The genetic population structure of lacustrine sockeye salmon, *Oncorhynchus nerka*, in Japan as the endangered species

Yuichiro Kogura[1] · James E. Seeb[2] · Noriko Azuma[3] · Hideaki Kudo[4] · Syuiti Abe[4] · Masahide Kaeriyama[4]

[1]Graduate School of Fisheries Sciences, Hokkaido University, 3-1-1, Minato-cho, Hakodate, Hokkaido 041-8611, Japan

[2]School of Aquatic and Fisheries Science, University of Washington, 1122 Boat ST NE Seattle, WA 98146, USA

[3]Nodai Bioresources Institute, Tokyo University of Agriculture, 196, Yasaka, Abashiri, Hokkaido, 099-2493, Japan

[4]Faculty of Fisheries Sciences, Hokkaido University, 3-1-1, Minato-cho, Hakodate, Hokkaido 041-8611, Japan

*Corresponding author: salmon@fish.hokudai.ac.jp Tel & Fax: (+81)138-40-5605

Abstract Lacustrine sockeye salmon (*Oncorhynchus nerka*) are listed as an endangered species in Japan despite little genetic information on their population structure. In order to clarify the genetic diversity and structure of Japanese populations for evaluating on the bottleneck effect and an endangered species, we analyzed the *ND5* region of mitochondrial DNA (mtDNA) and 45 single nucleotide polymorphisms (SNPs) in 640 lacustrine sockeye salmon in Japan and 80 anadromous sockeye salmon in Iliamna Lake of Alaska. The genetic diversity of the Japanese population in both mtDNA and SNPs was significantly less than that of the Iliamna Lake population. Moreover, all Japanese populations had SNP loci deviating from the HWE. In spite of low genetic diversity, the SNP analyses resulted that the Japanese population was significantly divided into three groups. These suggest that Japanese sockeye salmon populations should be protected as an endangered species and genetically disturbed by the hatchery program and transplantations.

Keywords Lacustrine sockeye salmon · *Oncorhynchus nerka* · Population structure · Single nucleotide polymorphisms · Mitochondrial DNA · Bottleneck effect

Introduction

Sockeye salmon (*Oncorhynchus nerka*) range widely in the North Pacific Ocean and northern adjacent Bering Sea and Okhotsk Sea. They spawn mainly in lakes and river systems around the Pacific Rim from the Columbia River Drainage to the southern portion of the Kamchatka Peninsula and the Kuril Islands. Life history patterns of sockeye salmon are classified into three types: anadromous, lake resident (lacustrine), and kokanee (Ricker 1940). Anadromous sockeye salmon migrate to the sea 1 or 2 years after living in freshwater and stay and grow in the ocean from 1 to 4 years (Burgner 1991). Kokanee salmon, which are derived from anadromous fish, live in a freshwater habitat all of their lives. Lacustrine sockeye salmon are classified as atavismus of kokanee; a part of them migrate to the sea as smolt due to a lack of resources, such as food and habitat, to satisfy their energy metabolism (Ricker 1940; Kaeriyama 1996). Although anadromous sockeye and kokanee salmon are sympatric in many watersheds and concurrently spawn in the same locations in some cases, these two types are reproductively isolated and genetically differentiated due to different growth rates and maturity timing (Foote and Larkin 1988; Wood and Foote 1996). Iliamna Lake in Alaska is the largest sockeye-producing site in the world (Burgner 1991). In Japan, although the anadromous sockeye salmon is not found naturally, lacustrine sockeye salmon are distributed among several lakes (Kaeriyama 1991; Figure 1). Moreover, the Japanese lacustrine sockeye salmon was recently listed as an endangered species by the Japanese Ministry of the Environment (Red List; Ministry of the Environment, Japan 2007).

The genetic population structure of the Pacific salmon has been clarified using various methods such as allozyme (Winans et al. 1994), mitochondrial DNA (mtDNA; Seeb and Crane 1999), and microsatellite DNA (msDNA; Beacham et al. 2006). The mtDNA and msDNA methods are powerful tools for genealogical identification, despite limitations of analytical precision (Zhang and Hewitt 2003). For example, mtDNA sequencing represents only single maternally inherited loci, and msDNA loci suffer from variable null alleles and mutation patterns, introducing ambiguity into data analyses. The loci can also be sparse in the genome and thus become difficult to isolate in some species (Navajas et al. 1998). In contrast, single nucleotide polymorphisms (SNPs) are a powerful tool for these applications (Smith et al. 2005). SNPs provide a better representation of the variation in a species genome than msDNA or allozymes. Additionally, SNPs are easily standardized across research groups and well suited for high-throughput genotyping (Brumfield et al. 2003). Studies using SNPs in *Oncorhynchus* spp. in Japan, hodwever, have not yet been conducted.

Japanese sockeye salmon population is uniquely located on the msDNA dendrogram of sockeye salmon in the North Pacific because of reducing genetic diversity relative to other populations (Beacham et al. 2006). Native Japanese lacustrine sockeye salmon have distributed in Akan Lake and Chimikeppu Lake in Hokkaido (Oshima 1934). The Akan

Lake lacustrine sockeye salmon were first transplanted to Shikotsu Lake in 1893 (Fujimura 1900). After that, they were re-transplanted from Shikotsu Lake to some oligotrophic lakes (Fig. 1). For instance, Lacustrine sockeye salmon were transplanted from Shikotsu Lake to Towada Lake in 1902, Tachibana Lake in 1911 and Abira River in 1985 at the first. Tachibana Lake had a transplant of Shikotsu Lake sockeye, despite multiple-transplant in other lakes. In the 1920s, the Shikotsu Lake lacustrine sockeye salmon experienced a catastrophic crash due to sexual imbalance and abnormal gonad development due to the overpopulation caused by over-released juveniles from the hatchery (Kaeriyama 1993). Kaeriyama (1991) also estimated that their effective population size was zero individuals in 1925 and 1927 based on the record of the hatchery reports. Subsequently in Shikotsu Lake, anadromous sockeye salmon eggs (4.53 million total) were transplanted 11 times between 1925 and 1940 from Urumobetsu Lake on Iturup Island in the Kril Islands (Tokui 1964). Because of these transplants after the population crush, Kaeriyama (1991) assumed that the Japanese native lacustrine sockeye salmon was displaced to the population of the Urumobetsu Lake anadromous sockeye salmon in Shikotsu Lake. However, Winans and Urawa (2000) noted genetic differences between the Japanese sockeye salmon (Shikotsu Lake, Towada Lake, and the Abira River) and Iturup Island's population (Sopochnoye Lake), despite no genetic differentiation among Japanese populations based on allozyme variability. Microsatellite DNA (msDNA) analysis revealed that a Japanese sockeye salmon population (the Abira River population) with the lowest genetic diversity was distinctive from North American and Russian populations including Kamchatka populations (Beacham et al. 2006). Genetic studies are very few regarding the lacustrine sockeye salmon population in Japan (Winans and Urawa 2000; Beacham et al. 2006), and no information on the genetic structure of the native Akan Lake and other lake populations including the Tachibana Lake population, although Japanese sockeye salmon is an endangered species.

The present study aimed to enhance 1) the genetic tool for assessing the genetic diversity of sockeye populations, and to find 2) the influence of transplantation especially from the Urumobetsu Lake anadromous sockeye salmon into the Japanese populations, and 3) the bottleneck effect in the Japanese sockeye population. We also analyzed the anadromous Iliamna Lake sockeye salmon populations in Iliamna Lake, Alaska, to compare with the population structure of Japanese sockeye salmon.

Materials and methods

Samples of Japanese sockeye salmon (n = 640) were collected from Akan Lake in 2004 and 2008 (AKA04, AKA08), Shikotsu Lake in 2003 and 2008 (SHI03, SHI08), Towada Lake in 2004 and 2008 (TOW04, TOW08), Tachibana Lake in 2003 (TAC03), and the Abira River in 1994 (ABI94) during the autumn. Sockeye salmon samples were also collected

from Iliamna Lake, the Kubichack River system, Alaska, USA, in the summer of 2003 (n = 81). These samples consisted of two spawning types: lake river-spawn (ILIR03) and lake-spawn types (ILIL03) (Fig. 1, Table 1). Samples of heart, dorsal fin, and adipose fin were kept in 100% ethanol at –20ºC until preparation of genomic DNA. Genomic DNA was extracted using a DNeasy 96 Tissue Kit (Qiagen, Valencia, CA). Extracted DNA was dissolved in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) and stored at –20ºC until use.

mtDNA assay

In this study, direct sequencing of the Japanese populations was performed to read the 520 base pair (bp) 5' half of the mtDNA NADH dehydrogenase subunit 5 (*ND5*) gene. The primer sets used in this study were as follows: forward primer, 5'-*TACCCCAATTGCCCTGTACG*-3', and reverse primer, 5'-*TAGGCTCCCGATTGTGAGAC*-3' (Kitanishi et al. 2007). Polymerase chain reaction (PCR) was performed under the following conditions: preheating at 95ºC for 5 min, followed by 40 cycles of denaturation at 95ºC for 45 s, annealing at 56ºC for 30 s, and extraction at 72ºC for 2 min, and completed with a final extension at 72ºC for 10 min. The PCR products were used for the sequencing reaction with a BigDye Terminator Cycle Sequencing Kit (ver. 3.1; Applied Biosystems, Foster City, CA). Direct nucleotide sequencing was performed using the same primers as described above with the following sequencing reaction: preheating at 96ºC for 1 min, followed by 25 cycles of denaturation at 96ºC for 20 s, annealing at 50ºC for 20 s, and an extension at 60ºC for 4 min.

SNP assay

Multiplex genotyping was performed using the BioMark 96.96 Dynamic Array (Fluidigm, South San Francisco, CA). The Dynamic Array contains a matrix of integrated channels and valves housed in an input frame. The input frame has 96 inlets for the sample DNA, and 96 inlets for the assays of 96 SNP markers. BioMark 96.96 Dynamic Arrays allowed for the genotyping of 95 individuals per 96-well plate (with one inlet used as a no-template control using TE buffer). Sample DNA from 95 individuals and the assays for the 45 sockeye salmon SNP markers (Table 2) were added to each plate using a Janus Automated Workstation (PerkinElmer, Fremont, CA). Plates were then pressurized to load the mixture into the array using a Fluidigm Integrated Fluidic Circuit Controller HX. The DNA was amplified by PCR using Fluidigm IFC thermal cyclers in 50 cycles of 92°C for 15 s and 60°C for 1 min. Endpoint reads for each Dynamic Array were performed using a Fluidigm Biomark. Each 96.96 dynamic array was scored by two separate members of the University

of Washington School of Aquatic and Fishery Sciences laboratory using the associated Fluidigm Biomark genotyping software.

Data analyses

After mtDNA sequencing, the sequence data were determined using Bioedit version 7.0.9.0 (Hall 1999). Haplotypes were analyzed using Clustal W (Thompson et al. 1994) for alignment. Haplotype diversity (h), nucleotide diversity (π), and the genetic differentiation index F_{ST} for pairwise populations were calculated with Arlequin version 3.11 (Excoffier et al. 2005). Statistical significance of haplotype frequencies was determined based on the exact test after sequential Bonferroni correction (Rice 1989). In the SNP analysis, genotype frequency conformity to Hardy–Weinberg equilibrium (HWE) in each locus and population was tested using GENEPOP version 1.2 (Raymond and Rousset 1995). Levels of significance for all tests in HWE were determined by population using sequential Bonferroni adjustments for simultaneous tests (Rice 1989). The differentiation index (F_{ST}) and genetic differentiation estimates, which were determined using the pairwise F_{ST} defined by Wright (1969), were calculated with Arlequin version 3.11 (Excoffier et al. 2005). Critical significance levels for simultaneous tests were evaluated using sequential Bonferroni adjustment (Rice 1989). Principal coordinate analysis (PCoA) was performed using GenAlEx 6.1 software (Peakall and Smouse 2006). Analysis of molecular variance (AMOVA) was performed using Arlequin version 3.11 (Excoffier et al. 2005) to examine genetic variation among sockeye salmon populations. To test for recent population bottlenecks, we used the program BOTTLENECK version 1.2.02 (Piry et al. 1999). During a bottleneck, alleles are lost from the population and levels of heterozygosity are temporarily higher than expected under mutation-drift equilibrium. Note that bottlenecks are detectable for only a few dozen generations until genetic drift and new mutations begin to reestablish mutation-drift equilibrium (Nei and Li 1976; Maruyama and Fuerst 1985; Cornuet and Luikart 1996). We assume an infinite allele model of mutation (IAM) for this analysis because SNPs were generated by distribution mutations on the simulated gene genealogies at a mutation rate of 6×10^{-5} per generation under an "infinite alleles" mutation model (Reich et al. 2001). We therefore used this model with the sign test to determine if a significantly greater proportion of loci with heterozygosity excess than expected existed for a population. The other test, which is considered more powerful and robust than the sign test (Piry et al. 1999), detects significant heterozygosity excess on average across loci using a standard difference test.

Results

151

152 mtDNA analysis

153

154 Both the Japanese and the Iliamna Lake populations possessed only one base substitution and had two haplotypes (Hap-1
155 and Hap-2) within the 520-bp sequence of the mtDNA *ND5* region. Nucleotide substitutions were pyrimidine (C–T)
156 transitions (260 bp; Hap-1: C; Hap-2: T). Within the Japanese populations, the AKA08, TOW04, TOW08, and TAC03
157 populations possessed both haplotypes, although the number of individuals that possessed Hap-2 were few. The other
158 Japanese populations, AKA04, SHI03, SHI08, and ABI94, possessed only Hap-1. The Iliamna Lake populations and the
159 lake and river spawning populations had both haplotypes, although the proportion of individuals that possessed each
160 haplotype was opposite (Fig. 2). The haplotype and nucleotide diversities of the Japanese populations were extremely
161 low compared to the Iliamna Lake populations (Table 3). The pairwise F_{ST} showed significant genetic difference between
162 the TAC03 and other Japanese populations ($P < 0.05$). A significant difference between the Japanese populations and the
163 Iliamna Lake populations was observed, but no significant difference between the Iliamna Lake populations was detected
164 (Table 4).

165

166 SNP analysis

167

168 Forty-five SNPs were amplified across populations. The Japanese populations had significantly more monomorphic loci
169 (22.3 ± 4.20) than the Iliamna Lake populations (6.50 ± 2.70; $P < 0.05$), and the heterozygotes in the Japanese
170 populations (0.10 ± 0.01) were also significantly smaller than that of the Iliamna populations (0.23 ± 0.01; $P < 0.05$; Fig.
171 3). Despite the low genetic diversity, all Japanese populations had SNP loci deviated from HWE, whereas the Iliamuna
172 population (IliL03) had only one locus deviated from HWE (Table 5). Within the Japanese populations, SHI08 and
173 SHI04 had six (One_zP3b, One_STC-410, One_GPDH2, One_U508-533, One_U502-167, and One_MARCKS-241) and
174 one (One_VIM-569) private alleles, respectively. The other populations did not have a private allele. Although the
175 Japanese populations had lower allelic diversity than the Iliamna Lake populations, the Japanese population had three
176 specific genotypes in three loci (One_IL8r_362, One_serpin, and One_Ots213_181). In these loci, the Japanese
177 populations had characteristic genotypes. For example, at One_IL8r_362, the Japanese populations had a "TT" genotype;
178 however, both river and lake spawning populations of the Iliamna Lake population had only "CC" or "CT" and did not
179 possess a "TT" genotype. The same situation was observed in the other loci (Fig. 4). The PCoA for both the Japanese and
180 the Iliamna Lake populations showed that genetic variations were expected as 93.01% on the first principal coordinate

and only 2.96% on the second principal coordinate (Fig. 5A). The PCoA for only Japanese populations showed 52.21% on the first principal coordinate and 19.88% on the second principal coordinate. According to the PCoA results, Japanese populations were categorized into three groups: the Akan Lake and Shikotsu Lake populations, the Towada Lake and Abira River populations, and the Tachibana Lake population (Fig. 5B). Variance components calculated by AMOVA for these three groups showed that 1.89% of the variation was accounted for among groups and 0.64% of the variation among populations ($P < 0.05$; Table 7).

Discussion

In this study, both Japanese and the Iliamna Lake sockeye salmon populations had only two haplotypes in the *ND5* mtDNA region. Anadromous sockeye salmon in south central Alaska and eastern Kamchatka possessed variable sites in the *ND5*/*ND6* mtDNA regions (Churikov et al. 2001; Brykov et al. 2003). The low genetic variability of the *ND5* region used in this study suggests that this region may be of limited use in discriminating among populations of sockeye salmon. Despite the low variability in this region, the Japanese populations had lower haplotype and nucleotide diversities than the Iliamna Lake populations, which indicates that Japanese lacustrine sockeye salmon populations are genetically homogeneous compared to the Iliamna Lake population. The Shikotsu Lake and the Abira River populations had only Hap-1. The Shikotsu Lake population experienced a sharp decline from 1985 to 1988; their effective population size was less than 100 (Kaeriyama 1991). Lacustrine sockeye salmon with Hap-2 in Shikotsu Lake may have been on the edge of extinction during this period. Transplants of the lacustrine sockeye salmon from Shikotsu Lake to the Abira River were carried out after 1985. The Abira River population did not possess Hap-2. This result also supports this hypothesis. According to pairwise F_{ST} comparisons within the Japanese populations, the Tachibana Lake population (TAC03) was significantly different from all Japanese populations. After the transplants from Shikotsu Lake in 1911–1912, the Tachibana Lake population was closed to further transplants from other Japanese populations and reproduced inside it (M. Kaeriyama unpublished data). Evidence on appearance in the Tachibana Lake population and absence in the Shikotsu Lake and the Abira River sockeye salmon of Hap-2 suggests that Hap-2 may be one of the genetic characteristics of native Japanese populations.

The SNP analysis showed that the Japanese lacustrine sockeye salmon populations had lower allelic and heterozygous diversity than the Iliamna Lake populations. All Japanese populations had SNP loci deviated from the HWE despite the low genetic diversity (Table 5). These results correspond to previous studies (Winans and Urawa 2000; Beacham et al. 2006). This suggests that the bottleneck effect and gene flow occurred to all Japanese populations.

211 However, the BOTTLENECK evaluated only ABI94 and TAC03 as the bottleneck effect, despite low allelic
212 heterozygositic diversity of SNP in all populations. Therefore, this BOTTLENECK could not estimate a bottleneck effect
213 for Japanese populations which had few SNP polymorphic loci. In spite of very low genetic diversity, the Japanese
214 populations had unique genotypes (One_IL8r_362, One_serpin, and One_Ots213_181) compared to the Iliamna Lake
215 populations. These genotypes did not appear in some of the Alaskan sockeye salmon populations (unpublished data, J.
216 Seeb). We may distinguish the Japanese populations from other populations using these genotypes. Why the "XX"
217 genotype did not appear despite the existence of the "XY" genotype in some loci of sockeye salmon populations is still
218 uncertain. For example, the locus of One_IL8r_362 in Iliamna Lake populations did not possess the "TT" genotype.
219 Genetic linkage, especially an association between the "X" genotype and viral disease loci, for example, may cause a
220 nonappearance of "XX" (Moen et al. 2008). To solve this problem, we need to perform whole-genome linkage
221 disequilibrium mapping of common disease genes (Kruglyak 1999).

222 According to a PCoA result including both the Japanese and Iliamna Lake populations, almost no genetic
223 differences exist among the Japanese populations because the second principal coordinate showed only 3.0% and the
224 Japanese populations showed a concentrated distribution (Fig. 5A). Namely, there are no genetic differentiations among
225 Japanese populations. This evidence suggests that the transplants of Urumobetsu Lake anadromous population would not
226 necessarily affect to the Shikotsu Lake population. The carrying capacity of Shikotsu Lake is constantly about 7,200
227 individuals and 12,000 at a maximum (Kaeriyama 1991). The first overpopulation occurred during the early 1920s
228 because of over-release of hatchery-derived juvenile into the lake (Kaeriyama 1993). Subsequently in Shikotsu Lake,
229 many anadoromous sockeye salmon eggs were transplanted from Urumobetsu Lake to Shikotsu Lake. However, this
230 anadoromous population could not reproduce in order to the density dependent effect (Rogers et al. 1980; Kaeriyama
231 1998) under the low carrying capacity. In addition, the high rate of mortality of anadromous and nonanadromous hybrids
232 (Ricker 1940; Wood and Foote 1996; Craig et al. 2005) and genetic differences between sockeye and kokanee formed as
233 a result of multiple, divergent selection pressures associated with the marine (sockeye) and lacustrine (kokanee)
234 environment (e.g., Wood and Foote 1990, 1996; Taylor and Foote 1991; Foote et al. 1992; Foote et al. 1999). This may be
235 the reason why the Urumobetsu Lake anadromous sockeye salmon population did not adapt in the lakes of Hokkaido.

236 In conclusion, this study demonstrated that Japanese lacustrine sockeye salmon have much less genetic diversity
237 than the anadromous sockeye in North America and Russia based on the bottleneck effect due to the overpopulation in
238 Shikotsu Lake and transplants of a few eggs as the effective population size to lakes. Therefore, there is a need for
239 Japanese sockeye salmon to be protected as an endangered species, and not to be disturbed by transplantations.
240

Acknowledgments We thank the following help and input with this project: Fisheries Cooperative Associations of Shikotu Lake, Towada Lake and Akan Lake for providing DNA samples; Carita Pascal, Eric Grau and Jessie Tsai (IPSEG, University of Washington, School of Aquatic and Fishery Sciences) for help with laboratory analysis. We also thank the Maeda Ippoenn foundation for admitting to Akan Lake. This study was supported by a Grant–in–Aid for Scientific Research (B) from the Ministry of Education Culture, Sports, Science and Technology, Japan (license number 21380113), and The Gordon and Betty Moore Foundation.

References

Beacham TD, McIntosh B, MacConnachie C, Miller KM, Withler RE (2006) Pacific rim population structure of sockeye salmon as determined from microsatellite analysis. Trans Am Fish Soc 135:174–187.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol 18:249–256.

Brykov VA, Polyakova NE, Podlesnykh AV (2003) Divergence of mitochondrial DNA in populations of sockeye salmon *Oncorhynchus nerka* Walbaum from Azabach'e Lake (Kamuchatka). Russian J Genet 39:1432–1437.

Burgner RL (1991) Life history of sockeye salmon (*Oncorhynchus nerka*). In: Groot C, Margolis L (eds) Pacific salmon life histories. University of British Columbia Press, Vancouver. pp. 3–117.

Churikov D, Matsuoka M, Luan X, Gray AK, Brykov VA, Gharrett AJ (2001) Assessment of concordance among genealogical reconstructions from various mtDNA segments in three species of Pacific salmon (genus *Oncorhynchus*). Mol Ecol 10:2329–2339.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent demographic bottlenecks from allele frequency data. Genetics 144:2001–2014.

Craig JK, Foote CJ, Wood CC (2005) Countergradient variation in carotenoid use between sympatric morphs of sockeye salmon (*Oncorhynchus nerka*) exposes nonanadromous hybrids in the wild by their mismatched spawning colour. Biol J Linn Soc 84:287–305.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Foote CJ, Larkin PA (1988) The role of male choice in the assortative mating of sockeye salmon and kokanee, the anadromous and nonanadromous forms of *Oncorhynchus nerka*. Behaviour 106:43–61.

278

279 Foote CJ, Wood CC, Clarke C, Blackburn J (1992) Circannual cycle of seawater adaptability in *Oncorhynchus nerka*:
280 genetic differences between sympatric sockeye salmon and kokanee. Can J Fish Aquat Sci 49:99–109.

281

282 Foote CJ, Moore K, Stenberg K, Craig JK, Wenburg JK, Wood CC (1999) Genetic differentiation in gill raker number
283 and length in sympatric anadromous and nonanadromous morphs of sockeye salmon, *Oncorhynchus nerka*. Environ
284 Biol Fish 54:263–274.

285

286 Fujimura S (1900) Transportation of sockeye salmon (Kabacheppo). Sci Rep Chitose Salmon Hatchery. pp. 32–42. (In
287 Japanese)

288

289 Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows
290 95/98/NT. Nucleic Acids Symp Ser 41:95–98.

291

292 Kaeriyama M (1991) Dynamics of the lacustrine sockeye salmon population in Lake Shikotsu, Hokkaido. Sci Rep
293 Hokkaido Salmon Hatchery 45:1–24. (In Japanese with English abstract, tables, and figures)

294

295 Kaeriyama M (1993) Life history of sockeye salmon released from hatchery. Tech Rep Hokkaido Salmon Hatchery
296 162:57–61. (In Japanese with English abstract, tables, and figures)

297

298 Kaeriyama M (1994) Life history strategy of sockeye salmon: diversity and endemism of life history pattern. In: Goto A,
299 Tsukamoto K, Maekawa K (eds) Diadromous fishes: life history and evolution. pp. 101–113. Tokai University Press,
300 Tokyo.

301

302 Kaeriyama M (1996) Effects of population density and habitat environment on life history strategy and migration of
303 juvenile sockeye (*Oncorhynchus nerka*) and chum salmon (*O. keta*). Sci Rep Hokkaido Salmon Hatchery 50:101–111.

304

305 Kaeriyama M (1998) Dynamics of chum salmon, Oncorhynchus keta, populations release from Hokaido, Japan. NPAFC
306 Bulletin 1:90-102.

307

308 Kitanishi S, Edo K, Yamamoto T, Azuma N, Hasegawa O, Higashi S (2007) Genetic structure of masu salmon
309 (*Oncorhynchus masou*) populations in Hokkaido, northernmost Japan, inferred from mitochondrial DNA variation. J
310 Fish Biol 71:437–452.
311
312 Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Gen
313 22:139–144.
314
315 Maruyama T, Fuerst PA (1985) Population bottlenecks and non-equilibrium models in population genetics. II. Number of
316 alleles in a small population that was formed by a recent bottleneck. Genetics 111:675–689.
317
318 Ministry of the Environment, Japan (2007) Red List. (http://www.biodic.go.jp/rdb/rdb_f.html. (In Japanese)
319
320 Moen T, Hayes B, Baranski M, Berg PR, Kjøglum S, Koop BF, Davidson WS, Omholt SW, Lien S (2008) A linkage
321 map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. BMC Genetics 9:223.
322
323 Navajas MJ, Thistlewood HMA, Lagnel J, Hughes C (1998) Microsatellite sequences are underrepresented in two mite
324 genomes. Insect Mol Biol 7:249–256.
325
326 Nei M, Li W-H (1976) The transient distribution of allele frequencies under mutation pressure. Genet Res Camb
327 28:205–214.
328
329 Oshima M (1934) Life-history and distribution of the fresh-water salmons found in the waters of Japan. Proc Fifth Pac
330 Sci Congr 5:3751–3773.
331
332 Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and
333 research. Mol Ecol Notes 6:288–295.
334
335 Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the
336 effective population size using allele frequency data. J Hered 90:502–503.
337

338 Raymond M, Rousset F (1995) Genepop (version-1.2): population genetics software for exact tests and ecumenicism. J
339 Hered 86:248–249.
340
341 Refsite T (1977) Effect of density on growth and survival of rainbow trout. Aquaculture 11:329–334.
342
343 Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farthadian SF, Ward R,
344 Lander EC (2001) Linkage disequilibrium in the human genome. Nature 411:199–204.
345
346 Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225.
347
348 Ricker WE (1940) On the origin of kokanee a fresh-water type of sockeye salmon. Trans R Soc Can 34:121–125.
349
350 Rogers DE (1980) Density-dependent growth of Bristol Bay sockeye salmon. p. 267-283. In W. J. McNeil and D. C.
351 Himsworth [eds.] Salmonid ecosystems of the North Pacific. Oregon State University Press, Corvallis, OR.
352
353 Seeb LW, Crane PA (1999) Allozymes and mitochondrial DNA discriminate Asian and North America population of
354 chum salmon in mixed-stock fisheries along the south coast of the Alaska Peninsula. Trans Am Fish Soc 128:88–103.
355
356 Smith CT, Templin WD, Seeb JE, Seeb LW (2005) Single nucleotide polymorphisms provide rapid and accurate
357 estimates of the proportions of U.S. and Canadian Chinook salmon caught in Yukon River fisheries. N Am J Fish
358 Manage 25:944–953.
359
360 Taylor EB, Foote CJ (1991) Critical swimming velocities of juvenile sockeye salmon and kokanee, the anadromous and
361 nonanadromous forms of *Oncorhynchus nerka* (Walbaum). J Fish Biol 38:407–419.
362
363 Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence
364 alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res
365 22:4673–4680.
366
367 Tokui T (1964) The study of himemasu. The transplant of himemasu in Japan. Sci Rep Hokkaido Salmon Hatchery

19:49–59. (In Japanese)

Winans GA, Urawa S (2000) Allozyme variability of *Oncorhynchus nerka* in Japan. Ichthyol Res 47:343–352.

Winans GA, Aebersold PB, Urawa S, Varnavskaya NV (1994) Detecting continent of origin of chum salmon (*Oncorhynchus keta*) using genetic stock identification techniques: status of allozyme baseline in Asia. Can J Fish Aquat Sci 51:95–113.

Wood CC, Foote CJ (1990) Genetic differences in the early development and growth of sympatric sockeye salmon and kokanee (*Oncorhynchus nerka*), and their hybrids. Can J Fish Aquat Sci 47:2250–2260.

Wood CC, Foote CJ (1996) Evidence for sympatric genetic divergence of anadromous and nonanadromous morphs of sockeye salmon (*Oncorhynchus nerka*). Evolution 50:1265–1279.

Wright S (1969) Evolution and the genetics of populations, Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol 12:563–584.

389 **Table 1** Samples of sockeye salmon in the Japanese and the Iliamna lakes used in this study.

390 *The samples in the Akan Lake in 2008 were harvested in two batches, the numbers were combined.

391

Population	Sample abbreviation	Date of collection (year/month/day)	n
Japan			
Akan Lake (2004)	AKA04	2004/10/10	74
Akan Lake (2008)	AKA08	2008/10/3 & 25	100*
Shikotsu Lake (2003)	SHI03	2003/10/23	59
Shikotsu Lake (2008)	SHI08	2008/10/17	100
Towada Lake(2004)	TOW04	2004/9/29	67
Towada Lake(2008)	TOW08	2008/10/28	100
Tachibana Lake (2003)	TAC03	2003/10/25	60
Abira River (1994)	ABI94	1994/10/3	80
Alaska, USA			
Iliamna Lake Lake Spawning (2003)	ILIL03	2003/8/14	35
Iliamuna Lake River Spawning (2003)	ILIR03	2003/8/15	46

392

393

394 **Table 2** Individual SNP markers (45loci) used in this study; N: Nucleotide DNA markers, M: mtDNA markers,

395 F_{ST}: the differentiation index for each marker in all samples.

396

Marker	DNA marker	F_{ST}	Reference
One_ACBP-79	N	0.237	A
One_ALDOB-135	N	0.182	A
One_CO1	M	0.309	A
One_ctgf-301	N	0.021	A
One_Cytb_17	M	0.009	A
One_Cytb_26	M	0.300	A
One_E2	N	0.089	B
One_GHII-2461	N	0.037	A
One_GPDH	N	0.035	B
One_GPDH2	N	0.041	B
One_GPH-414	N	0.484	A
One_hcs71-220	N	0.132	A
One_HGFA	N	0.731	B
One_HpaI-436	N	0.090	A
One_HpaI-99	N	0.015	A
One_IL8r-362	N	0.097	C
One_KPNA-422	N	0.264	A
One_LEI-87	N	0.251	A
One_MARCKS-241	N	0.017	C
One_MHC2_190v2	N	0.610	A
One_MHC2_251v2	N	0.422	A
One_Ots213-181	N	0.188	A
One_p53-576	N	0.027	A
One_pIns-107	N	0.053	B
One_Prl2	N	0.298	A
One_RAG1-103	N	0.039	A
One_RAG3-93	N	N/A	A
One_RF-112	N	0.198	B
One_RF-295	N	0.010	B
One_RH2op-395	N	0.007	A
One_serpin	N	0.171	B
One_STC-410	N	0.287	A
One_STR07	N	0.294	A
One_Tf_ex10-750	N	0.220	A
One_Tf_ex3-182	N	0.026	A
One_U301-92	N	0.018	A
One_U401-224	N	0.197	C
One_U404-229	N	0.061	C
One_U502-167	N	0.007	C
One_U503-170	N	0.198	C
One_U504-141	N	0.150	C
One_U508-533	N	0.026	C
One_VIM-569	N	0.046	A
One_ZNF-61	N	0.067	C
One_zP3b	N	0.133	B

A: Elfstrom et al. (2006), B: Smith et al. (2005), C: Habicht et al. (2007)

Table 3 Haplotype (*h*) and nucleotide (π) diversities of the mtDNA *ND5* region in populations of sockeye salmon in Japan and the Iliamna Lake, Alaska.

Population	h (±SD)	π (±SD)
AKA04	0.0000	0.0000
AKA08	0.0396 ± 0.0269	0.000071 ± 0.000210
SHI03	0.0000	0.0000
SHI08	0.0000	0.0000
TOW04	0.0299 ± 0.0287	0.000053 ± 0.000182
TOW08	0.0396 ± 0.0269	0.000071 ± 0.000210
TAC03	0.2350 ± 0.0646	0.000419 ± 0.000546
ABI94	0.0000	0.0000
ILIR03	0.494 ± 0.039	0.000881 ± 0.000859
ILIL03	0.487 ± 0.035	0.000868 ± 0.000845

Table 4 Population pairwise F_{ST} estimates of sockeye salmon collected in Japan and the Iliamna Lake. The data are based on the mtDNA *ND5* data. Statistical significance of F_{ST} value is tested by the exact test after sequential Bonferroni adjustments. *P<0.05, **P<0.01, ***P<0.001

	AKA04	AKA08	SHI03	SHI08	TOW04	TOW08	TAC03	ABI95	ILIR03	ILIL03
AKA04										
AKA08	0.00308									
SHI03	0.00000	0.00558								
SHI08	0.00000	0.00000	0.00000							
TOW04	0.00150	0.01158	0.00192	0.00616						
TOW08	0.00576	0.00669	0.00254	0.01010	0.01188					
TAC03	0.13355**	0.12152***	0.11754**	0.15912***	0.08665	0.09154*				
ABI95	0.00000	0.00227	0.00000	0.00000	0.00268	0.00687*	0.13967***			
ILIR03	0.49950***	0.51406***	0.46046***	0.55504***	0.43447***	0.47578***	0.16180***	0.51354***		
ILIL03	0.66415***	0.68197***	0.63265***	0.70751***	0.61836***	0.65706***	0.38936***	0.67527***	0.05991	

Table 5 The list of SNP loci deviated from HWE basec on GENEPOP. *P<0.05, **P<0.01, ***P<0.001

Population name	Marker	P value
AKA04	One_E2	*
	One_serpin	*
	One_RF112	*
AKA08	One_GPDH	**
SHI03	One_IL8r-362	*
SHI08	One_HGFA	*
TOW04	One_Prl2	**
TOW08	One_Ots213-181	*
TAC03	One_IL8r-362	***
	One_RF-112	**
ABI94	One_IL8r-362	*
	One_GPDH	*
IliL03	One_E2	*

Table 6 Results of a sign test and standard difference test indicator for a genetic bottleneck in eight populations of lacustrine sockeye salmon in Japan. IAM stands for the infinite alleles model of mutation (Reich et al. 2001)

Population	Sigh Test (IAM)	Standard difference Test (IAM)
AKA04	0.198	0.367
AKA08	0.173	0.334
SHI03	0.326	0.143
SHI08	0.321	0.121
TOW04	0.286	0.116
TOW08	0.091	0.053
TAC03	0.049	0.048
ABI96	0.016	0.025

Table 7 Results of analysis of molecular variance among sockeye salmon populations in Japan. The Japanese populations were partitioned into three groups based on the PCOA result (Fig. 5). *P<0.05.

Source of Variation	df	Sum of squares	Variance components	Persentage of variation
Among groups	2	41.38	0.04458	1.89*
Among populations within groups	5	23.79	0.01515	0.64*
Within populations	1262	2902.96	2.30029	97.47
Total	1269	2968.12	2.36002	

*P<0.005

Figure captions

Fig. 1 (A) Sampling locations of lacustrine sockeye salmon in Japan. Historically important transplants are indicated by arrows and transplant ages are shown in parentheses. Native Japanese lacusrtine sockeye salmon lived only in Akan Lake and Chimikeppu Lake (B) Sampling locations of sockeye salmon in Iliamna Lake, Alaska, USA. River and lake spawning populations are expressed in stars and circles, respectively. Refer to Table 1 for sample names

Fig. 2 Distribution of mtDNA *ND5* haplotypes of sockeye salmon in Japan. Abbreviations are referred to Table 1

Fig. 3 Number of monomorphic SNPs (A) and observed heterozygosity (B) of sockeye salmon populations in Japan and the Iliamna Lake populations. Bar indicates the standard error of each heterozygosity. Abbreviations are referred to Table 1

Fig. 4 SNP genotypes on One_IL8r_362, One_serpin, and One_ Ots213_181 loci of the Shikotsu Lake (A) and Iliamna Lake (B) sockeye salmon populations. Red, green, and blue dots indicate 'XX,' 'XY,' and 'YY' genotypes, respectively. Black and gray dots represent a standard point and samples beyond detection because of bubbles or low density of DNA, respectively

Fig. 5 Principal coordinate analysis estimated by 45 SNP loci of sockeye salmon among Japanese and Lake Iliamna Lake populations (A), and within Japanese populations (B). Abbreviations are referred to Table 1

Fig. 1

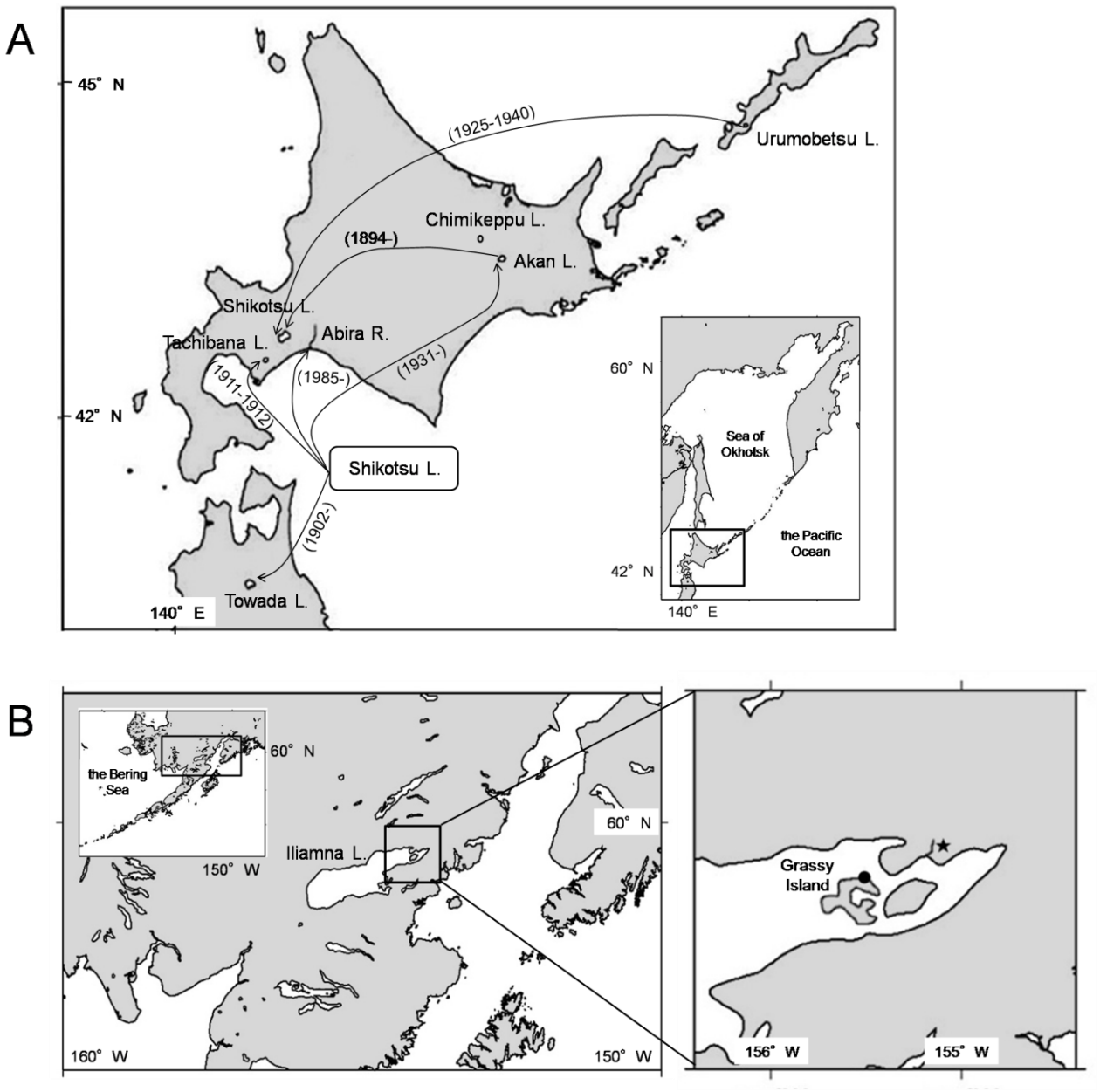

A
45° N
42° N
140° E
(1925-1940)
Urumobetsu L.
Chimikeppu L.
(1894-)
Akan L.
Shikotsu L.
Tachibana L.
Abira R.
(1911-1912)
(1985-)
(1931-)
Shikotsu L.
(1902-)
Towada L.
60° N
42° N
140° E
Sea of Okhotsk
the Pacific Ocean
B
the Bering Sea
60° N
150° W
Iliamna L.
60° N
160° W
150° W
Grassy Island
156° W
155° W

Fig. 2

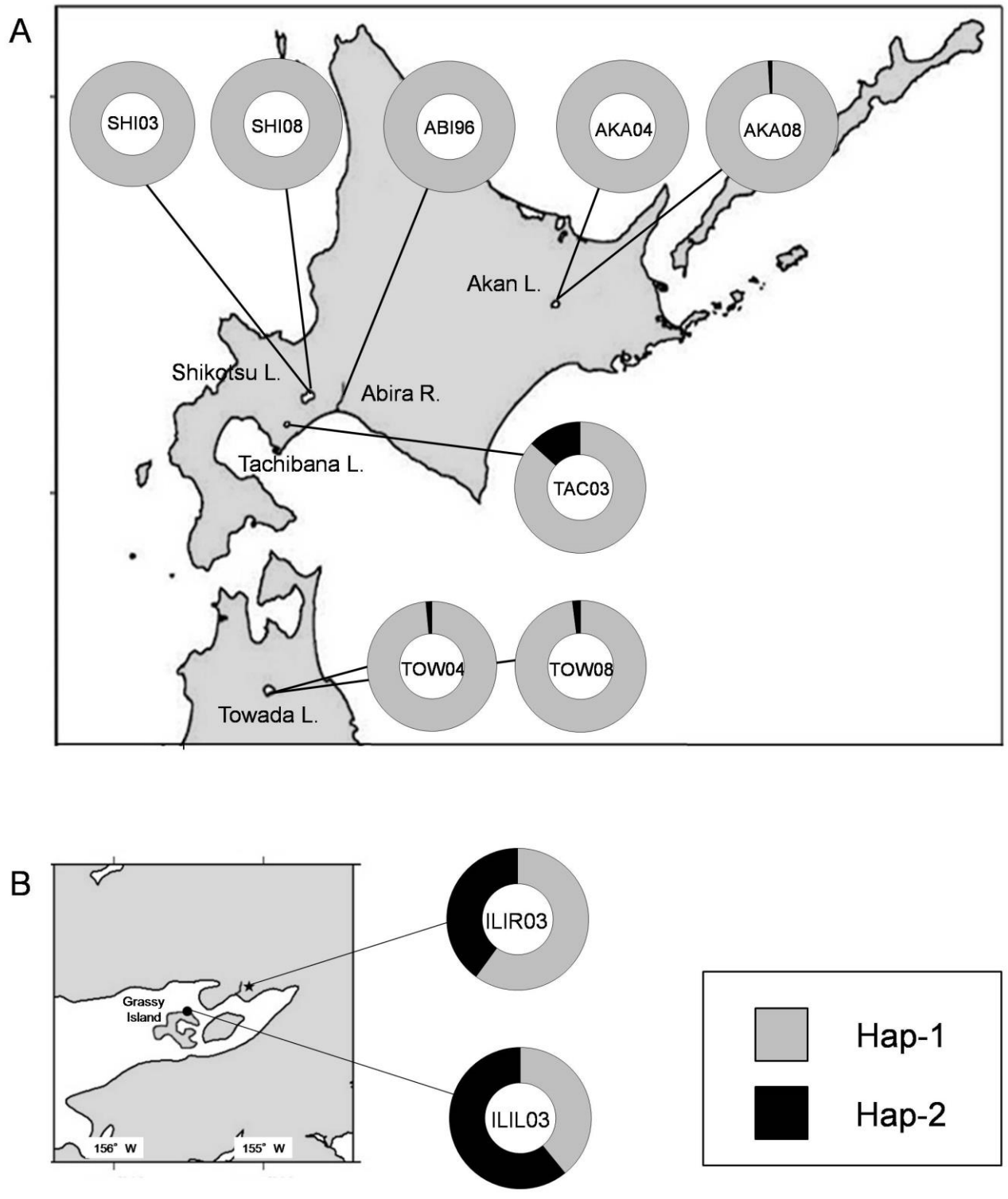

A
SHI03
SHI08
ABI96
AKA04
AKA08
Akan L.
Shikotsu L.
Abira R.
Tachibana L.
TAC03
TOW04
TOW08
Towada L.
B
Grassy Island
156° W
155° W
ILIR03
ILIL03
Hap-1
Hap-2

Fig. 3

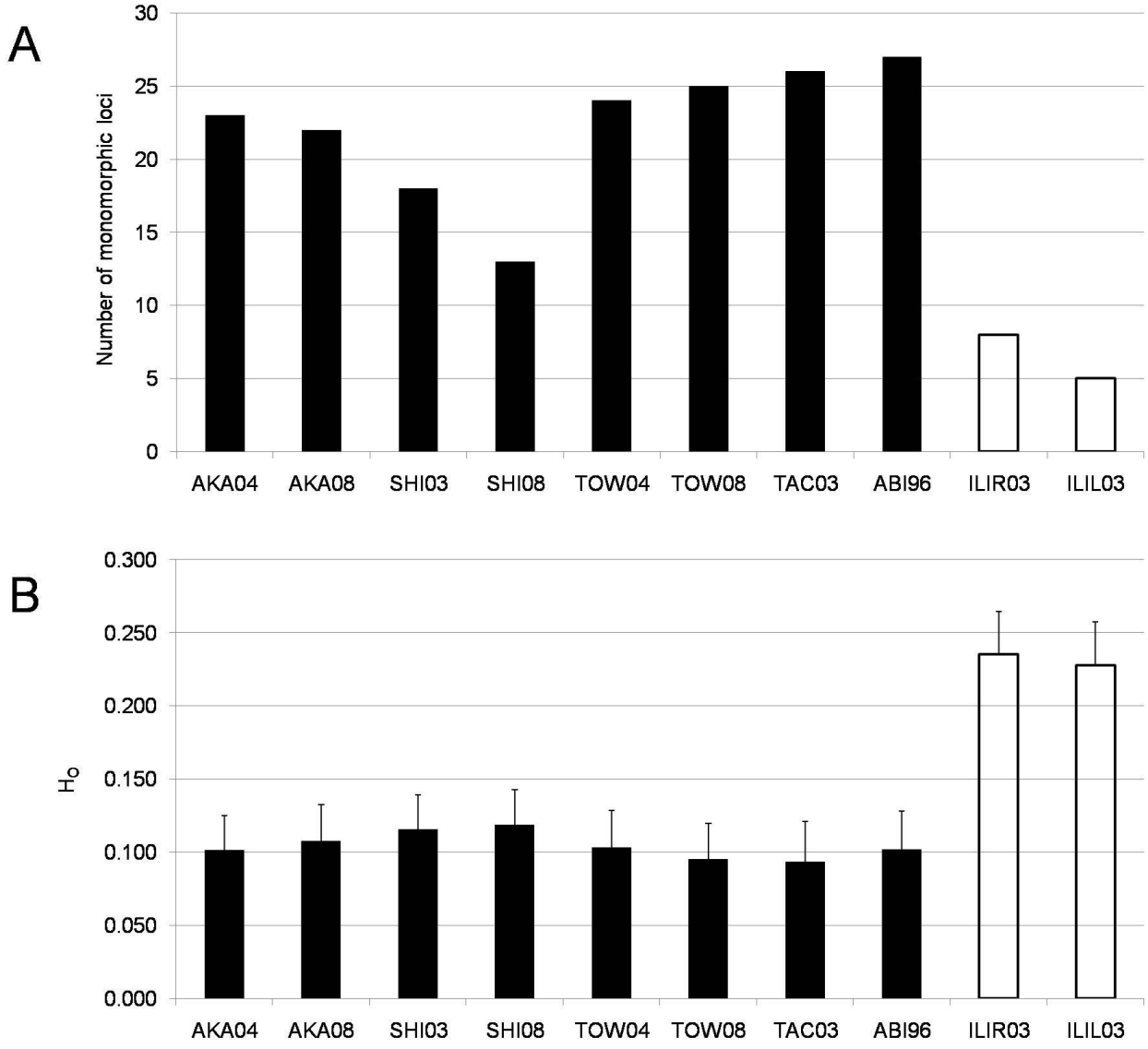

A
Number of monomorphic loci
30
25
20
15
10
5
0
AKA04
AKA08
SHI03
SHI08
TOW04
TOW08
TAC03
ABI96
ILIR03
ILIL03
B
H_O
0.300
0.250
0.200
0.150
0.100
0.050
0.000
AKA04
AKA08
SHI03
SHI08
TOW04
TOW08
TAC03
ABI96
ILIR03
ILIL03

Fig. 4

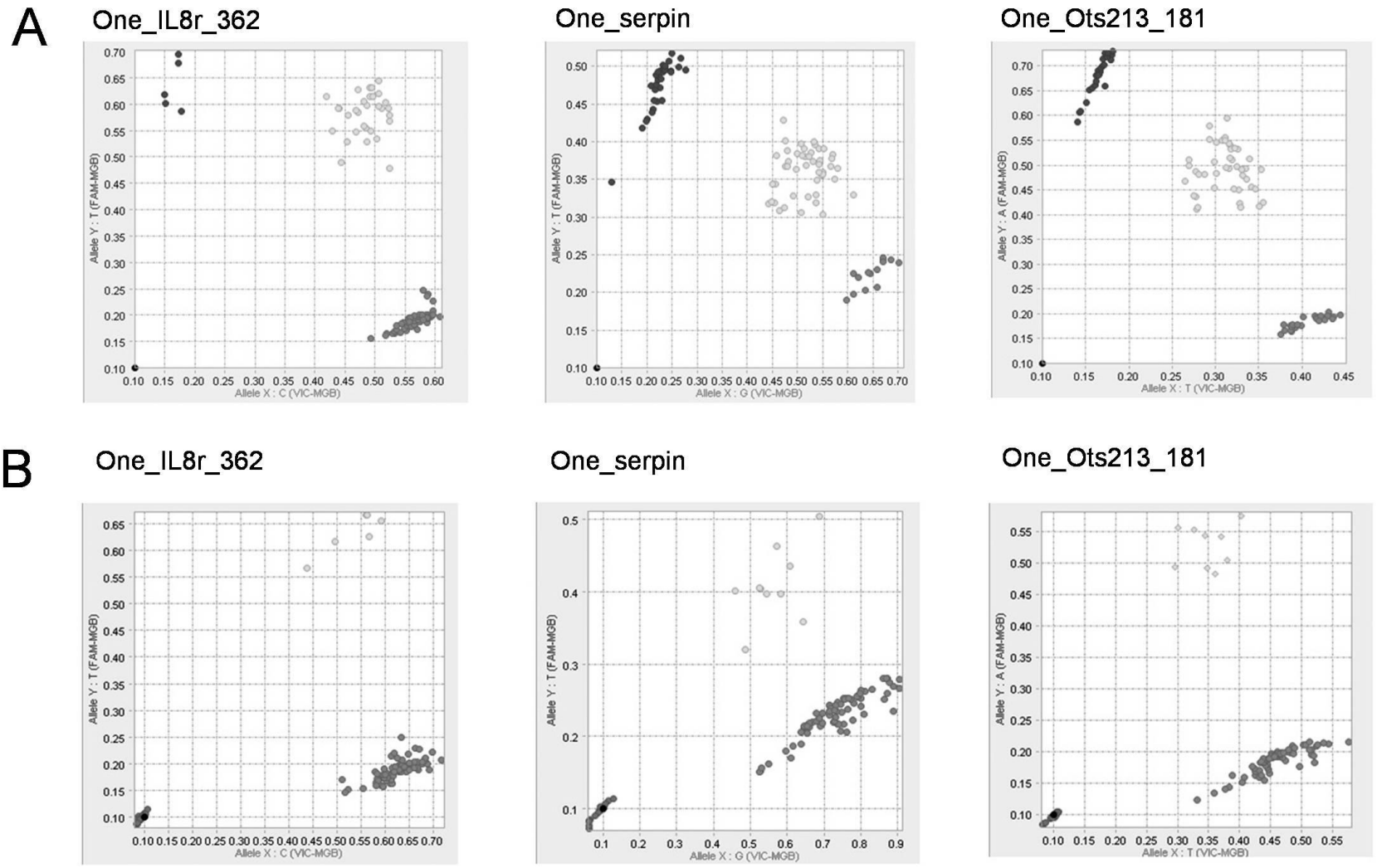
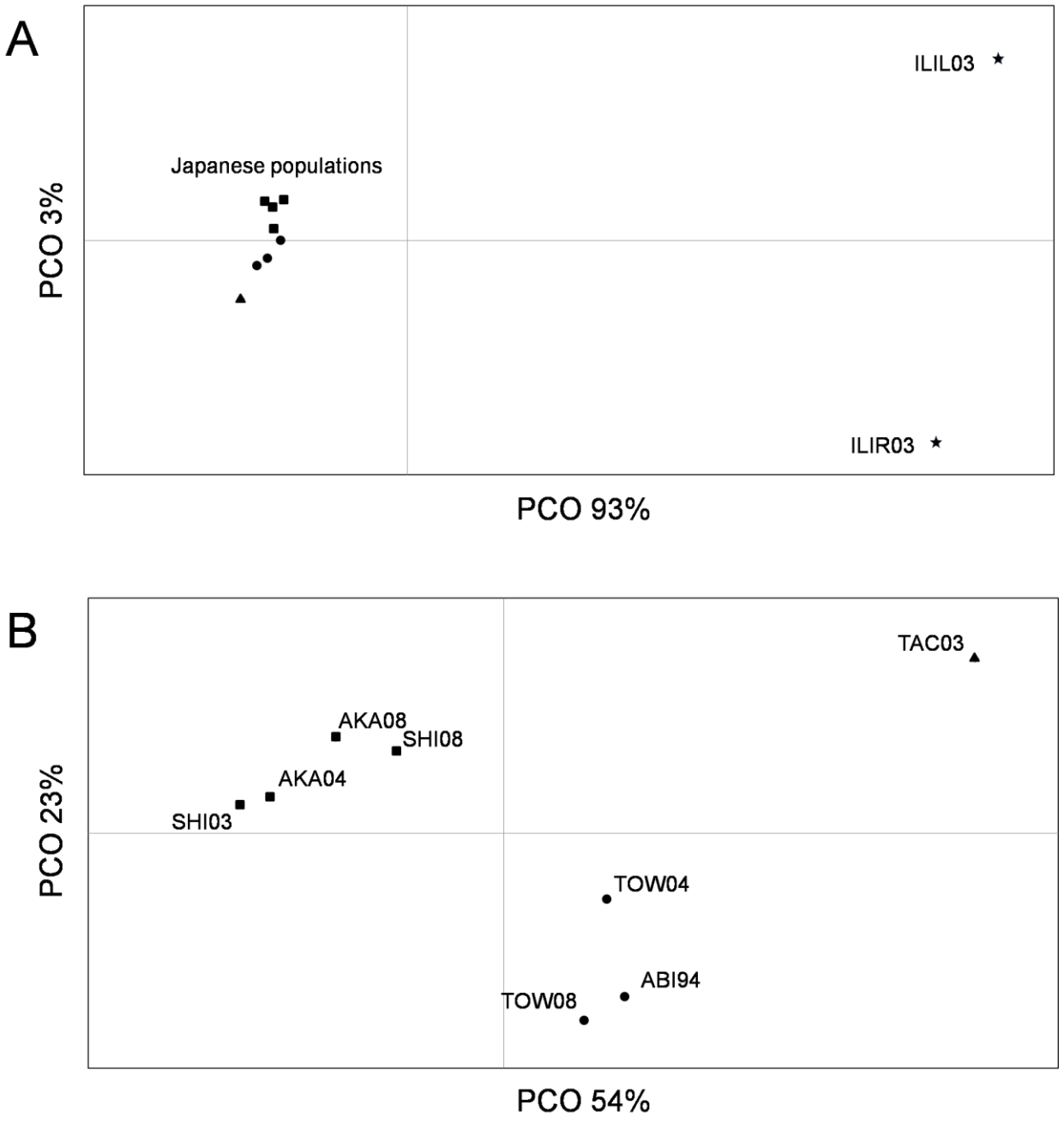

A
One_IL8r_362
Allele Y : T (FAM-MGB)
Allele X : C (VIC-MGB)
One_serpin
Allele Y : T (FAM-MGB)
Allele X : G (VIC-MGB)
One_Ots213_181
Allele Y : A (FAM-MGB)
Allele X : T (VIC-MGB)
B
One_IL8r_362
Allele Y : T (FAM-MGB)
Allele X : C (VIC-MGB)
One_serpin
Allele Y : T (FAM-MGB)
Allele X : G (VIC-MGB)
One_Ots213_181
Allele Y : A (FAM-MGB)
Allele X : T (VIC-MGB)

Fig. 5


A
ILIL03
Japanese populations
PCO 3%
ILIR03
PCO 93%
B
TAC03
AKA08
SHI08
AKA04
SHI03
TOW04
ABI94
TOW08
PCO 23%
PCO 54%