Fingerprinting of hatchery haplotypes by whole-mitogenome sequencing improves genetic studies of masu salmon *Oncorhynchus masou masou*

Short title: Fingerprinting of hatchery haplotypes by whole-mitogenome sequencing in *Oncorhynchus masou masou*

Yoko Kato-Unoki[1]*, Keitaro Umemura[2], Kosuke Tashiro[3]

[1] Center for Advanced Instrumental and Educational Supports, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

[2] Fishery Research Laboratory, Kyushu University, Fukuoka, Japan

[3] Laboratory of Molecular Gene Technology, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

* Corresponding author

E-mail: unoki.yoko.659@m.kyushu-u.ac.jp (YK)

Abstract

Stocking hatchery fish can lead to disturbance and extinction of the local indigenous population. Stocking similar lineages to the indigenous population makes it difficult to conduct genetic studies to determine conservation units. We considered that, due to the length of the mitogenome, whole-mitogenome analysis might overcome this problem by enabling identification of hatchery haplotypes. Here, to provide basic information for conservation of indigenous masu salmon *Oncorhynchus masou masou*, a commonly stocked fish in the Kase River system, Japan, we used whole-mitogenome analysis of fish to identify hatchery haplotypes in the river and in several hatcheries that might be used for stocking. Whole-mitogenome sequencing clearly identified hatchery haplotypes like fingerprints, identifying the tributaries contaminated with hatchery haplotypes. The results suggest that informal stocking of *O. m. masou* has been performed widely across the Kase River system. Non-hatchery haplotypes mainly belonged to clade I, which was not found in northern Hokkaido Island. Sites without hatchery haplotypes were estimated three, suggesting that these sites are suitable for conservation of the indigenous fish. This study demonstrated that the resolution of the

conventional partial analysis of the mitogenome is insufficient to distinguish between hatchery haplotype and indigenous haplotype. The whole-mitogenome sequences provided accurate information that was not available in the partial sequences, enabling various inferences: e.g., estimation of the origin of stocked fish and circulation of the reared strains. We conclude that whole-mitogenome analysis is useful for the genetic study of this species. The whole-mitogenome data produced will contribute to the conservation, resource management, and further study of *O. m. masou*.

Keywords: mitochondrial DNA; whole sequence; hatchery fish; indigenous; *Oncorhynchus masou masou*

Introduction

Stocking of hatchery fish has been carried out throughout the world to mitigate declines in natural production and enhance fishery production and recreational fishing [1]. However, it has become problematic because it can lead to disturbance and extinction of the local indigenous population, and loss of genetic diversity [2,3]. Such effects are observed among salmonid fishes (e.g., [4–7]). Determination of the conservation unit of the indigenous population requires genetic information to identify the indigenous individuals; however, stocking fish with a similar lineage to that of the indigenous fish makes such studies difficult.

Masu salmon *Oncorhynchus masou masou* is endemic across Japan, with diverse life histories [8]: there is a river-resident form and an anadromous form [8], with spawning adults homing to their natal river at high rates, similar to other Pacific salmons [9]. *Oncorhynchus masou masou* is one of commonly stocked fish for recreational fishing in Japan. Especially on Kyushu Island, public resource management is not sufficient in many rivers, and inappropriate and informal stocking is performed widely by individuals and/or fishing clubs. Recent research has suggested that *O. m. masou* can be divided into

64 different populations in terms of behavior, morphology, and genetics in each river [10–
65 13]. Stocking hatchery *O. m. masou* might reduce indigenous characteristics and local
66 adaptations because of hybridization between the stocked and indigenous fish. This could
67 ultimately lead to a loss of indigenous populations, as observed in other salmonid species
68 (e.g., [4,14,15]). To conserve the indigenous *O. m. masou* resource in a river, their
69 genetic information is required. Genetic study of *O. m. masou* in Japan has focused
70 mainly on northern Hokkaido Island and some areas of Honshu Island and its
71 surrounding waters (e.g., [10,13,16,17]); studies of this species in Kyushu Island are
72 scarce [18].

73 The Kase River system is located in northwestern Kyushu Island (Fig 1).
74 Inappropriate *O. m. masou* stocking, with no consideration of the indigenous population,
75 has been conducted for a long time by fishing clubs and personal game fishers in this
76 river. The distribution of this informal stocking and its impacts on indigenous resources
77 are unknown, and genetic information is required.

78

79 **Fig 1. Sampling locations and hatchery *MT* haplotype distribution detected in the**

80 **Kase River system.** The upper map shows Kyushu Island, and the lower map shows the

81 Kase River system. In the lower map, the main man-made dams (black bars with build

82 year) and waterfalls (blue bars) that may hinder the fish run are marked. Sites with prior

83 information about stocking history are show in bold letters. For details of the samples and

84 the results of whole-mitogenome sequencing, see Tables 1 and 4.

85

86 We considered that whole-mitogenome analysis might be helpful for genetic

87 studies to determine the conservation units of this indigenous fish. The mitogenome is

88 commonly used as a genetic marker for characterizing population structure and

89 identifying maternal lineages [19–21]. Recent reports suggest that whole-mitogenome

90 analysis can increase the resolution of matrilineal genetic structure patterns, especially in

91 low diversity species, and allow detailed and more accurate phylogenetic analysis (e.g.,

92 [22–25]). Thus, whole-mitogenome analysis might enable identification of hatchery

93 haplotype in populations of stocked hatchery fish that share similar lineages, and could

94 help estimate the indigenous genetic structure and lineage. Moreover, whole-mitogenome

sequencing will allow comparative phylogenetic analysis of any region of the mitogenome.

Here, to provide basic information for conservation of the indigenous *O. m. masou* in the Kase River system, we analyzed the whole-mitogenome in this species in this river system and in several hatcheries that might be used for stocking. Through this analysis we could identify hatchery haplotypes of fish in this river system and clarify their distribution. The whole-mitogenome analysis provided important information that was not available from conventional partial-mitogenome analysis.

Materials and methods

Ethics statement

This study was conducted with the permission of the Saga prefecture (permission number 3018). The fish were collected under appropriate fishing licenses that allowed the capture and sacrifice of the fish. The ethical approval of the Kyushu University Animal Experiment Committee was not required because approval is only necessary for fish reared at Kyushu University (according to the Kyushu University Animal

111 Experimentation Regulations). However, this study was carried out according to the
112 guidelines of the Ichthyological Society of Japan
113 (http://www.fish-isj.jp/english/guidelines.html).
114

115 Study area and samples

116 Information about the *O. m. masou* samples (sample numbers, sampling year, and
117 stocking history) used in this study is listed in Table 1, and the locations of sampling sites
118 are shown in Fig 1. The samples from the Kase River system were collected at 18 sites
119 from 16 tributaries (St1–St8, St9-1 and St9-2, St10–St16, and StM1) (Fig 1, Table 1).
120 Samples from Shiiba Research Forest, Kyushu University (sSRF), an unstocked site in
121 the Hitotsuse River system, were used as the control group. Information on the stocking
122 history and hatcheries used for the Kase River system were obtained by interviewing the
123 local populace, fishing clubs, and local hatcheries, similar to the approach described in
124 Nakamura [26]. As a result of these interviews, one research center (Miyazaki Prefectural
125 Fisheries Research Institute, sKFR) and six hatcheries (sHT1–sHT6) were selected as
126 putative sources of *O. m. masou* in this river system (Fig 1, Table 1): sHT1, a hatchery in

127 the Kase River system, is used for the public projects of Saga prefecture, a local
128 fishermen's cooperative, and local residents and volunteers—fish in sHT1 originate from
129 sKFR-reared wild eggs from the Ishido stream of the Hitotsuse River system; sHT2 is
130 recently used by a fishing club; sHT3 has been commonly used by fishing clubs,
131 including the same group that used sHT2; sHT4 is being used with sHT1 by a local
132 fishermen's cooperative; embryos from sHT5 can be easily purchased by individuals on
133 the internet; and sHT6 is the original hatchery used by a different fishing club than that
134 using sHT2. Information about the river of origin and mating history of hatchery fishes is
135 shown in the columns "River / tributary" and "Additional information", respectively, in
136 Table 1.
137

Table 1. Sample information used in this study.

	Site ID	**River / tributary[a]**	**Sampling year**	**Number of samples**	**Stocking history**	**Additional information**
Kase River	St1	Kase / tributary 1	2016	16	unknown	

system	St2	Kase / tributary 2	2016	21	unknown	
	St3	Kase / tributary 3	2016	5	stocked	
	St4	Kase / tributary 4	2015, 2016	23	stocked	
	St5	Kase / tributary 5	2016	21	unknown	
	St6	Kase / tributary 6	2016	12	unknown	
	St7	Kase / tributary 7	2016	8	stocked	
	St8	Kase / tributary 8	2016	6	unknown	
	St9-1	Kase / tributary 9	2016	22	unknown	
	St9-2	Kase / tributary 9	2016	11	unknown	
	St10	Kase / tributary 10	2016	26	unknown	
	St11	Kase / tributary 11	2016	22	unknown	
	St12	Kase / tributary 12	2016	20	unknown	
	St13	Kase / tributary 13	2016	16	unknown	
	St14	Kase / tributary 14	2016	11	unknown	
	St15	Kase / tributary 15	2016	11	stocked	
	St16	Kase / tributary 16	2016	11	unknown	

	StM1	Kase / tributary 16	2016	24	stocked	
Control	sSRF	Hitotsuse	2016	20	unstocked	prohibited fishing area
Hatchery	sKFR	Hitotsuse / Ishido	2016	7	-	from native fish
etc.	sHT1	Hitotsuse / Ishido	2016	11	-	from sKFR
	sHT2	Kita	2016	15	-	from native fish
	sHT3	Gokase	2016	15	-	mated with native fish and others
	sHT4	-	2016	10	-	mated with many strain
	sHT5	Ooyodo / Okimizu	2016	10	-	from native fish
	sHT6	Gokase / Kawabashiri	2016	9	-	mated with native fish and others

[a]For each hatchery, the original river of the reared fish is given.

138

139 # Sampling and DNA extraction

140 Wild fish were captured by electrofishing or normal fishing from 2015 to 2016. A small
141 piece of fin was taken from each specimen, and the captured fish were released at the
142 same point in the river. Tissues were immediately fixed in 99.5% ethanol or RNA*later*
143 solution (Thermo Fisher Scientific) and stored at −20 °C until use. Total genomic DNA

144 was extracted from preserved tissue using the QIAGEN Blood & Tissue DNeasy Kit

145 (QIAGEN).

146

147 DNA sequence and genetic analysis

148 Sequencing was performed in two steps: first, the NADH dehydrogenase subunit 5 gene

149 (*ND5*) sequences of all samples (383 individuals) were analyzed; next, some of the

150 samples (88 individuals) were selected to represent each site and each *ND5* haplotype,

151 and the whole-mitogenome sequences of these samples were analyzed.

152 The *ND5* region (1,597 bp) was amplified with the primer pair ND5-F1 and ND5-R

153 (Table 2). The PCR mixture was as follows: 10–50 ng genomic DNA, 1× Phusion HF

154 buffer (New England BioLabs), 0.2 mM of each dNTP, 0.5 μM of each primer, and 0.1

155 μL Phusion DNA polymerase (New England BioLabs) in a total volume of 10 μL. The

156 PCR program was as follows: one cycle at 98 °C for 30 s, followed by 40 cycles of 98 °C

157 for 10 s, 64 °C for 20 s, and 72 °C for 60 s, and a final extension at 72 °C for 2 min. Each

158 PCR product was sequenced with the primer ND5-F1 or ND5-F2 (Table 2). The *ND5*

159 sequences were assembled and multiple sequences were aligned using ATGC Ver. 4.3.5

160 software (GENETYX Co.). The identified 1,449 bp sequence variants (*ND5* haplotypes)

161 (*ND5* position, 358–1806; mitogenome position, 13299–14747 for reference sequence

162 NC_008747) were deposited in DDBJ/EMBL/GenBank as shown in S1 Table.

163

Table 2. List of primers used in this study.

Name	Sequence (5'–3')	Primer position[a]	
		mitogenome	*ND5*
ND5-F1[b]	TACCCCAATTGCCCTGTACG	13223–13242	282–301
ND5-F2	TCAGGCTCAATTATCCACAG	13986–14005	1045–1064
ND5-R[b]	ACTAACACGTGGGTTAGGTCGAG	14797–14819	(3' NCR)[c]
MT-F	AATTATCCACAGTTTAAATGACGAACAAG	13994–14022	1053–1081
MT-R	AAAAGTATAGCTTTAAAGAATGCGTGAGT	13950–13978	1009–1037

[a] The primer is positioned in the reference sequence (NC_008747). [b] Reference taken from Kitanishi et al. [10]; however, ND5-R varied slightly from this reference. [c] 3' terminal non-coding region of *ND5*

164

165

166 The whole-mitogenome fragment (16.6 kbp) was amplified with primer pair MT-F
167 and MT-R (Table 2). The PCR mixture was as follows: 20–100 ng genomic DNA, 1×
168 PrimeSTAR GXL Buffer (TaKaRa Bio Inc.), 0.2 mM each dNTP, 0.2 μM each primer,
169 and 0.3 μL of PrimeSTAR GXL DNA polymerase (TaKaRa Bio Inc.) in a total volume
170 of 15 μL. The PCR program was as follows: one cycle at 98 °C for 10 s, followed by 40
171 cycles of 98 °C for 10 s, 60 °C for 15 s, 68 °C for 14 min, and a final extension at 68 °C
172 for 5 min. Each PCR product was subjected to agarose gel electrophoresis, and the
173 mitogenome was extracted from the excised band by using the Wizard SV Gel and PCR
174 Clean-up System (Promega).

175 Mitogenome libraries for next-generation sequencing (NGS) were prepared using a
176 QIAseq FX DNA Library Kit (QIAGEN). For each library, the quality and fragment size
177 were checked using an Agilent BioAnalyzer and the concentration was quantified by
178 qPCR (Mx3000p, Agilent) using a KAPA LQ Kit (KAPA Biosystems). Pooled and
179 denatured library (8 pM) containing 5% volume of PhiX (control library; Illumina) was

180 sequenced using the Illumina MiSeq system with MiSeq Reagent Kit V3 (300-bp
181 paired-end reads) (Illumina).

182 The obtained NGS data of paired end reads were trimmed using Trimmomatic ver.
183 0.36 as follows: a cleanup adapter was applied and reads with low quality (Q score, <28)
184 and short-length (<50) were filtered out; then, the reads were mapped to the reference
185 mitogenome sequence (*O. m. masou* accession No: NC_008747) using Burrow–Wheeler
186 Aligner ver. 0.7.12, and the obtained SAM (Sequence Alignment Map) files being
187 converted to BAM (Binary Alignment Map) files using SAMtools ver. 1.4.1. The
188 resulting reads and the single nucleotide polymorphism (SNP) sites for the reference
189 sequence were visualized with Integrative Genomics Viewer version 2.3.83. and
190 TASSEL ver. 5.0. The primer sites and those outside the region of interest (mitogenome
191 position: 13950–14022) were replaced with the predetermined *ND5* sequence. The
192 identified full-length mitogenome sequence variants (*MT* haplotypes) were deposited in
193 DDBJ/EMBL/GenBank as shown in S1 Table.

194 The relationships among haplotypes were estimated using the TCS program
195 implemented in PopART [27]. To determine the relationships between haplotypes in past

196 studies and the current study, we also analyzed the haplotype network including data of

197 Kitanishi et al. [10] and Yu et al. [16].

198

199 Results

200 DNA sequence and determination of haplotype

201 The sequencing analysis of the 1,449 bp fragment containing the ND5 gene detected a

202 total of 18 *ND5* haplotypes with 28 SNP sites from 383 individuals (Tables 3 and S2). Of

203 these, 10 *ND5* haplotypes (KS4, KS6, KS8–KS10, KS12, KS14, HT3, HT4A, and HT4B)

204 were detected in hatcheries, and seven of these, KS4, KS6, KS8–KS10, KS12, and KS14,

205 were detected in the Kase River system. KS4 was also detected in the unstocked area,

206 sSRF, in the Hitotsuse River system.

207

Table 3. Distribution of *ND5* haplotypes in each sampling site.

		KS1	KS2	KS3	KS4*	KS5	KS6*	KS7	KS8*	KS9*	KS10*	KS11	KS12*	KS13	KS14*	KS15	HT3*	HT4A*	HT4B*	Total (n)
Kase River system	**St1**	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0				16
	St2	0	0	2	0	0	0	0	19	0	0	0	0	0	0	0				21
	St3	0	0	2	1	0	0	0	0	0	2	0	0	0	0	0				5
	St4	5	12	0	0	0	0	0	0	0	0	0	0	0	0	6				23
	St5	2	14	0	5	0	0	0	0	0	0	0	0	0	0	0				21
	St6	0	0	0	7	0	0	0	1	3	0	0	1	0	0	0				12
	St7	0	0	0	7	0	0	0	0	0	1	0	0	0	0	0				8
	St8	0	0	1	1	0	0	0	4	0	0	0	0	0	0	0				6
	St9-1	0	0	1	0	0	2	0	19	0	0	0	0	0	0	0				22
	St9-2	0	0	2	2	0	0	0	1	0	0	4	2	0	0	0				11
	St10	0	0	7	19	0	0	0	0	0	0	0	0	0	0	0				26
	St11	0	0	8	14	0	0	0	0	0	0	0	0	0	0	0				22
	St12	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0				20
	St13	0	0	3	0	0	0	0	0	0	1	0	0	5	0	7				16
	St14	0	0	0	0	0	0	2	0	0	0	0	0	0	9	0				11
	St15	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0				11
	St16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	10				11
	StM1	0	0	2	0	1	0	8	2	2	0	0	0	1	8	0				24
	Total (n)	7	26	44	56	21	2	10	46	16	4	4	3	6	18	23				286
Control	**sSRF**	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Hatchery etc.	**sKFR**	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
	sHT1	0	0	0	6	0	0	0	0	5	0	0	0	0	0	0	0	0	0	11
	sHT2	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	15
	sHT3	0	0	0	0	0	3	0	1	4	2	0	3	0	1	0	1	0	0	15
	sHT4	0	0	0	0	0	1	0	2	0	1	0	0	0	0	0	1	1	4	10
	sHT5	0	0	0	0	0	0	0	1	5	2	0	0	0	0	0	2	0	0	10
	sHT6	0	0	0	0	0	0	0	2	0	0	0	0	0	4	0	1	2	0	9
	Total (n)	0	0	0	13	0	4	0	6	14	20	0	3	0	5	0	5	3	4	77

Asterisked haplotypes indicate the hatchery haplotype in *ND5* haplotype.

Eighty-eight of the 383 individuals that underwent *ND5* sequencing were selected for whole-genome sequencing by NGS. The whole-mitogenome sequencing of all 88 individuals was successful with approximately >1500-fold depth of coverage. The

213 number of samples with each determined *MT* haplotype are shown for each sampling site
214 in Table 4. The majority of *ND5* haplotypes were associated with more than one *MT*
215 haplotype, with 32 haplotypes including 256 SNP sites being detected (Tables 4 and S2).
216 Sixteen haplotypes were detected in hatcheries (see asterisks, Table 4), and six of these,
217 mtKS4_1, mtKS8, mtKS9, mtKS10_1, mtKS12, and mtKS14_2, were detected in the
218 Kase River system. Hatchery haplotypes were distributed in 12 tributaries (13 sites)
219 including sites with unknown stocking history (Fig 1). Several hatchery haplotypes were
220 detected upstream of artificial barriers (e.g., dams) and/or natural barriers (e.g.,
221 waterfalls): mtKS4_1 at St11; mtKS8 at St2 and St9-1; mtKS9 at St15; and mtKS10_1 at
222 St13 (Fig 1). The whole-mitogenome sequences of KS6 in the Kase River system
223 (mtKS6_1) and KS4 in sSRF (mtKS4_2) did not match any hatchery *MT* haplotype in the
224 samples analyzed here (Table 4). Two *MT* haplotypes, mtKS4_1 and mtKS9, were
225 detected in hatchery sHT1, which was stocked with fish from sKFR; one of these *MT*
226 haplotypes, mtKS4_1, was the same as that detected in sKFR (Tables 1 and 4). Analyzed
227 MT haplotype in the ND5 haplotype KS4 samples in the Kase River system was
228 consistently mtKS4_1 except a sample at St8, mtKS4_2.

Table 4. Result of whole-mitogenome sequencing.

		KS 1	KS 2		KS 3			KS 4*			KS 5	KS 6*		KS 7	KS 8*	KS 9*	KS 10*			KS 11	KS 12*	KS 13		KS 14*				KS 15	HT3*			HT 4A*	HT 4B*	
		mtKS1	mtKS2_1	mtKS2_2	mtKS3_1	mtKS3_2	mtKS3_3	mtKS4_1*	mtKS4_2	mtKS4_HIT	mtKS5	mtKS6_1	mtKS6_HT3*	mtKS7	mtKS8*	mtKS9*	mtKS10_1*	mtKS10_HT2*	mtKS10_HT3*	mtKS11	mtKS12*	mtKS13_1	mtKS13_2	mtKS14_1	mtKS14_2*	mtKS14_HT3*	mtKS14_HT6*	mtKS15	mtHT3_1*	mtHT3_HT4*	mtHT3_HT6*	mtHT4A*	mtHT4B*	Total (n)
Kase River system	St1				4																														4
	St2					1									1																				2
	St3				1			1									1																		3
	St4	1	1																									1							3
	St5	1		3				1																											5
	St6							1							1	1					1														4
	St7							1									1																		2
	St8				1				1						1																				3
	St9-2							1												1	1														3
	St9-1				1							1			2																				4
	St10				1			2																											3
	St11						2	1																											3
	St12										4																								4
	St13						1										1					1						1							4
	St14													1										1											2
	St15															1																			1
	St16																								1			1							2
	StM1										1			1	1	1							1		1										6
Control	sSRF									1																									1
Hatchery	sKFR							1																											1
	sHT1							1								1																			2
	sHT2																	1																	1
	sHT3												1		1	1			1		1					1			1						7
	sHT4												1		1		1													1		1	1	6
	sHT5															3			2																5
	sHT6														1										1		3				1	1		7
***MT* haplotype (n)**		2	1	3	8	1	3	10	1	1	5	1	2	2	9	8	4	1	3	1	3	1	1	1	3	1	3	3	1	1	1	2	1	88
***ND5* haplotype (n)**		2	4		12			12			5	3		2	9	8	8			1	3	2		8				3	3			2	1	88

Number in each cell shows analyzed sample number. Asterisks indicate the hatchery haplotypes.

Some hatchery haplotypes were detected at multiple hatcheries (Table 4): mtKS6_HT3 was detected at sHT3 and sHT4; mtKS8 was detected at sHT3, sHT4, and sHT6; mtKS9 was detected at sHT1, sHT3, and sHT5; mtKS10_HT3 was detected at sHT3 and sHT5; and mtHT4A was detected at sHT4 and sHT6. On the other hand, six hatchery haplotypes (mtKS10_HT2, mtKS14_HT3, mtHT3_1, mtHT3_HT4, mtHT3_HT6, and mtHT4B) were not detected in any of the samples from the Kase River system and other hatcheries.

239

240 ## Haplotype network analysis

241 We analyzed the relationships among haplotypes by using the TCS program (see
242 Materials and Methods). The results showed a network of *ND5* haplotypes
243 (*ND5*-haplotype network) comprising four clades (Fig 2A). Analysis of *MT* haplotype
244 relationships showed a detailed network (*MT*-haplotype network) (Fig 2B). Clades I–III
245 in the *ND5*-haplotype network were similar to those in the *MT*-haplotype network. Clade
246 IV of the *ND5*-haplotype network included other *O. masou* subspecies and formed a
247 star-like topology radiating from haplotype KS14, whereas clade IV of the *MT*-haplotype
248 network was divided into several sub-clades. In both the *ND5*- and *MT*-haplotype
249 networks, all haplotypes in clade III were hatchery haplotypes. In clade I of the
250 *MT*-haplotype network, mtKS4_1, which was detected at sKFR (original river: Ishido
251 tributary in the Hitotsuse River system), was in the same branch as mtKS4_HIT, which
252 was detected at sSRF (original river: other tributary in the Hitotsuse River system).

253

254 **Fig 2. TCS network trees of *ND5* haplotypes and *MT* haplotypes obtained in this**

255 **study.** TCS network trees of (A) *ND5* haplotypes and (B) *MT* haplotypes are shown.

256 Each dash represents one single nucleotide difference between two neighboring

257 haplotypes. Reference information (abbreviations and sequence accession numbers) are

258 as follows: OMM, *O. m. masou* (NC_008747); OMI, *O. m. ishikawae* (DQ_864464);

259 OMF, *O. m. formosanus* (DQ_858456); OMB, *O. m.* Biwa subsp. (Biwa salmon)

260 (EF_105342).

261

262 In the *ND5*-haplotypes network, the non-hatchery haplotypes were mostly in clade I

263 (70%), followed by clade IV (23%) and clade II (7%) (from the data in Table 3). We

264 compared the haplotypes and cladal relationships obtained in this study with those

265 obtained using sequencing data of a 561 bp region in *ND5* in the reports by Kitanishi et al.

266 [10] and Yu et al. [16] (S3 Table and S1 Fig). KS1, KS2, and KS13, were new haplotypes

267 in this region of *ND5* in the current study. Clades I, II, III, and IV in the current study

268 appear to correspond to clade 1-3, 1-6, 2-3, and 2-1, reported by Yu et al. [16],

269 respectively.

Discussion

Identification of hatchery haplotypes by whole-mitogenome analysis

The whole-mitogenome sequencing clearly identified hatchery haplotypes in *O. m. masou*, like fingerprints. Tributaries contaminated with fish with hatchery haplotypes were clarified, suggesting that informal stocking was widely done across the study area. In pre-information, five tributaries (five sites) had a stocking history (Table 1; bold letters, Fig 1), however, hatchery haplotypes were detected in 12 tributaries (13 sites) (Fig 1). Hatchery haplotypes were also detected upstream of waterfalls and artificial barriers that were built before stocking become popular in the Kase River system in the early 1970s (Fig 1; local residents pers. comm.). *Oncorhynchus masou masou* has a high homing tendency for the natal stream, and the dispersal of anadromous females and river-resident individuals is low [28–30]. Hence, we consider that hatchery haplotypes in each tributary were derived from artificially stocked fish rather than natural dispersal.

285 This study suggests that the resolution of partial mitogenome sequences is
286 insufficient to distinguish between hatchery haplotypes and indigenous haplotypes in
287 similar lineages. For example, the *ND5* haplotype detected in the prohibited fishing area
288 sSRF was identical to that detected at the hatcheries sHT1 and sKFR (KS4; Table 3),
289 however, the *MT* haplotype detected at sSRF (i.e., mtKS4_HIT) and that detected at the
290 above hatcheries (i.e., mtKS4_1) differed by four nucleotides (Tables 4 and S2). Without
291 whole-mitogenome analysis, we could not get the accurate understanding of the result
292 indicating that the samples in the prohibited fishing area showed a hatchery *ND5*
293 haplotype.

294 Our results suggest that the *MT* haplotype mtKS4_HIT might be an indigenous
295 haplotype in sSRF for the following reasons. As mentioned above, the mtKS4_HIT
296 haplotype detected in sSRF differed by only four nucleotides from the mtKS4_1
297 haplotype found in the sHT1 hatchery and sKFR, and they diverged from the same node
298 in the *MT*-haplotype network (Fig 2B). Fish in the sHT1 hatchery are known to originate
299 from sKFR, which in turn sourced fish from the Ishido tributary of the Hitotsuse River
300 system. The site sSRF is in another tributary of the same Hitotsuse River system (Table

301 1). Therefore it is possible that fish with these two haplotypes diverged from a common
302 ancestor in the Hitotsuse River system and accumulated genetic mutations at each site
303 due to their homing ability and/or being landlocked [30,31]. These results also indicate
304 that the whole-mitogenome analysis can detect a subtle divergence of the mitogenome
305 among tributaries.
306

307 Other findings provided by whole-mitogenome analysis

308 The whole-mitogenome analysis provided accurate information that was not available by
309 conventional partial analysis of the mitogenome, enabling various inferences to be made.
310 For instance, it allowed estimation of the origin of stocked fish. Partial analysis showed
311 that the *ND5* haplotype KS4 was detected widely and frequently across the Kase River
312 system (Table 3). However, with whole-mitogenome analysis, fish with this haplotype
313 could be divided into three *MT* haplotypes, mtKS4_1, mtKS4_2, and mtKS4_HIT, and
314 these *MT* haplotype of analyzed samples in the Kase River system were all mtKS4_1
315 except a sample at St8, mtKS4_2 (Table 4). Hence, we could infer that the haplotype
316 mtKS4_1 in the Kase River system was derived from sKFR and was spread from the

317 hatchery sHT1 on the Kase River system, because this *MT* haplotype was detected at both
318 sHT1 and its original sKFR, consistent with our pre-information (Fig 1, Table 1).
319 The current study clarified that a reared strain is shared by multiple hatcheries and
320 is stocked at multiple sites: e.g., hatchery *MT* haplotype mtKS8 was detected at three
321 hatcheries (sHT3, sHT4, and sHT6) and four tributaries (five sites: St2, St6, St8, St9-1,
322 and StM1) (Fig 1, Table 4). Hatcheries sHT2–sHT6 are used by fishing clubs and/or
323 many personal people and are likely used to stock various rivers (fishing clubs and game
324 fishers pers. comm.). It is of concern that stocking of these rivers could cause loss of the
325 unique gene sequences of the indigenous fish, and lead to loss of genetic diversity in this
326 species in the river and/or tributaries [2,14,32].
327 Non-hatchery haplotypes habitats are estimated to three sites. Urgent conservation
328 might be required. Indigenous fish might be remained in the tributaries in which we did
329 not detect a hatchery haplotype, and no hatchery haplotypes were detected four tributaries
330 (4 sites) in the Kase River system (sites St1, St4, St12 and St14; Fig 1 and Table 4).
331 However, the mtKS15 haplotype detected at St4 could be a stocked hatchery haplotype
332 that could not be detected in this study, because (a) St4 has a stocking record (Table 1),

333 (b) mtKS15 was also detected from the sites, St13 and St16, which are geographically
334 distant from St4, as observed with other hatchery MT haplotypes. (Fig 1, Table 4).
335 Indigenous fish in the Kase River system may be revealed in further studies by
336 microsatellite analysis of the nucleic genome [21,33]. Addition, this suspicious haplotype
337 might be resolved by accumulating whole-mitogenome sequence data of hatcheries and
338 field samples from several other areas in Japan in future studies.
339

340 Future applications of whole-mitogenome analysis

341 The SNP data obtained here for *O. m. masou* (S2 Table) may be used for resource
342 management such as monitoring of hatchery fishes and/or putative indigenous fish by
343 testing environmental DNA [34,35].
344 The whole-mitogenome analysis will help estimate the indigenous haplotype,
345 because the indigenous haplotypes must be among the non-hatchery haplotypes.
346 Estimation of the indigenous haplotype might lead to research into the geographical and
347 phylogenetic evolution of this species. In this study, the proportion of non-hatchery
348 haplotypes was highest in clade I (70%), possibly suggesting that clade I is the main

349 population of indigenous *O. m. masou* in the Kase River system. As shown in S1 Fig, we
350 obtained three new haplotypes, KS1, KS2, and KS13, in the *ND5* 561 bp region in this
351 study. KS1 and KS2 are not hatchery haplotypes and belong to clade I (Fig 2). The only
352 haplotype corresponding to clade I in previous studies is H22 (S3 Table and S1 Fig) [16],
353 however this haplotype was not detected in northern Hokkaido Island and was detected at
354 a low rate in Honshu Island and in Korea [10,16]. These results may suggest that the
355 clade composition differs between the Kase River system on Kyushu Island and northern
356 Japan. A recent study using a cytochrome-*b* gene (1141 bp) in *O. masou* in the
357 northwestern Pacific including southern Kyushu Island suggested the possible existence
358 of a specific clade comprised of only *O. m. masou* in Kyushu Island (Group D in [18]).
359 Further study might be necessary to assess whether our clade I haplotypes are related to
360 Group D haplotypes in that report. Accumulating the whole-mitogenome data for this
361 species throughout its area of distribution will might allow for a detailed phylogeographic
362 analysis and may reveal the history of geographic evolution in this species.
363
364

Acknowledgments

The authors are grateful to Dr. Norio Onikura of the Faculty of Agriculture of Kyushu University for supporting this study; Dr. Yukiko Ogino of the Faculty of Agriculture of Kyushu University for help with English writing; Dr. Takamasa Kawamura of the Faculty of Agriculture of Kyushu University for technical support of NGS; Daisuke Kambayashi of Miyazaki Prefectural Fisheries Research Institute and staff at the Shiiba Research Forest of Kyushu University for providing samples; and the Fisheries Division, River & Sediment Control Division, and Forestry Preservation Division of Saga Prefecture for providing information. NGS was performed using an Illumina MiSeq at the Center for Advanced Instrumental and Educational Support, Faculty of Agriculture, Kyushu University. This work was supported by the Japan Society for the Promotion of Science KAKENHI Grant Number 16H00474.

Author Contributions

Conceptualization: YK.

Formal analysis: YK, KU, KT.

Funding acquisition: YK.

Investigation: YK, KU, KT.

Methodology: YK.

Project administration: YK.

Resources (NGS reagents): KT.

Supervision: YK.

Writing – original draft: YK.

Writing – review & editing: KU, KT.

References

1. Lever C. Naturalized Fishes of the World. London: Academic Press; 1996. Available: https://books.google.co.jp/books?id=Q5QWAQAAIAAJ
2. Hindar K, Ryman N, Utter F. Genetic Effects of Cultured Fish on Natural Fish Populations. Can J Fish Aquat Sci. 1991;48: 945–957.
3. Rhymer JM, Simberloff D. Extinction by hybridization and introgression. Annu Rev Ecol Syst. 1996;27: 83–109. doi:10.1146/annurev.ecolsys.27.1.83
4. Araki H, Berejikian BA, Ford MJ, Blouin MS. Fitness of hatchery-reared salmonids in the wild. Evol Appl. 2008;1: 342–355. doi:10.1111/j.1752-4571.2008.00026.x
5. Altukhov YP, Salmenkova EA. Stock transfer relative to natural organization, management, and conservation of fish populations. In: Ryman N, Utter FM, editors. Population genetics and fishery management. Seattle: University of Washington Press; 1987. pp. 333–343.

6. Miller LM, Close T, Kapuscinski AR. Lower fitness of hatchery and hybrid rainbow trout compared to naturalized populations in Lake Superior tributaries. Mol Ecol. 2004;13: 3379–3388. doi:10.1111/j.1365-294X.2004.02347.x

7. Gharrett AJ, Smoker WW, Reisenbichler RR, Taylor SG. Outbreeding depression in hybrids between odd- and even-broodyear pink salmon. Aquaculture. 1999;173: 117–129. doi:https://doi.org/10.1016/S0044-8486(98)00480-3

8. Kato F. Life histories of masu and amago salmon (*Oncorhynchus masou* and *Oncorhynchus rhodurus*). In: Groot C, Margolis L, editors. Pacific Salmon Life Histories. Vancouver: University of British Columbia Press; 1991. pp. 396–414. Available: http://ci.nii.ac.jp/naid/10008268078/en/

9. Ueda H. Sensory mechanisms of natal stream imprinting and homing in *Oncorhynchus* spp. J Fish Biol. 2019;95: 293–303. doi:10.1111/jfb.13775

10. Kitanishi S, Edo K, Yamamoto T, Azuma N, Hasegawa O, Higashi S. Genetic structure of masu salmon (*Oncorhynchus masou*) populations in Hokkaido, northernmost Japan, inferred from mitochondrial DNA variation. J Fish Biol. 2007;71: 437–452. doi:10.1111/j.1095-8649.2007.01689.x

11. Mayama H. *Oncorhynchus masou masou* ecological notes (in Japanese). Fish Egg. 1990;159: 7–21. Available: http://salmon.fra.affrc.go.jp/kankobutu/tech_repo/fe02/fishandegg159_p07-21.pdf

12. Noguchi D, Taniguch N. Studies on the genetic diversity of wild populations by microsatellite of masu *Oncorhynchus* DNA markers (in Japanese with English summary). Aquac Sci. 2007;55: 521–527. doi:10.10233/aquaculturesci1953.55.521

13. Kitanishi S, Yamamoto T, Urabe H, Shimoda K. Hierarchical genetic structure of native masu salmon populations in Hokkaido, Japan. Environ Biol Fishes. 2018;101: 699–710. doi:10.1007/s10641-018-0730-6

14. Finnegan AK, Stevens JR. Assessing the long-term genetic impact of historical stocking events on contemporary populations of Atlantic salmon, *Salmo salar*. Fish Manag Ecol. 2008;15: 315–326. doi:10.1111/j.1365-2400.2008.00616.x

15. Hansen MM, Fraser DJ, Meier K, Mensberg KLD. Sixty years of anthropogenic pressure: A spatio-temporal genetic analysis of brown trout populations subject to

stocking and population declines. Mol Ecol. 2009;18: 2549–2562. doi:10.1111/j.1365-294X.2009.04198.x

16. Yu J, Azuma N, Yoon M, Brykov V, Urawa S, Nagata M, et al. Population genetic structure and phylogeography of masu salmon (*Oncorhynchus masou masou*) inferred from mitochondrial and microsatellite DNA analyses. Zoolog Sci. 2010;27: 375–385. doi:10.2108/zsj.27.375

17. Yamamoto S, Morita K, Kikko T, Kawamura K, Sato S, Gwo J-C. Phylogeography of a salmonid fish, masu salmon *Oncorhynchus masou* subspecies-complex, with disjunct distributions across the temperate northern Pacific. Freshw Biol. 2019; 1–18. doi:10.1111/fwb.13460

18. Iwatsuki Y, Ineno T, Tanaka F, Tanaka K. The southernmost population of *Onchorhynchus masou masou* from Kyushu Island, Japan and gross genetic structure of the *O. masou* complex from the northwestern Pacific. In: Gwo J-C, Shieh Y-T, Burridge CP, editors. Proceedings of International Symposium on the 100th Anniversary of the discovery of Formosa landlocked salmon Bull Natl Taiwan Museum. National Taiwan Museum; 2019. pp. 101–119.

450 19. Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution.
451 Nature. 1987;325: 31–36. doi:10.1038/325031a0
452 20. Forgacs D, Wallen RL, Dobson LK, Derr JN. Mitochondrial Genome Analysis
453 Reveals Historical Lineages in Yellowstone Bison. PLoS One. 2016;11: e0166081.
454 Available: https://doi.org/10.1371/journal.pone.0166081
455 21. Fraser DJ, Bernatchez L. Adaptive evolutionary conservation: Towards a unified
456 concept for defining conservation units. Mol Ecol. 2001;10: 2741–2752.
457 doi:10.1046/j.1365-294X.2001.t01-1-01411.x
458 22. Miya M, Nishida M. Use of mitogenomic information in teleostean molecular
459 phylogenetics: a tree-based exploration under the maximum-parsimony optimality
460 criterion. Mol Phylogenet Evol. 2000;17: 437–455.
461 23. Shamblin BM, Bjorndal KA, Bolten AB, Hillis-Starr ZM, Lundgren I,
462 Naro-Maciel E, et al. Mitogenomic sequences better resolve stock structure of
463 southern Greater Caribbean green turtle rookeries. Mol Ecol. 2012;21: 2330–2340.
464 doi:10.1111/j.1365-294X.2012.05530.x

24. Feutry P, Kyne PM, Pillans RD, Chen X, Marthick JR, Morgan DL, et al. Whole mitogenome sequencing refines population structure of the Critically Endangered sawfish Pristis pristis. Mar Ecol Prog Ser. 2015;533: 237–244. doi:10.3354/meps11354

25. Bishop CR, Hughes JM, Schmidt DJ. Mitogenomic analysis of the Australian lungfish (*Neoceratodus forsteri*) reveals structuring of indigenous riverine populations and late Pleistocene movement between drainage basins. Conserv Genet. 2017;19: 587–597. doi:10.1007/s10592-017-1034-7

26. Nakamura T. Estimation of the distribution of genetically pure populations of the Japanese charr by inquiring survey. J Jpn Soc Eros Control. 2001;53: 3–9. doi:10.11475/sabo1973.53.5_3

27. Leigh JW, Bryant D. POPART: Full-feature software for haplotype network construction. Methods Ecol Evol. 2015;6: 1110–1116. doi:10.1111/2041-210X.12410

28. Miyakoshi Y, Takahashi M, Ohkuma K, Urabe H, Shimoda K, Kawamuka H. Homing of masu salmon in the tributaries of the Shiribetsu River evaluated by

returns of marked fish (in Japanese with English summary). Sci Rep Hokkaido Fish Res Inst. 2012;81: 125–129. Available: https://www.hro.or.jp/list/fisheries/marine/att/81-miyakoshi1.pdf

29. Ueda H. Physiological mechanisms of imprinting and homing migration in Pacific salmon *Oncorhynchus* spp. J Fish Biol. 2012;81: 543–558. doi:10.1111/j.1095-8649.2012.03354.x

30. Kitanishi S, Yamamoto T, Koizumi I, Dunham JB, Higashi S. Fine scale relationships between sex, life history, and dispersal of masu salmon. Ecol Evol. 2012;2: 920–929. doi:10.1002/ece3.228

31. Kitanishi S, Yamamoto T, Higashi S. Microsatellite variation reveals fine-scale genetic structure of masu salmon, *Oncorhynchus masou*, within the Atsuta River. Ecol Freshw Fish. 2009;18: 65–71. doi:10.1111/j.1600-0633.2008.00325.x

32. Lodge DM. Biological Invasions - lessons for ecology. Trends Ecol Evol. 1993;8: 133–137. doi:10.1016/0169-5347(93)90025-K

495 33. Toews DPL, Brelsford A. The biogeography of mitochondrial and nuclear
496 discordance in animals. Mol Ecol. 2012;21: 3907–3930.
497 doi:10.1111/j.1365-294X.2012.05664.x
498 34. Rees HC, Maddison BC, Middleditch DJ, Patmore JRM, Gough KC. REVIEW
499 The detection of aquatic animal species using environmental DNA - a review of
500 eDNA as a survey tool in ecology. J Appl Ecol. 2014;51: 1450–1459.
501 doi:10.1111/1365-2664.12306
502 35. Tsuji S, Maruyama A, Miya M, Ushio M, Sato H, Minamoto T, et al.
503 Environmental DNA analysis shows high potential as a tool for estimating
504 intraspecific genetic diversity in a wild fish population. bioRxiv. 2019; 1–24.
505 doi:http://dx.doi.org/10.1101/829770
506
507

508 # Supporting information captions

509 **S1 Table. List of GenBank accession numbers deposited in this study**.

510

511 **S2 Table. List of *MT* haplotype SNP sequences for the reference sequence (*O. m.***

512 ***masou* NC_008747).**

513

514 **S3 Table. Collation of haplotypes determined in past studies and this study.**

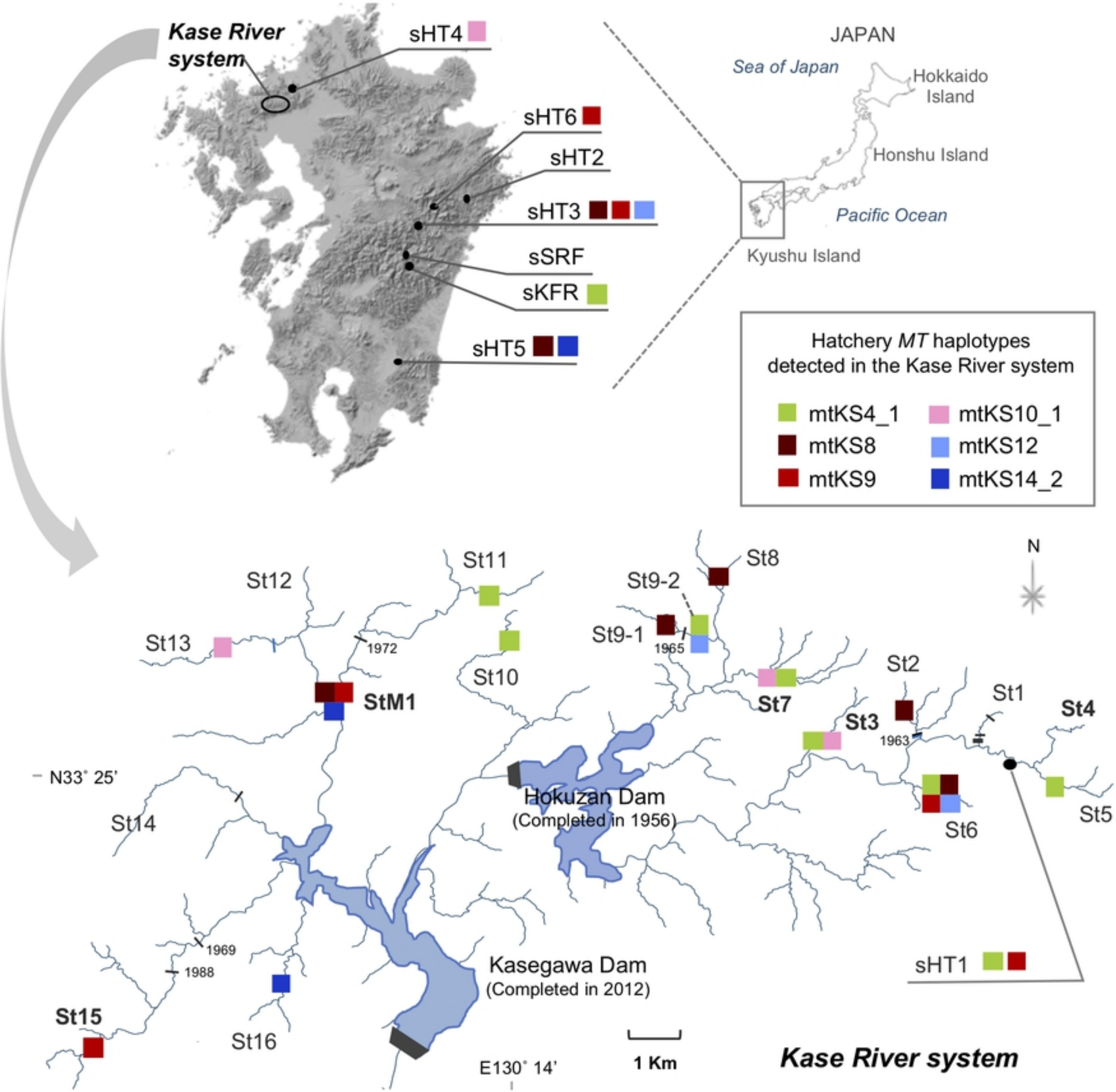
515

516 **S1 Fig. TCS network trees showing the relationship between data from past studies**

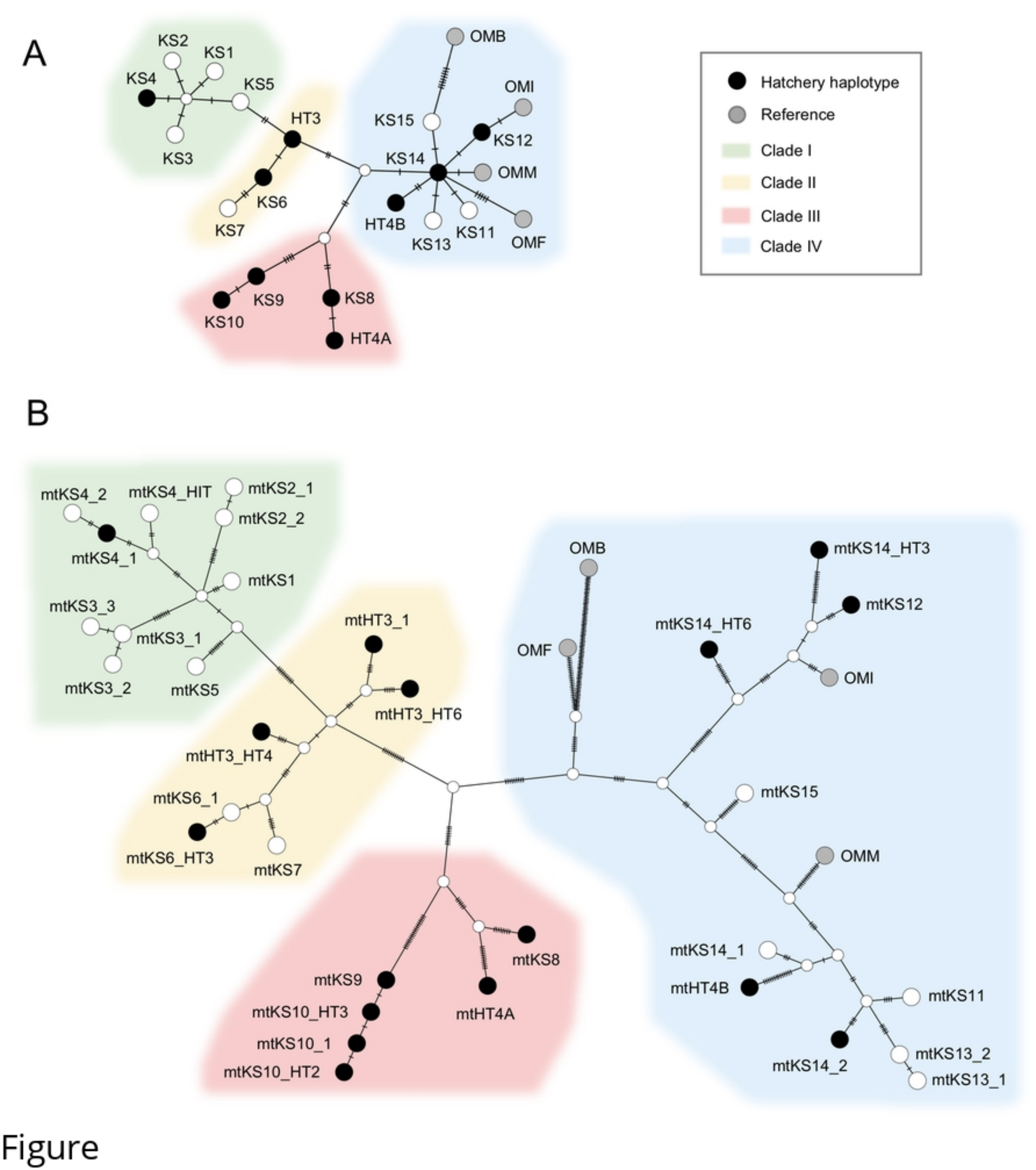
517 **and the present study.** Double circled haplotypes are new haplotypes obtained by this

518 study. Colors show the clades corresponded to this study in Fig 2. For details showing the

519 collation of haplotypes of past studies and this study see S3 Table.


Kase River system
sHT4
sHT6
sHT2
sHT3
sSRF
sKFR
sHT5
JAPAN
Sea of Japan
Hokkaido Island
Honshu Island
Pacific Ocean
Kyushu Island
Hatchery MT haplotypes detected in the Kase River system
mtKS4_1
mtKS10_1
mtKS8
mtKS12
mtKS9
mtKS14_2
N
St12
St11
St8
St9-2
St13
1972
St9-1
1965
St10
St2
StM1
St7
St1
St4
St3
1963
N33° 25'
Hokuzan Dam
(Completed in 1956)
St14
St6
St5
1969
1988
Kasegawa Dam
(Completed in 2012)
sHT1
St15
St16
E130° 14'
1 Km
Kase River system

Figure


A
KS2
KS1
KS4
KS5
KS3
HT3
KS6
KS7
OMB
OMI
KS15
KS12
KS14
OMM
HT4B
KS13
KS11
OMF
KS9
KS8
KS10
HT4A
Hatchery haplotype
Reference
Clade I
Clade II
Clade III
Clade IV
B
mtKS4_2
mtKS4_HIT
mtKS2_1
mtKS2_2
mtKS4_1
mtKS1
mtKS3_3
mtKS3_1
mtKS3_2
mtKS5
mtHT3_1
mtHT3_HT6
mtHT3_HT4
mtKS6_1
mtKS6_HT3
mtKS7
OMB
OMF
mtKS14_HT3
mtKS12
mtKS14_HT6
OMI
mtKS15
OMM
mtKS14_1
mtHT4B
mtKS11
mtKS14_2
mtKS13_2
mtKS13_1
mtKS8
mtKS9
mtKS10_HT3
mtHT4A
mtKS10_1
mtKS10_HT2

Figure