The differential between these two loci was less for region-specific accuracy (56.5% versus 39.7%).

Heterozygosity was a good predictor of the relative power of microsatellites for stock composition analysis ($r = 0.85$, $P < 0.01$). For microsatellites, loci with fewer alleles also tended to have lower heterozygosity (Table 3), and just as loci with more alleles were more powerful for stock identification analysis, more heterozygous loci were more powerful for stock identification analysis (Figure 3a). However, there was little increase in the accuracy of estimated stock composition for microsatellites once heterozygosity values reached 0.80, which corresponded to approximately 30 alleles. All of the SNPs evaluated displayed only two alleles, and there were only minor differences in stock identification accuracy among SNPs for loci with a heterozygosity value greater than 0.15 (Figure 3b). In this case, heterozygosity at SNP loci was a modest predictor of the relative power of the locus for stock identification analysis ($r = 0.55$, $P < 0.01$).

How Many Microsatellites for SNP Equivalency?

The application of 72 SNPs in estimating the stock composition of 60 single-population samples resulted in an average accuracy of 97.5% to reporting region, with a standard deviation of 0.8% (Table 4). These levels of accuracy and precision, along with the levels associated with the population-level estimates of stock composition and individual assignment to both region and population provided the reference points for estimating the number of microsatellites required to provide equivalent results in stock identification applications. The regional accuracy estimates for individual populations ranged from 79% for the Louis Creek population to 100% for the middle Shuswap River population. Starting with the microsatellite with the highest observed population-specific accuracy (*Ots107*), at least 14 microsatellites were required to achieve comparable levels of accuracy (97.3%) and precision (SD = 1.0%) in the regional estimates of stock composition (Table 4).
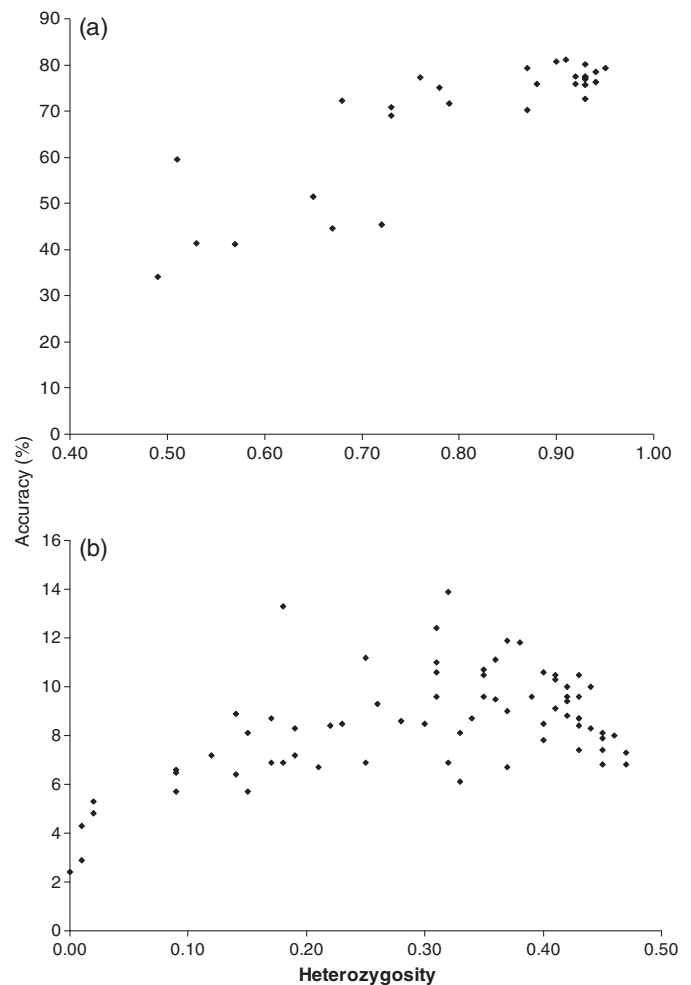

FIGURE 3. Mean accuracy of the stock composition estimation of single-population samples vis-à-vis the heterozygosity of **(a)** microsatellites and **(b)** single-nucleotide polymorphisms for 60 Chinook salmon populations in British Columbia.

The application of 72 SNPs in estimating stock composition resulted in an average population-specific accuracy of 83.1% (SD = 3.7%; Table 4). Individual population estimates ranged from 6.3% for the Puntledge River fall-run population to 99.9% for the middle Shuswap River population (Figure 4). Some populations, such as those of the Quinsam, upper Chilcotin, and Bulkley rivers, were clearly more differentiated than others, with high levels of population-specific accuracy (>97%) regardless of the suite of markers used to estimate stock composition. Similarly, the estimated stock compositions for the populations that displayed lower levels of accuracy, such as those of Shakes Creek, the Nahlin River, and the Nakina River, were consistently lower than than those for most other populations with both microsatellites and SNPs. The standard deviations of the estimated stock compositions were generally higher with SNPs than with either of the DFO or GAPS suites of microsatellites (Figure A.1 in the appendix). The standard deviations of the estimated stock compositions derived from the SNPs were greater than those

TABLE 4. Accuracy (means) and precision (SDs) of the estimated population and regional stock compositions of single-population samples determined with ONCOR from suites of markers incorporating the best 3–14 microsatellites from Table 2 as well as the best 72 single-nucleotide polymorphisms (SNPs). The percent correct assignment of individuals to specific populations and regions is also shown.

Loci	Stock composition (%)		SD (%)		Individual assignment (%)	
	Population	Region	Population	Region	Population	Region
Microsatellites						
3	75.8	92.9	4.7	2.3	35.6	62.2
4	76.3	93.2	4.2	2.1	42.0	68.0
5	77.5	94.1	3.9	1.8	46.9	73.7
6	80.3	94.3	3.4	1.7	52.0	76.3
7	81.5	94.9	3.3	1.6	55.9	79.7
8	82.2	95.3	3.1	1.4	58.2	81.9
9	83.4	96.2	2.8	1.2	60.7	84.5
10	83.9	96.4	2.7	1.1	62.2	85.6
11	83.3	96.2	2.7	1.1	64.3	87.0
12	83.8	96.5	2.6	1.0	64.5	87.5
13	84.8	97.1	2.5	1.0	65.3	88.1
14	84.9	97.3	2.5	1.0	67.9	89.3
SNPs	83.1	97.5	3.7	0.8	56.7	87.3

of the compositions derived from the DFO microsatellites for 39 of the populations evaluated (sign test analysis; $P < 0.05$) and greater than those of the compositions from the GAPS microsatellites for 43 of the populations ($P < 0.01$). On average, application of the best microsatellites resulted in comparable levels of accuracy (83.4%) and precision (SD = 3.4%) with nine and six microsatellites, respectively.

Assignment of individuals to the correct region of origin with the 72 SNPs was attained with an accuracy of 87.3%, and assignment to specific populations was attained with an accuracy of 56.7% (Table 4). Regional assignment accuracy ranged from 15% for the Bear River population to 100.0% for several populations, such as the Burman River population. Population assignment accuracy ranged from 8.7% for the Shakes Creek population to 98.6% for the middle Shuswap River population (Figure A.2). Comparable region-specific levels of assignment accuracy (87.5%) were achieved with the use of 12 microsatellites, comparable levels of population-specific assignment accuracy (58.2%) with 8 microsatellites (Table 4). In summary, the number of microsatellites required to produce stock identification results equivalent to those from applying 72 SNPs depended on the specific task. Regional estimates of stock composition required at least 14 microsatellites, population-specific estimates of stock composition required 6–9 microsatellites, and individual assignment to population and region required 8–12 microsatellites.

Comparing DFO and GAPS Microsatellites

The average estimates of population-specific stock composition for single-population samples were 83.0% for the DFO microsatellites and 83.1% for the GAPS microsatellites

(Table 5). Both suites of loci provided essentially equal average population-specific accuracy over the 60 populations of Chinook salmon surveyed in British Columbia. The addition of the four most powerful non-DFO microsatellites (*Ots201b*, *Ots213*, *Omm1080*, and *Ots212*) to the suite of DFO microsatellites improved the estimated accuracy of the single-population samples to 85.6% (Table 5). Improvement of the level of accuracy of estimated stock compositions to this level could also be achieved by the addition of about 25 of the top-ranked SNPs to the 12 DFO microsatellites (Table 5). The addition of the 2 most powerful non-GAPS microsatellites (*Ots107* and *Ots100*) to the GAPS suite improved the estimated accuracy to an average of 85.5% (Table 5), essentially the same as that of the enhanced DFO suite. Approximately 20–25 of the top-ranked SNPs would be required to be added to the suite of GAPS microsatellites to provide the same population resolution as that provided by the addition of the 2 most powerful non-GAPS microsatellites. In summary, at least two options are available for improving the accuracy and precision of stock composition estimates for the DFO and GAPS suites of microsatellites if the number of injections on the automated sequencer is capped at three. For the DFO microsatellites, either 4 microsatellites or 25 SNPs can be added to the existing suite. For the GAPS microsatellites, either 2 microsatellites or 20–25 SNPs can be added to the existing suite. The addition of either set of loci to both sets of microsatellites would provide generally equivalent results.

In stock identification applications, different levels of resolution may be required for populations in different reporting regions. For example, population-specific estimates of stock composition may be required in a specific region, while regional-level estimates are satisfactory in other regions. If an existing
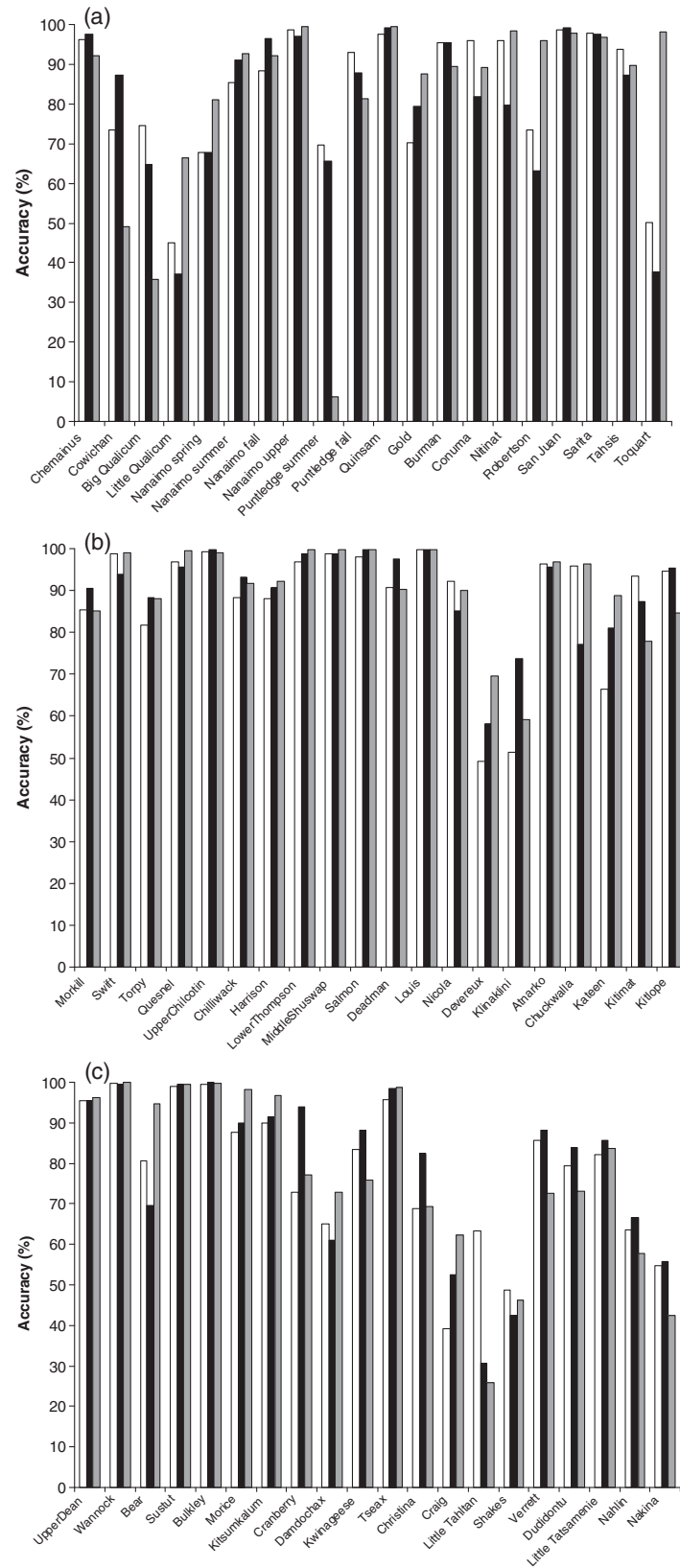
FIGURE 4.   Mean accuracy of assignment to population derived from the DFO microsatellite (white bars), GAPS microsatellite (black bars), and single-nucleotide polymorphism (gray bars) suites of loci for **(a)** Vancouver Island, **(b)** Fraser River, southern mainland, and northern mainland, and **(c)** additional northern mainland, Skeena River, Nass River, Stikine River, and Taku River Chinook salmon populations for simulated single-population samples.

TABLE 5. Accuracy (means) and precision (SDs) of the estimated population and regional stock compositions of single-population samples determined with ONCOR from suites of markers incorporating the best 13 microsatellites (micros), the best 72 single-nucleotide polymorphisms (SNPs), the Fisheries and Oceans Canada (DFO) microsatellites, the DFO microsatellites plus the best 20–25 SNPs, the DFO microsatellites plus 4 additional microsatellites (*Ots201b*, *Ots213*, *Omm1080*, and *Ots212*), the Genetic Analysis of Pacific Salmon (GAPS) microsatellites, the GAPS microsatellites plus the best 20–25 SNPs, and the GAPS microsatellites plus 2 additional microsatellites (*Ots107* and *Ots100*) over 60 Chinook salmon populations in British Columbia. The percent correct assignment of individuals to specific populations and regions is also shown.

	Stock composition (%)		Standard deviation (%)		Individual assignment (%)	
Loci	Population	Region	Population	Region	Population	Region
Optimum 13	84.8	97.1	2.5	1.0	65.3	88.1
72 SNPs	83.1	97.5	3.7	0.8	56.7	87.3
DFO	83.0	96.9	3.1	1.1	61.5	87.0
DFO + 20 SNPs	84.8	97.7	2.7	0.8	66.5	91.7
DFO + 4 micros	85.6	97.5	2.4	0.9	69.1	91.0
DFO + 25 SNPs	85.7	97.9	2.6	0.8	67.8	92.2
GAPS	83.1	96.7	2.8	1.0	61.8	87.3
GAPS + 20 SNPs	84.9	97.6	2.6	0.8	67.1	91.4
GAPS + 2 micros	85.5	97.1	2.5	0.9	66.9	89.2
GAPS + 25 SNPs	86.3	97.8	2.5	0.8	68.3	91.6

microsatellite baseline is to be enhanced to improve accuracy, the choice of loci to add may depend to some degree on the specific application. The levels of accuracy of estimated stock composition at the population level clearly differed among the reporting regions. For example, in applications utilizing the DFO loci and centered on providing higher resolution of East Coast Vancouver Island populations (a region of conservation concern), enhancing the DFO baseline with four additional microsatellites provided the most effective means of increasing the resolution among populations (Table 6).

Alternatively, enhancing the accuracy of identification of Stikine River populations was most effective by adding 25 SNPs. For the GAPS microsatellite baseline, enhancement with two microsatellites provided the best resolution of upper Fraser River populations, but for West Coast of Vancouver Island populations a combined microsatellite–SNP approach provided the best population resolution (Table 6).

In some applications, individual Chinook salmon are required to be identified to either their region or population of origin. The regional origin of individuals from lower Thompson River

TABLE 6. Average accuracy of estimated population stock compositions by region for the Fisheries and Oceans Canada (DFO), Genetic Analysis of Pacific Salmon (GAPS), single-nucleotide polymorphism (SNP), DFO plus 4 microsatellites (micros), GAPS plus 2 microsatellites, DFO plus 25 SNPs, and GAPS plus 25 SNPs baselines incorporating the regions and populations indicated in Table 1. See Table 5 for more information about the loci utilized.

Region	DFO	GAPS	SNPs	DFO + 4 micros	GAPS + 2 micros	DFO + 25 SNPs	GAPS + 25 SNPs
ECVI	80.6	81.1	72.4	83.5	82.0	81.9	81.7
WCVI	85.7	80.1	93.6	87.1	81.1	91.9	90.7
UPFR	88.7	90.9	90.8	93.3	93.4	89.0	92.6
MDFR	98.0	97.7	99.2	98.2	98.3	99.6	99.5
LOFR	88.2	92.0	92.0	89.5	90.9	88.2	92.4
SOTH	97.9	99.2	99.8	99.0	99.3	99.6	99.9
LOTH	94.3	94.1	93.4	95.5	94.9	95.3	95.4
SOML	50.2	63.0	64.3	55.4	63.1	58.3	68.6
NOML	91.7	90.2	91.5	94.7	92.8	94.1	93.1
Upper Skeena River	89.8	84.5	97.1	89.5	84.6	93.9	90.2
Skeena River, Bulkley	93.7	94.9	99.0	95.6	96.2	94.9	96.1
Lower Skeena River	83.5	88.2	76.0	92.2	90.0	96.0	97.0
Nass River	79.3	85.4	81.1	85.4	86.0	82.2	87.2
Stikine River	61.1	59.2	55.3	63.2	60.2	63.8	61.8
Taku River	70.0	71.7	64.2	71.3	71.4	72.1	75.1

TABLE 7. Average accuracy of regional assignment for individual Chinook salmon for the Fisheries and Oceans Canada (DFO), Genetic Analysis of Pacific Salmon (GAPS), single-nucleotide polymorphism (SNP), DFO plus 4 microsatellites (micros), GAPS plus 2 microsatellites, DFO plus 25 SNPs, and GAPS plus 25 SNPs baselines incorporating the regions and populations indicated in Table 1. See Table 5 for more information about the loci utilized.

Region	DFO	GAPS	SNPs	DFO + 4 micros	GAPS + 2 micros	DFO + 25 SNPs	GAPS + 25 SNPs
ECVI	97.6	98.3	97.5	99.1	98.6	99.3	99.5
WCVI	98.0	98.3	96.8	99.1	99.3	99.8	99.6
UPFR	94.2	87.7	93.9	96.8	89.7	97.7	92.4
MDFR	88.5	83.9	83.5	89.1	86.1	95.2	90.8
LOFR	92.8	89.6	99.1	96.7	94.4	98.3	99.3
SOTH	92.1	93.2	98.7	97.1	93.5	96.3	97.9
LOTH	96.5	95.9	99.3	98.4	98.7	98.7	100.0
SOML	88.4	91.6	78.6	94.9	95.1	95.9	93.9
NOML	85.7	82.0	88.8	89.3	85.7	92.5	91.7
Upper Skeena River	63.7	68.0	51.7	74.7	69.0	67.7	70.9
Skeena River, Bulkley	77.4	85.1	81.6	84.4	85.2	81.2	88.1
Lower Skeena River	49.5	54.5	66.1	68.0	55.5	75.7	77.0
Nass River	84.3	88.3	92.5	86.2	89.1	93.0	95.0
Stikine River	66.4	65.0	61.1	75.0	69.8	78.0	69.0
Taku River	69.4	74.2	66.1	76.6	76.3	76.2	75.8
Mean	83.0	83.7	83.7	88.4	85.7	89.7	89.4

populations was identified with a high degree of accuracy regardless of the suite of loci used in the procedure (Table 7). With enhanced baselines, assignment of individuals to the correct region was typically accomplished with an accuracy greater than 90% for regions in southern British Columbia. In northern British Columbia, identification of individuals to the Stikine River, Taku River, and regional groups in the Skeena River displayed the lowest assignment accuracy (69–78% for the Stikine and Taku River regions and 55–88% for the Skeena River regions) when enhanced baselines were considered.

The most difficult problem encountered in stock identification applications is the correct assignment of individuals to specific populations. The highest accuracy of assignment was typically observed for the Fraser River populations, the lowest accuracy for the Stikine and Taku River populations (Table 8). There was no consistent ranking within regions of the

TABLE 8. Average accuracy of population assignment for individual Chinook salmon for the Fisheries and Oceans Canada (DFO), Genetic Analysis of Pacific Salmon (GAPS), single-nucleotide polymorphism (SNP), DFO plus 4 microsatellites (micros), GAPS plus 2 microsatellites, DFO plus 25 SNPs, and GAPS plus 25 SNPs baselines incorporating the regions and populations indicated in Table 1. See Table 5 for more information about the loci utilized.

Region	DFO	GAPS	SNPs	DFO + 4 micros	GAPS + 2 micros	DFO + 25 SNPs	GAPS + 25 SNPs
ECVI	56.4	58.3	41.9	63.2	61.0	57.0	59.5
WCVI	66.5	61.5	61.9	73.1	73.6	77.1	72.2
UPFR	73.4	62.5	69.4	81.5	68.6	74.9	65.0
MDFR	85.2	83.4	79.8	87.6	84.5	92.9	90.8
LOFR	66.5	72.3	81.4	77.2	79.5	72.2	80.0
SOTH	85.8	85.9	96.5	92.1	87.7	91.1	93.6
LOTH	81.1	79.3	80.9	81.9	83.4	84.4	88.5
SOML	45.1	47.1	47.8	47.2	53.4	53.4	52.2
NOML	68.7	63.5	64.5	78.5	72.4	76.2	74.6
Upper Skeena River	62.5	66.0	49.2	74.5	66.6	67.7	70.9
Skeena River, Bulkley	77.4	85.1	78.8	84.4	85.2	81.2	88.1
Lower Skeena River	49.5	54.5	66.1	68.0	55.5	75.7	77.0
Nass River	54.3	63.7	52.4	66.9	65.4	60.7	71.0
Stikine River	33.0	35.2	25.0	40.3	39.3	41.2	40.3
Taku River	41.2	47.8	30.0	49.7	50.1	50.8	50.9
Mean	63.1	64.4	61.7	71.1	68.4	70.4	71.6

TABLE 9. Accuracy (means) and precision (SDs) of the estimated population and regional stock compositions of single-population samples determined with ONCOR from suites of markers incorporating either the best or worst 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, and 72 single-nucleotide polymorphisms (SNPs) from Table 2. The percent correct assignment of individuals to specific populations and regions is also shown.

Number of SNPs	Class	Stock composition (%)		Standard deviation (%)		Individual assignment (%)	
		Population	Region	Population	Region	Population	Region
10	Best	59.8	88.3	10.5	5.0	21.9	56.8
	Worst	28.9	59.5	10.2	11.7	5.6	21.9
15	Best	68.0	91.7	8.2	3.5	27.7	64.2
	Worst	47.5	79.0	10.7	7.9	10.0	27.4
20	Best	71.9	94.1	7.0	2.6	33.1	70.5
	Worst	63.4	85.9	8.5	5.6	16.7	36.6
25	Best	75.5	95.2	6.2	2.2	38.0	74.2
	Worst	67.9	86.9	7.5	4.2	21.2	44.6
30	Best	80.1	95.8	5.4	1.9	41.8	77.1
	Worst	71.2	91.3	6.9	3.5	25.5	50.1
35	Best	79.8	96.1	5.0	1.7	44.3	79.3
	Worst	74.1	93.0	6.2	2.8	31.2	58.1
40	Best	80.1	96.4	4.8	1.6	46.6	81.0
	Worst	75.1	93.8	5.7	2.6	34.7	63.6
45	Best	80.6	96.6	4.5	1.5	48.9	82.7
	Worst	76.0	94.6	5.4	2.1	38.7	69.2
50	Best	81.6	96.6	4.5	1.4	51.2	84.5
	Worst	78.6	95.6	4.9	1.9	42.6	73.5
55	Best	81.7	96.6	4.2	1.3	52.0	84.9
	Worst	80.4	96.4	4.4	1.6	47.1	78.5
60	Best	81.9	97.0	4.1	1.2	53.8	85.0
	Worst	80.9	96.7	4.2	1.4	50.0	81.4
65	Best	82.7	97.0	3.9	1.2	55.7	86.8
	Worst	82.1	97.0	4.0	1.3	51.8	83.5
70	Best	83.1	97.3	3.8	1.1	56.7	87.2
	Worst	82.6	97.1	3.9	1.2	54.6	85.7
72	All	83.2	97.2	3.8	1.1	56.7	87.3

relative assignment accuracy provided by the DFO, GAPS, or SNP baselines (Table 8). Enhancing the GAPS baseline with 2 microsatellites provided the least relative increase in assignment accuracy, whereas enhancing the GAPS baseline with 25 SNPs provided the greatest increase.

How Many SNPs for Microsatellite Equivalency?

The average stock composition accuracy of the 60 single-population samples incorporating the 12 DFO microsatellites was 96.9% to geographic region, that for the 13 GAPS microsatellites was 96.7% (Table 5), and that for the 72 SNPs was 97.2% (Table 9). In essence, the 72 SNPs employed for stock composition estimation provided an accuracy in the estimation of stock composition to geographic region equivalent to or better than that of the DFO and GAPS microsatellites. The average precision of the regional estimates of stock composition was again essentially equivalent to that of the DFO microsatellites, and 80 SNPs of the average quality evaluated in our study were projected to be required to produce regional estimates of stock composition equivalent to those available from the GAPS microsatellites. When the enhanced microsatellite baselines were evaluated, 79–88 SNPs were projected to be required for the equivalency of regional accuracy and precision for the enhanced DFO microsatellites, and 68–88 SNPs were projected to be required for the GAPS baseline (Table 10).

The average population-specific accuracy derived from the DFO microsatellites was 83.0% to specific populations, 83.1% for the GAPS microsatellites, 85.6% for the enhanced DFO microsatellites, and 85.5% for the enhanced GAPS microsatellites (Table 5). The 72 SNPs evaluated already provided population-specific accuracy equivalent to that available from the existing DFO and GAPS microsatellites for most populations (Figure A.2). The accuracy available from the enhanced DFO or GAPS suites of microsatellites was projected to require between 118 and 122 SNPs (Table 10). The variability of the population-specific estimates derived from the existing DFO

TABLE 10. Estimated number of single-nucleotide polymorphisms (SNPs) required to equal the performance of different suites of microsatellites with respect to accuracy and precision (SD) of population and regional estimates of stock composition, as well as the percent correct individual assignment to population and region, based on fitting functions to the results in Table 9. The mean accuracy and precision and the accuracy of individual assignments obtained from the different suites of microsatellites are given in Table 5.

	Stock composition (%)		Standard deviation (%)		Individual assignment (%)	
Loci	Population	Region	Population	Region	Population	Region
DFO	72	64	110	73	93	73
DFO + 4	122	79	179	88	137	90
GAPS	73	60	133	80	94	74
GAPS + 2	118	68	166	88	121	82

and GAPS microsatellites was less than that available from the 72 SNPs for most populations (Figure A.1). The precision of the population-specific estimates available from the existing microsatellite baselines was projected to require 122 SNPs for equivalency with that of the DFO microsatellites and 118 SNPs for equivalency with that of the GAPS microsatellites. The enhanced microsatellite baselines were projected to require 179 SNPs and 166 SNPs, respectively, for equivalency of the precision of population-specific estimates (Figure 5; Table 10).

The average accuracy of the assignment of individuals to specific regions was 87.0% for the DFO microsatellites and 87.3% for the GAPS microsatellites, and 73–74 SNPs of the average quality evaluated in the study were estimated to be required to provide equivalent accuracy in the assignment of individuals to region. The SNPs surveyed in the study provided equivalent levels of regional assignment accuracy for individuals. The level of regional accuracy of individual assignment available from the enhanced DFO and GAPS suites of microsatellites was projected to require 90 and 82 SNPs, respectively (Figure 5; Table 10). Assignment of individuals to specific populations was achieved with an average accuracy of 61.5% for the DFO microsatellites and 61.8% for the GAPS microsatellites. These levels of accuracy were projected to be achieved with 93–94 SNPs of the average quality surveyed. The levels of accuracy of individual assignment available from the enhanced DFO and GAPS suites of microsatellites were projected to require 137 and 121 SNPs, respectively (Table 10).

DISCUSSION

Sample Size

Two alleles were observed at the SNP loci, but up to 62 alleles were observed at the microsatellite loci. Therefore, microsatellites require more fish to be sampled in a population to obtain estimates of allele frequencies with similar accuracy and precision than do SNPs. Beacham et al. (2011) demonstrated that microsatellites required larger baseline samples than SNPs to reduce the sampling variation in the estimation of allele frequencies and thus increase the accuracy of the estimated stock compositions. The population sample size was required to be about two to three times larger in the microsatellite baselines than in the SNP baselines before equivalent levels of accuracy relative to the asymptotic value were obtained. As most population sample sizes were approximately 100 individuals for the survey of SNP variation in our study, microsatellite sample size was capped at a maximum of 200 individuals per population to estimate microsatellite allele frequencies. The accuracy of the estimated stock composition of single-population samples of salmon derived from microsatellites can be limited by the baseline population sample size, with sample sizes of approximately 200 individuals being required before there is little effect of sample size on the accuracy of population-specific estimates (Beacham et al. 2006a). For sockeye salmon *O. nerka*, Beacham et al. (2010) showed that once approximately 95 individuals within a population had been sampled at SNP loci, there was virtually no increase in the accuracy of estimated stock compositions or individual assignments. Thus, the accuracy of the estimated stock compositions and assignments of individuals derived from SNPs was not limited by population sample size in our study. Comparison of the utility of SNPs and microsatellites for stock identification requires that adequate sample sizes be available for the populations included in the analyses for both classes of markers.

Relative Ranking of the Markers

In the current study, the survey of microsatellite variation included loci with 7–62 alleles. The number of alleles observed at a microsatellite locus was related to the accuracy and precision of the estimated stock compositions. Microsatellites with larger numbers of alleles provided more accurate and precise estimates than did microsatellites with few alleles. Similar empirical results have previously been reported for Chinook salmon in the Yukon River drainage (Beacham et al. 2008b) and on a Pacific Rim basis (Beacham et al. 2006a). Of the top 10 microsatellites surveyed in the current study, 5 were also in the top 10 evaluated in the Yukon River; 2 of the remaining loci were not surveyed in the previous study and 3 were of lesser value. The 5 loci in the top 10 in both studies were *Ots100*, *Ots107*, *Oki100*, *Omm1080*, and *Ots211*, which, given the widely divergent geographical
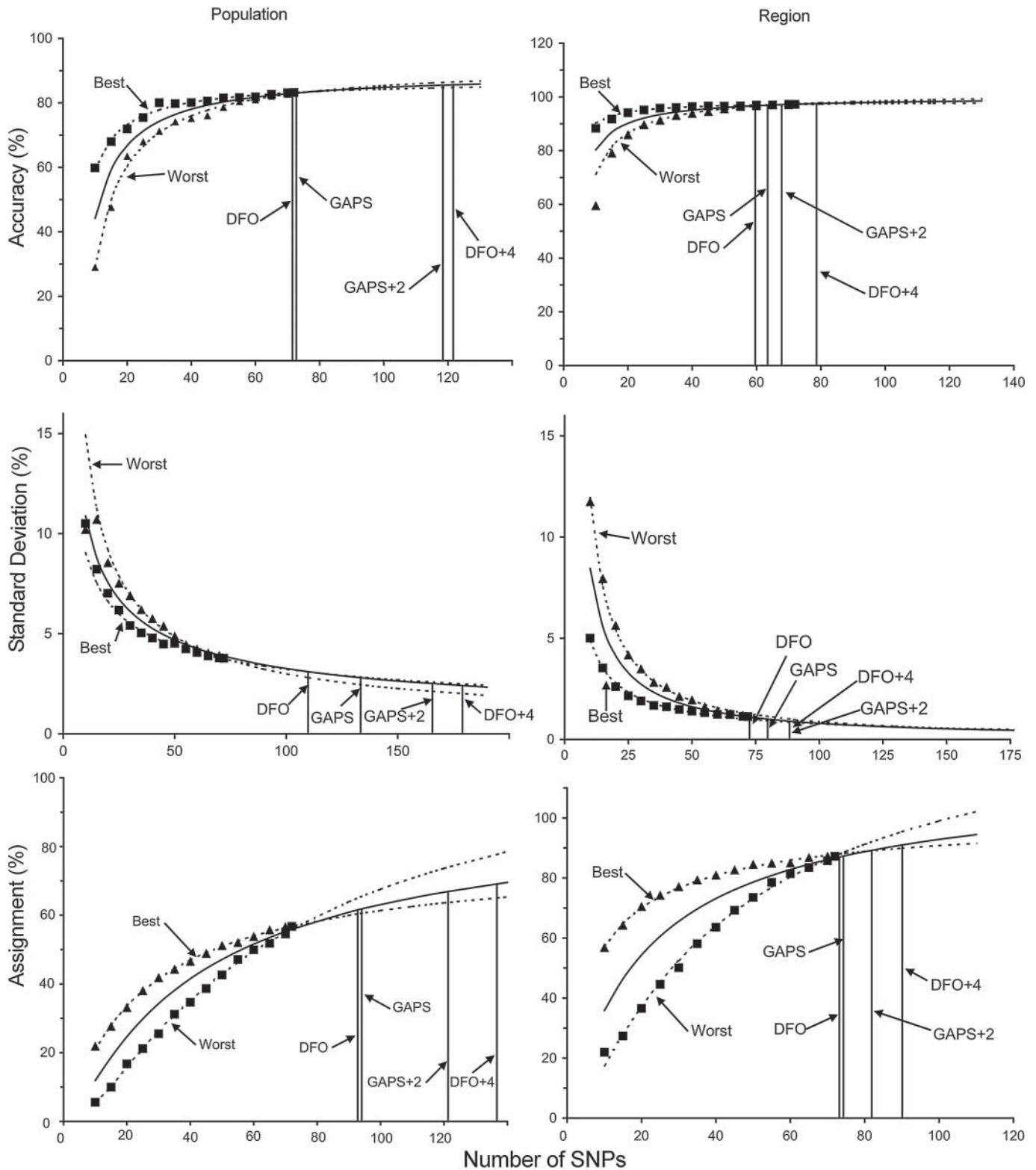
FIGURE 5. Accuracy and standard deviations of regional and population estimates of stock composition vis-à-vis the number of single-nucleotide polymorphisms (SNPs) employed in the estimation procedure, along with the accuracy of individual assignments to specific regions or populations. The vertical lines indicate the number of SNPs projected to be required to provide results comparable to those of different suites of microsatellites. The estimates were bounded by initially incorporating the SNPs with the highest average accuracy into the analyses of the single-population mixtures, with lower-accuracy SNPs being added sequentially; the average population accuracy and precision were recorded for each set of SNPs (squares). Subsequently, the SNPs with the lowest average accuracy were initially incorporated in the analyses, with progressively higher-accuracy SNPs being added sequentially; the average accuracy and precision were again recorded (triangles). The overall mean accuracy and precision for each specified number of SNPs were determined by averaging the results from both processes; these values were then used in the projections.

settings of the two studies, is indicative of the general power of these loci for stock identification applications.

Anderson (2010) reported that the accuracy of population assignment for a set of genetic markers can be overestimated in routine applications if the suite of markers has been chosen specifically for the observed differentiation in allele frequency among existing population samples. Sampling error in the estimation of baseline allele frequencies may lead to an optimistic assessment of their accuracy in routine applications. Accuracy would ideally be tested with individuals not included in the baseline samples. In our study, all of the samples available for a population were used to determine SNP allele frequencies, and all of the samples were used for 62% of the populations surveyed for microsatellites. Although the accuracy of population assignment was not tested with new individuals, the trends in accuracy among the different suites of microsatellites and SNPs would probably not have been affected.

Although other methods for evaluating the relative power of loci for population differentiation and individual assignment have been applied (Rosenberg et al. 2003; Hedrick 2005; Narum et al. 2008), power for stock identification was the sole basis of evaluating the loci surveyed in the study. In a previous study evaluating the effectiveness of 13 microsatellites and 37 SNPs for assigning individuals to 29 specific populations of Chinook salmon, Narum et al. (2008) reported that the best 10 loci for correct individual assignment were microsatellites and that the best 15 loci included 12 microsatellites and 3 SNPs. Assignment accuracy was higher than when using either SNPs or microsatellites alone. In our study, which centered on the accuracy of population-specific estimates of stock composition as a measure of power of the locus for stock identification and which incorporated 29 microsatellites and 73 SNPs across 60 populations in British Columbia, the top 29 of the 103 markers evaluated in our study were all microsatellites, with the accuracy of the estimated population-specific stock compositions produced by incorporating the least-informative microsatellite being approximately 2.5 times that of the highest-ranked SNP. The 12 highest-ranked SNPs in the study of Narum et al. (2008) were incorporated in our survey, and all of these SNPs were found to be less valuable in estimating stock composition than the least-powerful microsatellite.

DFO and GAPS Microsatellites

The accuracy and precision of the stock composition estimates provided by the 12-locus DFO and the 13-locus GAPS sets of microsatellites were essentially equivalent, although both were lower than that provided by an optimum set of 13 microsatellites. If increased accuracy and precision of stock composition estimates is required for either set of loci, there are three potential solutions. The first potential solution is to increase the number of microsatellites incorporated in the suites of loci, the second is to augment the current suite of microsatellites with higher-resolution SNPs, and the third is to replace the microsatellites entirely with a set of SNPs. Increasing the

resolution of stock composition estimates in the real world is constrained by the cost of producing the estimates for a particular sample.

On a practical basis, when microsatellites are used for stock composition estimation, constraining the cost typically centers on limiting the number of injections on the automated sequencer. If the number of injections is arbitrarily capped at three for analysis of Chinook salmon microsatellite variation, then four microsatellites can be added to the DFO suite and two to the GAPS suite. When these additional microsatellites were incorporated into our analysis, the improvements in the accuracy and precision of the estimated stock compositions were similar with both suites of loci, with an average population-specific accuracy of 85.5% over all populations and an average standard deviation of 2.5%. When the second option (augmenting the microsatellites with a suite of high-resolution SNPs) was employed, the addition of 20–25 SNPs achieved the accuracy and precision of the population-specific stock composition estimates achieved with the augmented suite of microsatellites. If the 85.5% average population-specific accuracy and 2.5% standard deviation observed in our survey are considered acceptable for management applications, the choices are clear as to how to improve the accuracy of estimated stock compositions. For the GAPS microsatellites, adding two microsatellites to the suite will increase the cost of analysis by the expense of conducting two polymerase chain reactions and the time required to analyze two additional loci. The resulting cost would be compared with that of surveying an additional 20–25 SNPs. These cost comparisons can be made by the individual laboratories that apply the GAPS microsatellites; in our laboratory, it is more efficient to add two microsatellites to the survey than it is to add 20–25 SNPs.

Microsatellites or SNPs?

Accurate and precise regional estimates of stock composition are generally the easiest to produce, followed by population-specific estimates, with the assignment of individuals to specific populations being the most difficult. Assessment of the accuracy and precision of the regional estimates of stock composition produced by the application of 72 SNPs indicated that they were equivalent to the accuracy and precision available from the existing DFO and GAPS microsatellites. If that level of resolution is all that is required, either the existing sets of microsatellites or the SNPs could be utilized. However, if the existing DFO and GAPS microsatellites are enhanced by either four or two microsatellites, respectively, the SNP baseline evaluated would require some enhancement. If population-level estimates of stock composition are required for some regions of the baseline utilized, an additional 46–50 SNPs of the average quality evaluated in the study would be required to provide population-specific accuracy and precision comparable to those obtained with the microsatellites. An additional 94–107 SNPs were projected to be required if population-specific results comparable to those available from the enhanced microsatellite baselines for British Columbia Chinook salmon were required.

Single-nucleotide polymorphism arrays containing 100–200 loci were also estimated to be required to meet management standards for fine-scale resolution of Columbia River Chinook salmon (Hess et al. 2011). If individual assignment is part of the stock identification application, then no enhancement of the existing SNP baseline is required if the regional levels of accuracy provided by the DFO and GAPS microsatellites is acceptable, and only about 20 additional SNPs would be required if the enhanced microsatellite baselines were utilized. An additional 49–65 SNPs was projected to be required if population-specific assignment results comparable to those available from the enhanced microsatellite baselines were required. However, if SNPs are developed that provide higher resolution than the average of the 72 SNPs used in our projections, fewer additional SNPs will be required to produce stock composition results of a quality equivalent to that of the microsatellites.

The cost of laboratory analysis for individual fish is a key factor in deciding the appropriate technology to use in a particular laboratory. Technologies are replaced when one provides a clear advantage over another, such as substituting microsatellites or SNPs for allozymes in analyses of the genetic variation in Pacific salmon. Although a number of techniques are available to survey SNP variation, we are aware of none that will allow well over 100 SNPs to be analyzed in our laboratory at a cost comparable to that of analyzing up to 16 microsatellites with three injections on an automated DNA sequencer. For the present, the combined microsatellite–SNP approach outlined by Narum et al. (2008) and Hess et al. (2011) may be a practical approach to incorporating the power of both classes of markers.

ACKNOWLEDGMENTS

REFERENCES

Anderson, E. C. 2010. Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. Molecular Ecology Resources 10:701–710.

Banks, M. A., M. S. Blouin, B. A. Baldwin, V. K. Rashbrook, H. A. Fitzgerald, S. M. Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in Chinook salmon (*Oncorhynchus tshawytscha*). Journal of Heredity 90:281–288.

Beacham, T. D., J. R. Candy, K. L. Jonsen, J. Supernault, M. Wetklo, L. Deng, K. M. Miller, and R. E. Withler. 2006a. Estimation of stock composition and individual identification of Chinook salmon across the Pacific Rim using microsatellite variation. Transactions of the American Fisheries Society 135:861–888.

Beacham, T. D., J. R. Candy, B. McIntosh, C. MacConnachie, A. Tabata, K. Kaukinen, L. Deng, K. M. Miller, R. E. Withler, and N. Varnavskaya. 2005. Estimation of stock composition and individual identification of sockeye salmon on a Pacific Rim basis using microsatellite and major histocompatibility complex variation. Transactions of the American Fisheries Society 134:1124–1146.

Beacham, T. D., J. R. Candy, K. J. Supernault, M. Wetklo, B. Deagle, K. Labaree, J. R. Irvine, K. M. Miller, R. J. Nelson, and R. E. Withler. 2003. Evaluation and application of microsatellites for population identification of Fraser River Chinook salmon (*Oncorhynchus tshawytscha*). U.S. National Marine Fisheries Service Fishery Bulletin 101:243–259.

Beacham, T. D., K. L. Jonsen, J. Supernault, M. Wetklo, L. Deng, and N. Varnavskaya. 2006b. Pacific Rim population structure of Chinook salmon as determined from microsatellite variation. Transactions of the American Fisheries Society 135:1604–1621.

Beacham, T. D., M. Lapointe, J. R. Candy, B. McIntosh, C. MacConnachie, A. Tabata, K. Kaukinen, L. Deng, K. M. Miller, and R. E. Withler. 2004. Stock identification of Fraser River sockeye salmon (*Oncorhynchus nerka*) using microsatellites and major histocompatibility complex variation. Transactions of the American Fisheries Society 133:1117–1137.

Beacham, T. D., B. McIntosh, and C. Wallace. 2010. A comparison of stock and individual identification for sockeye salmon (*Oncorhynchus nerka*) in British Columbia provided by microsatellites (STRs) and single-nucleotide polymorphisms (SNPs). Canadian Journal of Fisheries and Aquatic Sciences 67:1274–1290.

Beacham, T. D., B. McIntosh, and C. Wallace. 2011. A comparison of polymorphism of genetic markers and population sample sizes required for mixed-stock analysis of sockeye salmon (*Oncorhynchus nerka*) in British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 68:550–562.

Beacham, T. D., S. Urawa, K. D. Le, and M. Wetklo. 2008a. Population structure and stock identification of chum salmon from Japan determined with microsatellite DNA variation. Fisheries Science 74:983–994.

Beacham, T. D., M. Wetklo, C. Wallace, J. B. Olsen, B. G. Flannery, J. K. Wenburg, W. D. Templin, A. Antonovich, and L. W. Seeb. 2008b. The application of microsatellites for stock identification of Yukon River Chinook salmon. North American Journal of Fisheries Management 28:283–295.

Beacham, T. D., I. Winther, K. L. Jonsen, M. Wetklo, L. Deng, and J. R. Candy. 2008c. The application of rapid microsatellite-based stock identification to management of a Chinook salmon troll fishery off the Queen Charlotte Islands, British Columbia. North American Journal of Fisheries Management 28:849–855.

Buchholz, W. G., S. J. Miller, and W. J. Spearman. 2001. Isolation and characterization of chum salmon microsatellite loci and use across species. Animal Genetics 32:162–165.

Cairney, M., J. B. Taggart, and B. Hoyheim. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. Molecular Ecology 9:2175–2178.

Campbell, N. R., and S. R. Narum. 2008. Identification of novel single-nucleotide polymorphisms in Chinook salmon and variation among life history types. Transactions of the American Fisheries Society 137:96–106.

Danzmann, R. G., M. Cairney, W. S. Davidson, M. M. Ferguson, K. Gharbi, R. Guyomard, L. E. Holm, E. Leder, N. Okamoto, A. Ozaki, C. E. Rexroad III, T. Sakamoto, J. B. Taggart, and R. A. Woram. 2005. A comparative analysis of the rainbow trout genome with 2 other species of fish (Arctic charr and Atlantic salmon) within the tetraploid derivative Salmonidae family (subfamily: Salmoninae). Genome 48:1037–1051.

Debevec, E. M., R. B. Gates, M. Masuda, J. Pella, J. M. Reynolds, and L. W. Seeb. 2000. SPAM (version 3.2): statistics program for analyzing mixtures. Journal of Heredity 91:509–511.

Glover, K. A., M. M. Hansen, S. Lien, T. D. Als, B. Høyheim, and Ø. Skaala. 2010. A comparison of SNP and STR loci for delineating population

structure and performing individual genetic assignment. BMC (BioMed Central) Genetics 11:2.

Goudet, J. 1995. FSTAT: a program for IBM PC compatibles to calculate Weir and Cockerham's (1984) estimators of *F*-statistics (version 1.2). Journal of Heredity 86:485–486.

Grieg, C. A., and M. A. Banks. 1999. Five multiplexed microsatellite loci for rapid response run identification of California's endangered winter Chinook salmon. Animal Genetics 30:318–320.

Grieg, C., D. P. Jacobson, and M. A. Banks. 2003. New tetranucleotide microsatellites for fine-scale discrimination among endangered Chinook salmon (*Oncorhynchus tshawytscha*). Molecular Ecology Notes 3:376–379.

Haasl, R. J., and B. A. Payseur. 2010. Multilocus inference of population structure: a comparison between single-nucleotide polymorphisms and microsatellites. Heredity 106:158–171.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633–1638.

Hess, J. E., A. P. Matala, and S. R. Narum. 2011. Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River basin. Molecular Ecology Resources 11(Supplement 1):137–149.

Kalinowski, S. T. 2002. How many alleles per locus should be used to estimate genetic distances? Heredity 88:62–65.

Kalinowski, S. T. 2004. Genetic polymorphism and mixed-stock fisheries analysis. Canadian Journal of Fisheries and Aquatic Sciences 61:1075–1082.

Kalinowski, S. T., K. R. Manlove, and M. L. Taper. 2007. ONCOR: software for genetic stock identification. Montana State University, Bozeman. Available: www.montana.edu/kalinowski/Software/ONCOR.htm. (September 2008).

Miller, K. M., T. D. Beacham, and T. J. Ming. 2008. Report to the Northern Endowment Fund on project NF-2007-I4: Chinook salmon SNP development. Pacific Salmon Commission, Vancouver.

Narum, S. R., M. Banks, T. Beacham, R. Bellinger, M. Campbell, J. DeKoning, A. Elz, C. Guthrie, C. Kozfkay, K. Miller, P. Moran, R. Phillips, L. Seeb, C. Smith, K. Warheit, S. Young, and J. C. Garza. 2008. Differentiating populations at broad and fine geographic scales with microsatellites and single-nucleotide polymorphisms. Molecular Ecology 17:3464–3477.

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

O'Connell, M., R. G. Danzmann, J. M. Cornuet, J. M. Wright, and M. M. Ferguson. 1997. Differentiation of rainbow trout stocks in Lake Ontario and the evaluation of the stepwise mutation and infinite allele mutation models using microsatellite variability. Canadian Journal of Fisheries and Aquatic Sciences 54:1391–1399.

Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7:1087–1090.

O'Reilly, P. T., L. C. Hamilton, S. K. McConnell, and J. M. Wright. 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Sciences 53:2292–2298.

Parken, C. K., J. R. Candy, J. R. Irvine, and T. D. Beacham. 2008. Genetic and coded wire tag results combine to allow more precise management of a complex Chinook salmon aggregate. North American Journal of Fisheries Management 28:328–340.

Pereira da Silva, W. P., and C. D. Pereira da Silva. 2007. Lab Fit curve fitting software (nonlinear regression and treatment of data program), version 7.2.46 (1999–2009). Available: www.labfit.net. (August 2009).

PSC (Pacific Salmon Commission). 2008. Recommendations for application of genetic stock identification (GSI) methods to management of ocean salmon fisheries: special report of the GSI Steering Committee and the Pacific Salmon Commission's Committee on Scientific Cooperation. PSC Technical Report 23, Vancouver.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the USA 94:9197–9201.

Rexroad, C. E., R. L. Coleman, A. L. Gustafson, W. K. Hershberger, and J. Killefer. 2002. Development of rainbow trout microsatellite markers from repeat enriched libraries. Marine Biotechnology 4:12–16.

Rexroad, C. E., R. L. Coleman, A. M. Martin, W. K. Hershberger, and J. Killefer. 2001. Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 32:283–319.

Rosenberg, N. A., L. M. Li, R. Ward, and J. K. Pritchard. 2003. Informativeness of genetic markers for inference of ancestry. American Journal of Human Genetics 73:1402–1422.

Sakamoto, T., R. G. Danzmann, K. Gharbi, P. Howard, A. Ozaki, S. K. Khoo, R. A. Woram, N. Okamoto, M. M. Ferguson, L. E. Holm, R. Guyomard, and B. Hoyheim. 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. Genetics 155:1331–1345.

Seeb, L. W., A. Antonovich, M. A. Banks, T. D. Beacham, M. R. Bellinger, M. Campbell, N. A. Decovich, J. C. Garza, C. M. Guthrie III, P. Moran, S. R. Narum, J. J. Stephenson, K. J. Supernault, D. J. Teel, W. D. Templin, J. K. Wenburg, S. F. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32:540–552.

Small, M. P., T. D. Beacham, R. E. Withler, and R. J. Nelson. 1998. Discriminating coho salmon within the Fraser River, British Columbia, using microsatellite DNA markers. Molecular Ecology 7:141–155.

Smith, C. T., A. Antonovich, W. D. Templin, C. M. Elfstrom, S. R. Narum, and L. W. Seeb. 2007. Impacts of marker class bias relative to locus-specific variability on population inferences in Chinook salmon: a comparison of single-nucleotide polymorphisms with short tandem repeats and allozymes. Transactions of the American Fisheries Society 136:1674–1687.

Smith, C. T., J. Baker, L. Park, L. W. Seeb, C. Elfstrom, S. Abe, and J. E. Seeb. 2005a. Characterization of 13 single-nucleotide-polymorphism markers for chum salmon. Molecular Ecology 14:259–262.

Smith, C. T., C. Elfstrom, L. W. Seeb, and J. E. Seeb. 2005b. Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Molecular Ecology Notes 5:4193–4203.

Smith, C. T., B. F. Koop, and R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1613–1621.

Smith, C. T., J. E. Seeb, P. Schwenke, and L. W. Seeb. 2005c. Use of the 5′-nuclease reaction for single-nucleotide-polymorphism genotyping in Chinook salmon. Transactions of the American Fisheries Society 134:207–217.

Smith, C. T., W. D. Templin, J. E. Seeb, and L. W. Seeb. 2005d. Single-nucleotide polymorphisms (SNPs) provide rapid and accurate estimates of the proportions of U.S. and Canadian Chinook salmon caught in Yukon River fisheries. North American Journal of Fisheries Management 25:944–953.

Waples, R. S., D. J. Teel, J. M. Myers, and A. R. Marshall. 2004. Life history divergence in Chinook salmon: historic contingency and parallel evolution. Evolution 58:386–403.

Williamson, K., S. Cordes, F. Jan, and B. May. 2002. Characterization of microsatellite loci in Chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. Molecular Ecology Notes 2:17–19.
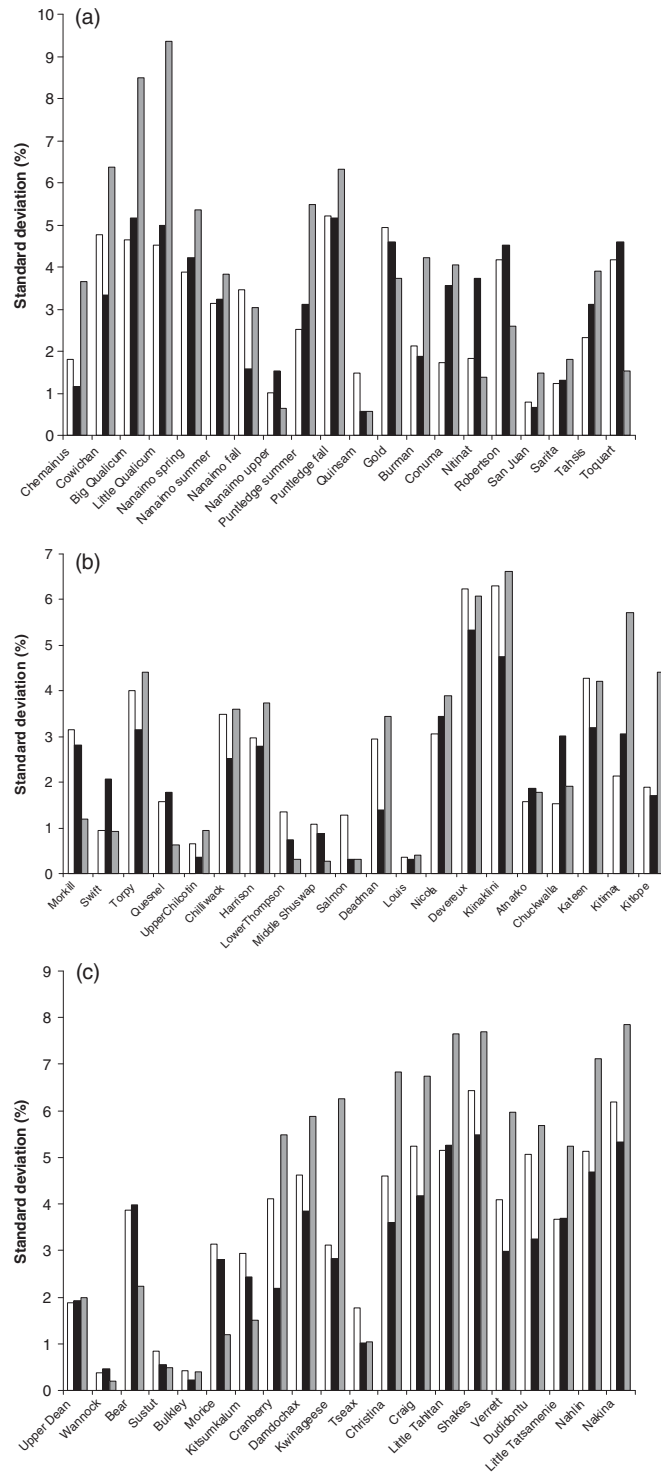
Appendix: Additional Details

FIGURE A.1. Mean standard deviations of assignment to population derived from the DFO microsatellite (white bars), GAPS microsatellite (black bars), and single-nucleotide polymorphism (gray bars) suites of loci for **(a)** Vancouver Island, **(b)** Fraser River, southern mainland, and northern mainland, and **(c)** additional northern mainland, Skeena River, Nass River, Stikine River, and Taku River Chinook salmon populations for simulated single-population samples.
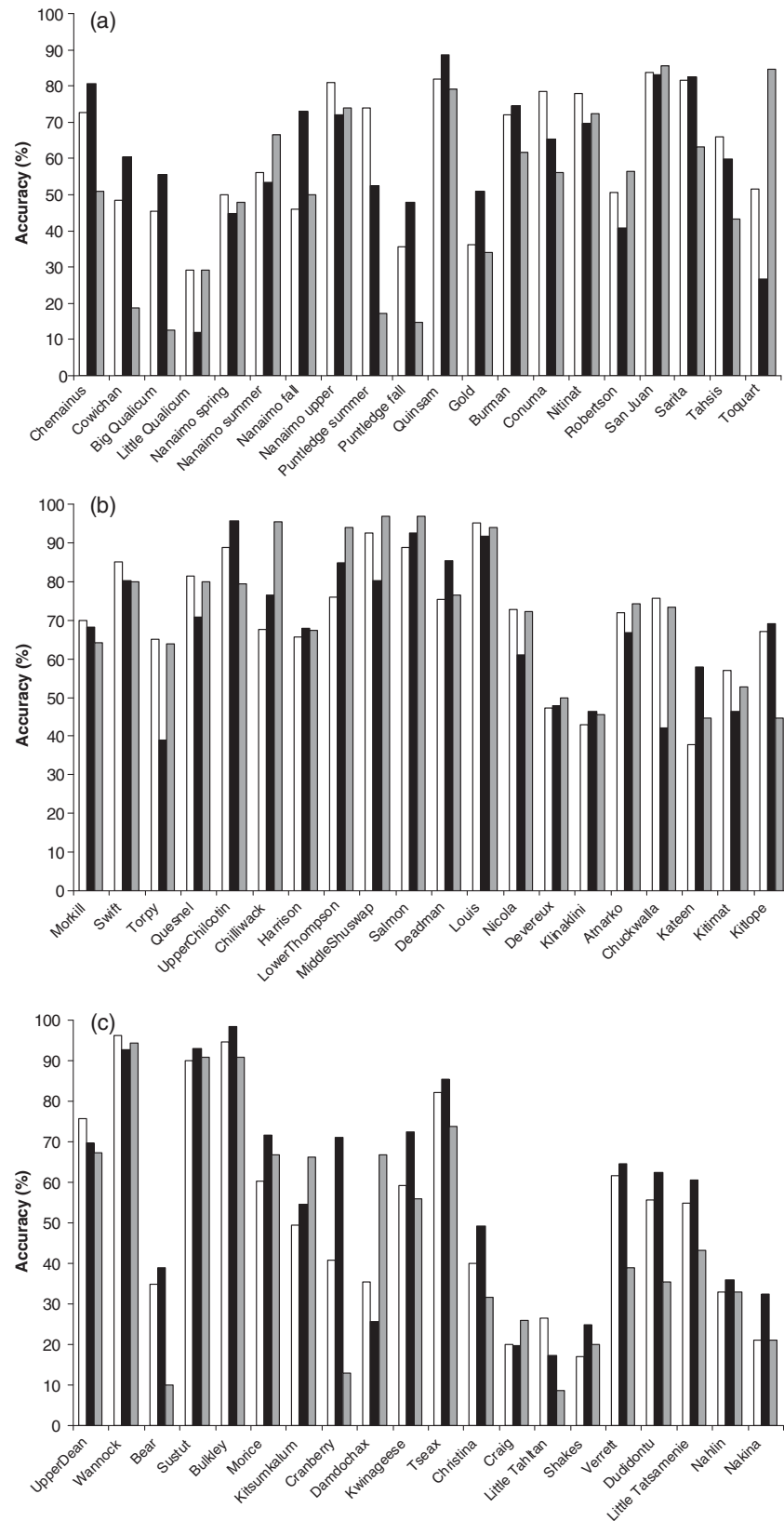
FIGURE A.2.   Mean population accuracy of individual assignment by population derived from the DFO microsatellite (white bars), GAPS microsatellite (black bars), and single-nucleotide polymorphism (gray bars) suites of loci for **(a)** Vancouver Island, **(b)** Fraser River, southern mainland, and northern mainland, and **(c)** additional northern mainland, Skeena River, Nass River, Stikine River, and Taku River Chinook salmon populations for simulated single-population samples.