

Large-scale hybridization of Japanese populations of Hinamoroko, *Aphyocypris chinensis*, with *A. kikuchii* introduced from Taiwan

**Katsutoshi Watanabe¹ · Ryoichi Tabata² · Jun Nakajima³ · Midori Kobayakawa⁴ ·
Masanari Matsuda² · Kosuke Takaku⁵ · Kazumi Hosoya⁶ · Kenichi Ohara⁷ · Motohiro
Takagi⁸ · Nian-Hong Jang-Liaw⁹**

✉ Katsutoshi Watanabe

watanak@terra.zool.kyoto-u.ac.jp

¹ Department of Zoology, Division of Biological Sciences, Graduate School of Science, Kyoto University, Kitashirakawa-oiwake, Sakyo, Kyoto 606-8502, Japan

² Lake Biwa Museum, 1091 Oroshimo, Kusatsu 525-0001, Japan

³ Fukuoka Institute of Health and Environmental Sciences, 39 Mukaizano, Dazaifu, Fukuoka 818-0135, Japan

⁴ Midorigaoka, Higashi, Fukuoka 813-0021, Japan

⁵ Japan Wildlife Research Center, 3-3-7 Kotobashi, Sumida, Tokyo 130-8606, Japan

⁶ Kindai University, 3327-204 Nakamachi, Nara 631-8505, Japan

⁷ Gifu Prefectural Research Institute for Fisheries and Aquatic Environment,

Kawashima-kasada, Kakamigahara, Gifu 501-6021, Japan

⁸ South Ehime Fisheries Research Center, Ehime University, 1289-1 Funakoshi, Ainan, Ehime
798-4292, Japan

⁹ Conservation and Research Center, Taipei Zoo, No. 30, Sec. 2, Xinguang Road, Wenshan
District, Taipei 11656, Taiwan

Type of paper: Full paper

Running title: Genetic disturbance in Japanese *Aphyocypris chinensis*

Number of text pages: 25; Number of figures: 6; Number of tables: 2; ESM figures: 2; ESM
tables: 2

Abstract

Aphyocypris chinensis is a small cyprinid that is widely distributed in lowland areas of continental China, the Korean Peninsula, and the northwestern part of Kyushu, Japan. However, this species is severely threatened in Japan due to human impacts; thus, several facilities and citizen groups implemented captive breeding and reintroduction/reinforcement programs using several dozen founder fish collected from Tanushimaru, the location with the only known natural population, in the 1980s and 1990s. To determine the phylogenetic position and genetic authenticity of Japanese populations of *A. chinensis*, we conducted mtDNA and genome-wide SNP analyses using a total of 254 specimens from 31 wild and captive samples, including specimens of Chinese and Korean *A. chinensis* and the close relative *A. kikuchii* from Taiwan. The mtDNA divergence and phylogeny indicated that the haplotypes found exclusively in Japan were differentiated from the Chinese/Korean haplotypes (uncorrected $p = 2.6\%$ in the cytochrome *b* gene) to the same extent as they were from *A. kikuchii* haplotypes ($p = 2.9\%$). Results from mtDNA sequences and 47–359 SNPs, obtained using the MIG-seq method with different parameter sets, revealed that the initial captive populations and an extinct wild population in Ukiha, adjacent to Tanushimaru, were genuine Japanese populations, whereas all extant captive and wild populations in Japan are hybrids between Japanese *A. chinensis* and *A. kikuchii*, or *A. kikuchii* itself. The details of the captive breeding and exchange programs, as well as evidence for the aquarium trade of *A. kikuchii* since the 1990s, strongly suggest that the mixture of *A. kikuchii* into Japanese populations occurred within the first several years of the establishment of captive populations in 1994. The present case of the highly probable extinction of genuine Japanese populations of *A. chinensis* emphasizes the importance of confirmation and management of the genetic authenticity of conservation-targeted species.

Keywords Captive breeding · Endangered species · Ex situ conservation · Hybridization ·

MIG-seq · mtDNA cytochrome *b*

Introduction

Ex situ conservation efforts are increasingly required in conservation programs for endangered species that cannot survive in current natural environments and are awaiting future reintroduction into restored habitats (United Nations 1992; Frankham et al. 2002; IUCN/SSC 2014). The main purpose of ex situ conservation, or captive breeding programs, is to avoid the extinction of target species by increasing the number of individuals in captivity, with ‘freezing’ evolution of the species (Frankham et al. 2002). Evolution, i.e., genetic changes over generations, in captive populations consists of loss of genetic diversity and adaptation to captivity. Because genetic changes occurring in captivity reduce the ability of captive populations to reproduce and survive when returned to the wild, such changes must be minimized during captive breeding. To this end, one effective measure is to manage a population in medium-sized fragmented subpopulations with occasional exchange of individuals (e.g., Margan et al. 1998).

Aphyocypris chinensis, “Hinamoroko” in Japanese, is a small (<70 mm total length) cyprinid fish that is among the most threatened species in Japan (Takaku et al. 2007; Kobayakawa 2016). This species is widely distributed in lowland areas of continental China (from the Amur to the Changjiang river basins) and the Korean Peninsula, whereas the distribution range in Japan is limited to the area around the Ariake Sea and Hakata Bay on northwestern Kyushu Island (Kimura 1997; Takaku et al. 2007). This species was nearly extinct before 1980 due to loss of habitat and farmland improvement, including changes in irrigation management (Takaku et al. 2007; Kobayakawa 2015). In 1982, a small number of *A. chinensis* individuals were found in Tanushimaru, Fukuoka (an agricultural ditch near the Kose River, the Chikugo River system), and about five specimens were introduced into captivity for breeding (Akiyama 1988; Maehata 1997; Ohara and Takagi 2005). Several

dozen individuals were rediscovered at the same site in 1994, and several institutions, aquaria, and citizen groups began captive breeding and reintroduction/reinforcement activities using these fish (Kimura 1997; Ohara and Takagi 2005). Aside from observations of other wild fish during a short period in the latter half of the 2000s in Ukiha, near Tanushimaru, all extant wild and captive fish are considered to originate from the captive population founded in 1994 (and partially in 1982) (Ohara and Takagi 2005; Kobayakawa 2016).

From the initial stage of the captive breeding of *A. chinensis*, genetic management efforts involving the exchange of individuals among facilities have been partially implemented (Ohara and Takagi 2005). However, the low genetic diversity of major captive populations of *A. chinensis* in Japan was documented in the early 2000s (Ohara et al. 2003; also see Ohara and Takagi 2005, 2007). Ohara and colleagues have conducted several genetic assessments of this species using allozymes (Ohara et al. 2003), mitochondrial DNA (mtDNA) RFLPs (Ohara and Takagi 2005), and microsatellite markers (Ohara and Takagi 2007). These authors found that some populations have lost genetic diversity, whereas others, especially those tended by a citizen group, have maintained high genetic diversity. Based on their genetic assessment, Ohara and Takagi (2005, 2007) suggested that exchanges of individuals among breeding groups should effectively prevent inbreeding so that the breeding scheme could yield an effective population size of 500 or more (Soulé 1980; Lehmkuhl 1984). To date, the management of captive populations using the exchange of individuals among facilities has been implemented by the Biodiversity Committee of the Japanese Association of Zoos and Aquariums (JAZA), administrative organizations, and a citizen group.

In addition to genetic diversity, the phylogenetic position and population structure of endangered species are essential information for the determination of conservation units and priorities, as well as for the development of strategies of reintroduction and reinforcement of captive individuals to the wild (Frankham et al. 2002). To determine the phylogenetic position

of Japanese populations of *A. chinensis*, Nakajima et al. (2012) conducted a preliminary analysis using partial mtDNA sequence data and unexpectedly discovered two divergent mtDNA lineages among Japanese specimens. On the other hand, because living *A. chinensis* (or similar species) from unknown localities were traded widely among Japanese aquarium shops in the 1990s (Akiyama et al. 2003; Nihon-Tanshuigyorui-Aigokai 2003; Ohara and Takagi 2005), some doubt the indigenesness of the population rediscovered in Tanushimaru in 1994 (e.g., Nihon-Tanshuigyorui-Aigokai 2003).

From a contemporary view, the genetic diversity, phylogenetic position, and potential hybridization of Japanese *A. chinensis* should be assessed using genome-wide polymorphism data, as well as mtDNA phylogenetic information, with the inclusion of data from closely related populations and species. One available method for this purpose is the MIG-seq [multiplexed inter-simple sequence repeat (ISSR) genotyping by sequencing] technique, which is a recently developed polymerase chain reaction (PCR)-based method for genome-wide single-nucleotide polymorphism (SNP) genotyping using a high-throughput sequencing platform (Suyama and Matsuki 2015). The MIG-seq method can potentially produce data on several hundred SNPs from a small amount of tissue samples in a relatively easy and cost-effective manner. The method is applicable not only to recently collected samples, but also to samples stored in museums, such as those used by Ohara et al. (2003), as long as they remain viable for PCR.

Using mtDNA phylogenetic and genome-wide SNP data from past and recent samples of captive and wild (reintroduced) populations, the present study aimed to resolve the following three issues: (1) the phylogenetic position of Japanese populations of *A. chinensis*; (2) genetic features of past and present populations, with particular focus on the presence or absence of hybridization with populations (or species) from other countries; and (3) the timing and details of hybridization, if detected. Based on these results, we discuss a

conservation policy and direction for Japanese populations of *A. chinensis*. In brief, the current situation regarding Japanese *A. chinensis* is grim; genuine Japanese populations are very likely no longer extant as a result of hybridization with a closely related species.

Materials and methods

Recent wild and captive populations of Japanese Aphyocypris chinensis. Since the 1980s, by which point wild *A. chinensis* had almost disappeared in Japan, the fish have been discovered twice in irrigation canals at Tanushimaru in Kurume, Fukuoka Prefecture (five fish in 1982 and several dozen fish in 1994). These fish were introduced to captivity, and breeding programs were immediately launched (T82 and T94 strains, respectively; Fig. 1). Because the genuine T82 strain nearly went extinct in the 1990s, present *A. chinensis* populations are considered to have originated from the T94 strain or from partially mixed T94/T82 stocks. Not all introductions to captive populations were formally recorded; hypothesized unrecorded introductions, based on circumstantial or genetic evidence, are herein referred to as “unknown.”

Fish of the T94 strain were directly or indirectly distributed to several aquaria, facilities of administrative organizations, and citizen groups, and thereafter kept in captivity (Fig. 1; Table 1). Some fish have been released to natural habitats, mainly by a citizen group [Hinamoroko Fosterparent Club (HFC)]. Captive fish have occasionally been exchanged among sub-populations (Fig. 1; Table 1).

Wild and captive population samples analyzed. In the present study, a total of 221 specimens from 25 Japanese population samples were analyzed; the sample included specimens from almost all extant wild and ex situ populations (except for some privately

maintained populations), past wild and captive populations, fish traded for hobby aquaria, and an introduced wild population from Shizuoka, which is located 800 km from the natural habitat (Fig. 1; Table 1). Specimens of *A. chinensis* from China (16 specimens from two population samples) and Korea (10 specimens from one population sample) and those of *A. kikuchii* from Taiwan (seven specimens from three population samples) were also analyzed. The mtDNA sequences of these species (AB218688, AF307452, JX184925, AP012123; He et al. 2004; Saitoh et al. 2006; Jang-Liaw et al. 2013; M. Miya, unpublished data) were also obtained from the DNA database (DDBJ/EMBL/GenBank) and used for phylogenetic analysis. The sequences of the following six species, selected based on Tang et al. (2010), were used as outgroups: *Yaoshanicus arcus* (AP011398), *Nicholsicypris normalis* (AP011396), *Macrochirichthys macrochirus* (AP011234), *Parachela williaminae* (HM224301) (Tang et al. 2010), *Pararasbora moltrechti* (JX311312; Jang-Liaw et al. 2013), and *Zacco platypus* (AP012115; Miya et al. 2015).

Sequencing and phylogenetic analyses of mtDNA cytb. Total genomic DNA was isolated from a piece of fin or muscle preserved in 99% ethanol using a Genomic DNA Purification kit (Promega, Tokyo, Japan). Nucleotide sequence data for the mitochondrial cytochrome *b* (*cytb*) gene (1,141 bp) were used to infer phylogenetic relationships among populations of *A. chinensis* and related species. PCR amplification was conducted using the primer pair L14724 (5'- TGA CTT GAA RAA CCA YCG YYG -3'; Palumbi et al. 1991) and H15915 (5'-ACC TCC GAT CTY CGG ATT ACA AGA C -3'; Aoyama et al. 2000). For parts of degenerate specimens, L15172 (5'- TGA GGA CAA ATA TCN TTY TGA GG -3'; Harada et al. 2002) was used instead of L14724 to amplify shorter segments (5' half of *cytb*). PCR conditions consisted of 30 cycles of denaturation (94 °C for 15 s), annealing (48 °C, or 50 °C for L15172, for 15 s), and extension (72 °C for 60 s) performed on a thermal cycler (ASTEC, Fukuoka, Japan). Then PCR products were purified with Illustra ExoStar (GE

Healthcare Japan, Tokyo, Japan) at 37 °C and sequenced on an automated DNA sequencer (GA3130xl; Applied Biosystems, Foster City, CA) using amplification primers and the BigDye Terminator Cycle sequencing FS Ready Reaction kit ver. 3.1 (Applied Biosystems). The *cytb* sequence data (1,141 bp for most specimens) were successfully obtained for a total of 220 specimens of *A. chinensis* and *A. kikuchii* (Table 1). Among these, the sequences of 29 specimens were incomplete (~800 bp), but they could still be assigned to major clades (see “Results”). The resulting sequences were deposited in the DNA database DDBJ/EMBL/GenBank (accession numbers LC497041–497059). The haplotype frequencies of each population were deposited in the Genetic Diversity and Distribution Map (GEDIMAP) freshwater fish database (<http://gedimap.zool.kyoto-u.ac.jp>; Watanabe et al. 2010) with population IDs P2354–2382.

The *cytb* nucleotide sequences were edited and aligned with ClustalW and by eye using MEGA 7.0 (Kumar et al. 2016). Phylogenetic analyses were conducted using the maximum likelihood (ML) and Bayesian inference (BI) methods. The ML tree was estimated using PAUP*4.0a (Swofford 2002) under the HKY + G model selected by Akaike’s information criterion (AIC; base frequencies of A = 0.319, C = 0.274, G = 0.115, and T = 0.292; ti/tv ratio = 9.577 and gamma shape = 0.174), as implemented in jModelTest 2.1.1 (number of substitution schemes = 11; Darriba et al. 2012). The robustness of the ML tree was assessed using the bootstrap method (BP) with 200 replicates by PAUP*. The BI trees were calculated using MrBayes version 3.2.7a (Ronquist et al. 2012) using the HKY+ G model with three codon position partitioning. The Markov chain Monte Carlo (MCMC) process was set so that four chains (three heated and one cold) ran simultaneously. We determined the Bayesian posterior probabilities (BPPs) of the branches based on the pooled trees from two independent runs for 3×10^6 cycles, with every 100 cycles sampled after reaching stationarity of the

likelihood scores. The first 25% of trees were discarded as burn-in. The trees were visualized using FigTree v.1.4.3 (Rambaut 2016).

MIG-seq library construction and population genetic analysis. Genome-wide SNP detection was conducted using MIG-seq, in which loci between two ISSRs were amplified by PCR and sequenced using a next-generation sequencer (Suyama and Matsuki 2015). A MIG-seq library was prepared following the protocol outlined in Suyama and Matsuki (2015), with minor modifications as outlined below. The first round of PCR was conducted using eight ISSR primer sets with tail sequences at an annealing temperature of 38 °C. The Multiplex PCR Assay kit ver. 2 (TaKaRa, Shiga) was used for the first PCR. The products of the first PCR reaction were diluted and used for the second round of PCR. The second PCR was conducted using primer pairs including tail sequences, adapter sequences for Illumina sequencing, and six-base (forward) and eight-base (reverse) barcode sequences to identify each individual sample, using Illumina Nextera/TruSeq DNA dual index adapters 701–724 (i7) and A501–508 (i5; Illumina, San Diego, CA; Document #1000000002694 v10). Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB, Ipswich, MA) was used for the second PCR. The second PCR consisted of 20 cycles of 98 °C for 30 s, 54 °C for 15 s, and 68 °C for 30 s. The products of the second PCR for each individual were mixed and purified using a MinElute PCR Purification Kit (Qiagen, Hilden). Fragments with size ranges of 300–800 bp were isolated and purified using SPRIselect (Beckman Coulter, Brea, CA) and a GeneRead Size Selection kit (Qiagen). Extraction by agarose gel (1%) electrophoresis was also conducted for some specimens for size selection and purification. TapeStation (Agilent Technologies, Santa Clara, CA) was used to confirm fragment size and to estimate the final concentration. Subsequently, the products were sequenced on an Illumina MiSeq platform (Illumina) using the MiSeq Reagent kit v3 (150 cycles; Illumina) following the manufacturer's protocol. MIG-seq data were successfully obtained for a total of 196

specimens of *A. chinensis* and *A. kikuchii* (Table 1). The sequence data were deposited in the DDBJ Sequence Read Archive (accession no. DRA008899).

To analyze population structure and potential hybridization, we primarily applied a Bayesian clustering approach to the MIG-seq data using the following procedures. From the raw sequence data, low-quality ends of reads, primer regions, anchors, and index tags were removed using the FASTX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). We then used the STACKS v1.46 pipeline (Catchen et al. 2013) to generate a format file for STRUCTURE v2.3.4 (Pritchard et al. 2000) and GENEPOP (Rousset 2008). In the STACKS analysis, we first used the program ‘ustacks’ to pile a minimum of two intra-individual reads (parameter $m = 2$) to create a stack (as a putative allele) with maximum mismatches (M) of 2 (a putative locus). The program ‘cstacks’ was used to create a “catalog” of loci across specimens that allowed two mismatches between alleles (parameter $n = 2$). Then, the sets of stacks produced by ‘ustacks’ were searched against the catalog with the program ‘sstacks’. Single nucleotide polymorphisms (SNPs) were retrieved using the program ‘populations’. All samples were treated as a single population (parameter $p = 1$), and only loci with ≥ 15 read depth ($m = 15$) that were present in at least 70% ($r = 0.7$) of individuals were considered. The ‘write_single_snp’ option was used to select only the first SNP of each locus. Only specimens with less than 60% missing SNPs were included in the analysis (average missing rate = 0.17 ± 0.10). The final dataset consisted of 240 SNPs for 196 specimens, and was exported as STRUCTURE and GENEPOP format files. Other parameter values of ‘populations’ (i.e., all combinations of $r = 0.7, 0.8, 0.9$, $m = 10, 15, 20$), producing datasets with 47–359 SNPs, were also analyzed to confirm the robustness of the results.

In the STRUCTURE analysis, 20 replicates for each predefined $K = 1–6$ were obtained with runs of 10,000 burnin and 100,000 MCMC iterations, applying the admixture model with correlated alleles. To select the most likely value of K , the method of Evanno et al.

(2005) was implemented in the program STRUCTURE HARVESTER (Earl and vonHoldt 2012), which determines the second-order rate of change in the distribution of $L(K)$. The final result with the best (and the second best) K was calculated using three runs of 100,000 burn-in and 1,000,000 MCMC iterations. The genetic diversity of each population sample was evaluated as the average expected heterozygosity (H_E) values for SNP sites in the MIG-seq data. H_E was calculated using ARLEQUIN 3.5 (Excoffier and Lischer 2010), the file for which was converted from GENEPOP format.

To visualize the genetic similarities among specimens from different population samples, the same dataset with 240 loci for 196 specimens was used to perform a principal component analysis (PCA), implemented in GenoDive ver. 2.0b23 (Meirmans and Van Tienderen 2004). Missing data were filled by values chosen randomly from all samples, according to the sample allele frequencies. A correlation matrix of allele frequencies among individuals was used to calculate principal component scores. The first two major principal component scores were plotted.

Results

Phylogenetic relationships. In total, 18 haplotypes were identified from 207 specimens of *Aphyocypris chinensis* and *A. kikuchii*. The ML and BI analyses revealed four major clades in the two species with high support values; i.e., a Japan clade (JP), two continental clades (C1 and C2), and a Taiwan clade (TW; Fig. 2). The relationships among clades were not completely resolved, but the same topology ((JP (C1 TW)) C2) was supported in both analyses. The average uncorrected sequence differences (p -distances) between clades were

2.48–4.97% (0.0286–0.0680 in HKY + G distance; Table 2). JP, C1, and TW were relatively close to each other, with 2.48–2.89% *p*-distances (0.0286–0.0344 in HKY + G distance).

The JP clade included five haplotypes [JPN-1–5; Electronic Supplementary Material (ESM) Table S1], all of which were obtained from parts of specimens belonging to captive and wild populations from Tanushimaru, and all three wild specimens from Ukiha (Fig. 2). The C1 clade consisted of eight haplotypes from Chinese and Korean specimens (CK-1–8) and a sequence from the database (AF307452) reported from China. The C2 clade consisted only of haplotypes from Sichuan specimens (CSC-1 and 2). One very notable result was obtained for the TW clade (KKC-1–3), which consisted of haplotypes from Taiwanese specimens (i.e., *A. kikuchii*) and the majority of captive breeding (89%) and wild (79%) specimens of *A. chinensis* in Japan, as well as commercially traded fish in Japan (100%; OKF, *n* = 4). KKC-1 was the only TW haplotype found in Japan (*n* = 116 in total). All specimens in the introduced population of Shizuoka (IZU; *n* = 18) also possessed the KKC-1 haplotype.

Population structure and hybridization. Admixture analysis using STRUCTURE supported division into two or three genomic components for specimens of *A. chinensis* and *A. kikuchii* (maximum $\Delta K = 726.94$ at *K* = 2, second $\Delta K = 439.87$ at *K* = 3 with *m* = 15 and *r* = 0.7; Fig. 3). The results were robust for the datasets (47–359 SNPs) produced from different parameters of ‘populations’ except whether the best *K* was 2 or 3 (ESM Fig. S1). Because the division with *K* = 3 simply included that with *K* = 2, we adopted the results with *K* = 3.

Two of the three components corresponded well to *A. chinensis* from continental China and Korea (samples 25–27) and *A. kikuchii* (Taiwan; samples 28–30); hence, they are hereafter referred to as “continental chinensis” and “kikuchii” elements, respectively. The Japanese samples were composed of two genetic elements, i.e., a Japanese endemic element (“Japanese chinensis”) and a “kikuchii” element, to various degrees. Three samples (3.UKH07+08, 8.LBM00, and 9.LBM01) were composed almost completely of the “Japanese

chinensis” element, whereas others (e.g., 5.IZU18, 10.LBM_HCA00, 14.LBM17, 15.HCA01, 24.OKF16) contained only the “kikuchii” element. Samples from the wild population (1.TNS08, 2.TNS18) and the majority of captive populations (e.g., 6.HFC01, 7.HFC16, 16–18.SML01, 18o, 18n, 19.MWU01, 21.CPR18) showed admixed genomic features; the average proportions of the “kikuchii” elements ranged from 37% to 88% (Table 1). In the results with $K = 2$, the “Japanese chinensis” and “continental chinensis” elements were merged.

The STRUCTURE analysis of the MIG-seq data ($K = 2–6$) did not support the result from mtDNA, in which the Sichuan specimens were the most distant from all others, although the reason for this discrepancy was not clear. However, the analyses using only continental samples ($n = 21$) or all samples except the Japanese samples with the “kikuchii” element ($n = 47$), differentiation between Liaoning + Korean samples (samples 26 and 27) and Sichuan (28) samples was well supported (ESM Fig. S2).

The inferred genomic composition of each population was concordant with the data for mtDNA haplotype groups (JP, C1, C2, and TW; Fig. 3). That is, all Japanese specimens with the pure “Japanese chinensis” element possessed JP haplotypes and those with the pure “kikuchii” element possessed a TW haplotype, whereas those with admixed genomes possessed either JP or TW haplotypes (Fig. 3). The average H_E values for SNP sites in the Japanese wild and captive population samples with $n \geq 3$ ranged from 0.019 (3.UKH07+08) to 0.145 (11.LBM_HCA05). The samples of the pure “Japanese chinensis” element (3.UKH07+08, 8.LBM00, and 9.LBM01) had significantly low H_E values (0.019–0.028) compared with other Japanese samples (0.041–0.145) (Mann–Whitney U test, approximate $P = 0.009$; Table 1; Fig. 4).

The scatter plots for the first two principal components clearly supported the genetic separation among the putatively genuine Japanese *A. chinensis* samples (3.UKH07+08,

8.LBM00, and 9.LBM01), *A. chinensis* from continental China and Korea, and *A. kikuchii* of Taiwan (Fig. 5). Specimens from other Japanese samples were located between the putatively genuine Japanese *A. chinensis* and *A. kikuchii* of Taiwan, or overlapped with them. The putatively genuine Japanese *A. chinensis* samples had 7.8–16.1% unique alleles compared to other continental and Taiwan samples (ESM Table S2).

Collectively, mtDNA and MIG-seq data strongly suggest that the Japanese populations of *A. chinensis* have hybridized with *A. kikuchii*, with the exception of a few lost populations.

Discussion

Current situation of *Aphyocypris chinensis* in Japan. The results of mtDNA and genome-wide SNP analyses consistently demonstrated that the Japanese populations of *Aphyocypris chinensis* have hybridized with *A. kikuchii*, the endemic species of Taiwan. The Japanese *A. chinensis* populations could have undergone some extent of genetic drift because of small population size and bottleneck/founder effects. However, genetic drift alone cannot explain our results on the sharing of mtDNA and nuclear DNA elements with *A. kikuchii*. Also, the presence of unique alleles in the putatively genuine Japanese samples denies that the “Japanese *chinensis*” element resulted from analytical artifacts.

Our data provide evidence that genuine Japanese populations of *A. chinensis* existed in captivity until the early 2000s since the discovery of the T94 strain in 1994 (LBM00, 01; Fig. 6). In addition, specimens obtained from a wild habitat (UKH) in 2007 and 2008 were determined to be from a genuine Japanese population. However, all known extant wild and captive populations of Japanese *A. chinensis* have hybridized with *A. kikuchii*. In other words, genuine Japanese populations of *A. chinensis* have very likely gone extinct.

Genuine Japanese individuals of *A. chinensis*, if they still exist, may be found in captive populations derived from early generations of the T94 strain and subsequently kept independent from the others. The T94 strain was inferred to originally have been a Japanese indigenous population, as LBM00 and LBM01, which were from the T94 strain that had been maintained independently since 1994, were not affected by introgression of genetic elements of *A. kikuchii*.

However, whether the T94 strain originated from a surviving native population near the discovery site or from a captive population privately introduced to the site is not clear.

Another possibility is that genuine Japanese *A. chinensis* may persist in wild habitats to which no recent captive population has been introduced. A population at Ukiha, about 10 km from Tanushimaru, was temporarily found in the latter half of the 2000s, and was identified as a genuine Japanese population in the present study. Although the species was not found at Ukiha in a recent survey (autumn 2018; Kobayakawa and Takaku, unpublished data), further investigation is necessary to find genuine Japanese populations in the wild.

The mixture of *A. kikuchii* into Japanese *A. chinensis* populations likely occurred through the introduction of aquarium fish. In the 1990s, *A. kikuchii* was undoubtedly circulated in the hobby aquarium market in Japan based on the following evidence. The semi-wild population OKF16 has been maintained by ‘OK Fish Farm’, a private ex situ conservation facility, since the founder fish were purchased from an aquarium shop and introduced in the 1990s. The wild population in Shizuoka (IZU) was also introduced in the 1990s by informal activity or by accident. Both of these populations are remote from the native range of the species and are revealed to be of *A. kikuchii*, although they have been considered to be *A. chinensis* (e.g., Kitahara 2009). Also, several private records exist regarding the aquarium trade of “*A. chinensis*” from overseas in the 1990s (e.g., Nihon-Tanshuigyorui-Aigokai 2003), and captive-bred fish of “*A. chinensis*” are still

commonly traded as of 2019. The morphological differences between *A. chinensis* and *A. kikuchii* are quite subtle; i.e., the lateral line of the former, but not the latter, extends to above the pelvic-fin base (Wu et al. 1964; Zhu et al. 2013). Therefore, traded fish of *A. kikuchii* mislabeled as *A. chinensis* could have been introduced into a captive *A. chinensis* population(s) without recording this activity. According to our genetic tracing, the initial introduction must have happened before 2001, and the genetic elements of *A. kikuchii* spread throughout all captive and reintroduced/reinforced populations until the mid2000s (Fig. 6). The low mtDNA diversity of the Japanese *A. kikuchii* (only one haplotype) suggests that the introduced population of *A. kikuchii* was not very large.

Although morphological differences between *A. chinensis* and *A. kikuchii* are unclear, they are genetically well differentiated. Our mtDNA analysis revealed that the Japan (JP), one continental (C1), and Taiwanese (TW, *A. kikuchii*) clades are clearly differentiated from each other at a similar level (2.5–2.9% *p*-distances). These genetic distances would be equivalent to 1 to 2 million years, if the molecular rate of 1–2% pairwise differentiation per million years in the cyprinid *cytb* is adopted (e.g., 1.5%, Zardoya and Doadrio 1999; 1.2–2.5%, Tominaga et al. 2016; 1.2–2.2%, Watanabe et al. 2018; 1.5–1.8%, Jang-Liaw et al. 2019; the ranges show 95% credible intervals). Considering their distinct differentiation, as well as the species status of *A. kikuchii*, the Japanese

populations of *A. chinensis* could be treated as an independent species. In other words, the loss of genuine Japanese populations of *A. chinensis* does not just mean the extinction of one regional population of a widely distributed species, but rather is equivalent to the loss of an endemic species of Japan.

Lessons from the case of *A. chinensis* in Japan. Two major lessons can be inferred from the highly probable extinction of genuine Japanese populations of *A. chinensis*. First, the origin and authenticity of individuals used for ex situ conservation programs should be

carefully examined using molecular genetic techniques at the initial stage of program activity. Genetic management of captive populations that includes the exchange of individuals among fragmented subpopulations is reasonable to alleviate inbreeding, loss of genetic diversity, and adaptation to captivity, all of which are important issues to consider in captive breeding programs (Frankham et al. 2002; Williams and Hoffman 2009). This type of management strategy was adequately implemented since the initial steps of captive breeding of the T94 strain and has been further supported by genetic diversity assessments (Ohara et al. 2003; Ohara and Takagi 2005, 2007). Based on the experience of fish managers, the exchange of fish actually improves reproduction and alleviates inbreeding depression (e.g., M. Matsuda, personal observation). However, there have also been situations in which fish labeled “*A. chinensis*” became common in the aquarium market (in the 1990s), and a wild population was established at a remote site (IZU). Considering present-day molecular techniques and the accumulation of genetic data in the database, similar failures that mix different species in captivity should never be repeated.

Second, as many habitats and populations as possible should be preserved. Because the agricultural ditches in Tanushimaru had become the only known wild habitat of *A. chinensis* in Japan by 1994, the Tanushimaru population (T94, partly with survivors of the T82 strain from the same locality) has been used for ex situ preservation and stocking as the only remaining Japanese *A. chinensis* population. As revealed in this study, this captive population had been unintentionally mixed with *A. kikuchii* within the first several years. On the other hand, another wild population of *A. chinensis* was found at Ukiha in the mid-2000s and was determined to be a genuine Japanese population in the present study. However, the Ukiha population has not been preserved in captivity and does not currently exist in the wild. The present case emphasizes the importance of the independent management of multiple populations for endangered species.

Perspective for conservation of *A. chinensis* in Japan. We propose the following three issues for the future conservation of *A. chinensis* in Japan. First, further investigation should be conducted immediately to determine whether any genuine Japanese populations of *A. chinensis* still exist. As discussed above, such populations may exist in both captivity and the wild. Wide-ranging and high-resolution surveys using environmental DNA would be useful (e.g., Takahara et al. 2012; Miya et al. 2015), as well as typical fish capture surveys. Second, until a genuine population can be discovered, hybrid populations of Japanese *A. chinensis* with *A. kikuchii* should be preserved as the only remaining populations with genetic features of Japanese *A. chinensis*. Genetic management is necessary to prevent an increase of genetic elements of *A. kikuchii* in captive populations. It should be essential that genetic information of each captive population is shared among responsible managers and administrators. Third, to achieve the first two goals, genuine *A. kikuchii* populations in the wild (IZU) and in captivity should be clearly recognized as a species distinct from *A. chinensis* and should be managed separately as populations of an exotic species.

Acknowledgements We sincerely thank N. Suzuki (Chiba Prefecture), K. Okura (Fukuoka City), O. Masuda, T. Miki, and M. Takeda (HCA), M. Murakami (HFC), T. Nagata, S. Yoshikawa, K. Yamada, T. Suzuki (LBM), Y. Suzuki, M. Nakamura (MWU), M. Onuma (National Institute for Environmental Studies, Japan), Y. Nakamura (OK Fish Farm), H. Nakashima (OMC), T. Satonaka (SML), and N. Akiyama (TKU) for providing specimens and valuable information, Y. Suyama (Tohoku University) for providing advice for the MIG-seq experiment, Y. Fuke (Kyoto University) for helping in molecular experiments, S. Nonaka (Dojyo Club) for providing a photo, and H. Akiyama and N. Daiki for helping arrangement of the photo collections in LBM. This study was supported in part by JSPS KAKENHI (nos. 18K14802 and 18H01330 to R. Tabata).

References

- Akiyama N (1988) Artificial breeding of *Rhodeus suigensis* and *Aphyocypris chinensis*.
Tansuigyo-hogo (1):54–55
- Akiyama N, Kitano T, Ueda M (2003) Complete guide for breeding freshwater fish. MPJ,
Tokyo
- Aoyama S, Watanabe S, Ishikawa S, Nishida M, Tsukamoto K (2000) Are morphological
characters distinctive enough to discriminate between two species of freshwater eels,
Anguilla celebesensis and *A. interioris*? Ichthyol Res 47:157–161
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool
set for population genomics. Mol Ecol 22:3124–3140
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new
heuristics and parallel computing. Nat Methods 9:772
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for
visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet
Resour 4:359–361
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using
the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform
population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567
- Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge
University Press, Cambridge

- Harada S, Jeon SR, Kinoshita I, Tanaka M, Nishida M (2002) Phylogenetic relationships of four species of floating gobies (*Gymnogobius*) as inferred from partial mitochondrial cytochrome *b* gene sequences. *Ichthyol Res* 49:324–332
- He S, Liu H, Chen Y, Kuwahara M, Nakajima T, Zhong Y (2004) Molecular phylogenetic relationships of Eastern Asian Cyprinidae (Pisces: Cypriniformes) inferred from cytochrome *b* sequences. *Sci China C Life Sci* 47:130–138
- IUCN/SSC (2014). Guidelines on the use of *ex situ* management for species conservation. Version 2.0. Gland, Switzerland: IUCN Species Survival Commission.
<https://portals.iucn.org/library/sites/library/files/documents/2014-064.pdf>. Accessed 25 March 2019
- Jang-Liaw NH, Tsai CL, Watanabe K (2013) Complete mitochondrial genome of the Kikuchi's minnow *Aphyocypris kikuchii* (Teleostei, Cyprinidae). *Mitochondrial DNA* 24:11–13
- Jang-Liaw NH, Tominaga K, Zhang C, Zhao Y, Nakajima J, Onikura N, Watanabe K (2019) Phylogeography of the Chinese false gudgeon, *Abbottina rivularis*, in East Asia, with special reference to the origin and artificial disturbance of Japanese populations. *Ichthyol Res* 66:460–478
- Kimura S (1997) *Aphyocypris chinensis*. In: Nagata Y, Hosoya K (eds) *Circumstances in endangered Japanese fishes and their protection*. Midori-shobo, Tokyo, pp 54–63
- Kitahara Y (2009) Establishment of *Aphyocypris chinensis* Günther at Matsuzaki town, Izu District, Shizuoka Prefecture, Japan. *Nanki Seibutsu* 51:9–12
- Kobayakawa M (2015) Hinamoroko: *Aphyocypris chinensis* Günther, 1868. In: Ministry of the Environment (ed) *Red data book 2014 – Threatened wildlife of Japan*–, vol 4, Pisces –brackish and fresh water fishes. Gyosei Corp, Tokyo, pp 34–35

- Kobayakawa M (2016) Conservation and reintroduction with the cooperation of administrations, citizens, and researchers: endangered species Hinamoroko (*Aphyocypris chinensis*). In: The Nature Conservation Committee of Ichthyological Society of Japan (ed) The challenges of freshwater fish conservation: concepts and practice for restoration of aquatic biodiversity. Tokai University Press, Hadano, pp 161–172
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874
- Lehmkuhl JF (1984) Determining size and dispersion of minimum viable populations for land management planning and species conservation. *Environ Manag* 8:167–176
- Maehata M (1997) Circumstances and problems in preservation at aquarium. In: Nagata Y, Hosoya K (eds) Circumstances in endangered Japanese fishes and their protection. Midori-shobo, Tokyo, pp 205–217
- Margan SH, Nurthen RK, Montgomery ME, Woodworth LM, Lowe EH, Briscoe DA, Frankham R (1998) Single large or several small? Population fragmentation in the captive management of endangered species. *Zoo Biol* 17:467–480
- Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4:792–794
- Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *R Soc Open Sci* 2:150088
- Nakajima J, Kobayakawa M, Nakamura T, Takaku K, Kano Y, Inui R, Oishi S, Onikura N, Hosoya K (2012) Phylogeography of *Aphyocypris chinensis* and proposal of

- conservation policy in Fukuoka Prefecture, northern Kyushu, Japan. Advance abstracts for the 45th annual meeting, 2012, the Ichthyological Society of Japan:41
- Nihon-Tansuigyori-Aigokai (2003) Towns protecting introduced fish.
<https://tansuigyo.net/a/link0.html>. Accessed 20 August 2018
- Ohara K, Takagi M (2005) Hinamoroko—the fish extincted in wild. In: Katano O, Mori S (eds), *Kishotansuigyo no Genzai to Mirai* (The present and future of endangered freshwater fishes of Japan). Shinzansha, Tokyo, Japan, pp 157–167
- Ohara K, Takagi M (2007) Survey of genetic variation at three microsatellite loci in captive populations of endangered Japanese minnow *Aphyocypris chinensis* with implications for reduction of inbreeding. *Fish Sci* 73:156–160
- Ohara K, Takagi M, Kaneko Y, Takei M (2003) Allozymic variation in an endangered Japanese minnow *Aphyocypris chinensis*. *Ichthyol Res* 50:86–89
- Palumbi S, Martin A, Romano S, McMillian WO, Stice L, Grabowski G (1991) The simple fool's guide to PCR. University of Hawaii, Honolulu
- Pritchard JK, Stephens P, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Rambaut A (2016) FigTree Version 1.4.3. <http://tree.bio.ed.ac.uk/software/figtree/>. Accessed 20 August 2018
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MRBAYES 3.2: efficient Bayesian phylogenetic inference and model selection across a large model space. *Syst Biol* 61:539–542
- Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol Ecol Resour* 8:103–106
- Saitoh K, Sado T, Mayden RL, Hanzawa N, Nakamura K, Nishida M, Miya M (2006) Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii: Ostariophysi): the first evidence toward resolution of higher-level relationships of the

- world's largest freshwater fish clade based on 59 whole mitogenome sequences. *J Mol Evol* 63:826–841
- Soulé ME (1980) Thresholds for survival: maintaining fitness and evolutionary potential. In: Soulé ME, Wilcox BA (eds) *Conservation biology: an evolutionary-ecological perspective*. Sinauer, Sunderland, pp 151–169
- Suyama Y and Matsuki Y (2015) MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Sci Rep* 5:16963. <https://doi.org/10.1038/srep16963>
- Swofford DL (2002) PAUP* 4.0b10a: phylogenetic analysis using parsimony (and other methods). Sinauer, Sunderland
- Takahara T, Minamoto T, Yamanaka H, Doi H, Kawabata Z (2012) Estimation of fish biomass using environmental DNA. *PLoS ONE* 7:e35868. <https://doi.org/10.1371/journal.pone.0035868>
- Takaku K, Kobayakawa M, Onikura N, Ohara K, Hosoya K (2007) Disappearance of Hinamoroko (*Aphyocypris chinensis*) and its traditional habitat from Japanese rice fields. *Jpn J Ichthyol* 54:231–234
- Tang KL, Agnew MK, Hirt MV, Sado T, Schneider LM, Freyhof J, Sulaiman Z, Swartz E, Vidithayanon C, Miya M, Saitoh K, Simons AM, Wood RM, Mayden RL (2010) Systematics of the subfamily Danioninae (Teleostei: Cypriniformes: Cyprinidae). *Mol Phylogenet Evol* 57:189–214
- Tominaga K, Nakajima J, Watanabe K (2016) Cryptic divergence and phylogeography of the pike gudgeon *Pseudogobio esocinus* (Teleostei: Cyprinidae): a comprehensive case of freshwater phylogeography in Japan. *Ichthyol Res* 63:79–93
- United Nations (1992) Convention of biological diversity. <https://www.cbd.int/doc/legal/cbd-en.pdf>. Accessed: 25 March 2019

- Watanabe K, Kano Y, Takahashi H, Mukai T, Kakioka R, Tominaga K (2010) GEDIMAP: a database of genetic diversity for Japanese freshwater fishes. *Ichthyol Res* 57:107–109
- Watanabe K, Sakai H, Sanada T, Nishida M (2018) Comparative phylogeography of diadromous and freshwater daces of the genus *Tribolodon* (Cyprinidae). *Ichthyol Res* 65:383–397
- Williams SE, Hoffman EA (2009) Minimizing genetic adaptation in captive breeding programs: a review. *Biol Conserv* 142:2388–2400
- Wu XW et al. (1964) The cyprinid fishes of China. Part I. Shanghai Science and Technology, Shanghai
- Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *J Mol Evol* 49:227–237
- Zhu Y, Zhao Y, Huang K (2013) *Aphyocypris pulchrilineata*, a new miniature cyprinid species (Teleostei: Cypriniformes: Cyprinidae) from Guangxi, China. *Ichthyol Res* 60:232–236

Figure legends

Fig.1 Details of the captive breeding, exchange of fish, and reintroduction/reinforcement into the wild of Japanese populations of *Aphyocypris chinensis*. Two source populations are known to have contributed to the captive populations since 1982; T82 and T94 were collected at Tanushimaru in 1982 and 1994, respectively. In addition, an “unknown” source was likely introduced into HCA before 2001. Different bar colors correspond to different or unknown sources. The figures near arrowheads indicate the year and number of fish (in parentheses) translocated as far as are known (i.e., no year data means that the timing of the translocation is unknown). *BWB* Biwakobunkakan, *CPF* Conservation Pond in Fukuoka City, *CPR* Clean Park Rinkai, *FEC* Fukuoka Eastern Clean Plant, *HCA* Himeji City Aquarium, *HFC* Hinamoroko Foster-parent Club, *KEE* Kyushu Environmental Evaluation Association, *LBM* Lake Biwa Museum, *MWU* Marine World Uminonakamichi, *OMC* Omuta captive population, *SML* Shima Marineland, *TKU* Tokai University, *TNS* Tanushimaru wild population, *UKH* Ukiha wild population. See Table 1 for more details

Fig.2 The ML tree of *Aphyocypris chinensis* and *A. kikuchii* with selected outgroups based on *cytb* sequences (1,141 bp). Only the closest outgroup clade is shown. The OTU labels are represented as the haplotype name, number of specimens (in parentheses), and sample abbreviation (see Table 1). The clade names (JP, C1, C2, and TW) and localities (or origins of captive populations) are shown. Numbers at major internodes represent ML bootstrap probability values for 200 replicates and Bayesian posterior probability values. Photographs of fish from representative samples are shown: from top, fish collected from Saga Prefecture in 1971 (by S. Nonaka), T82 specimens kept at Biwako-bunkakan (1991 by H. Akiyama), T82 + unknown origin from Himeji City Aquarium kept at the Lake Biwa Museum (2008 by

M. Matsuda), and *A. kikuchii* from Hualian, Taiwan in 2006 [National Museum of Natural Science (Fish), Taichung (NMNSF) 01501-F, by N.-H. Jang-Liaw]

Fig. 3 Results of admixture analysis for captive and wild populations of *Aphyocypris chinensis* and *A. kikuchii* based on data for 240 SNPs obtained using STRUCTURE (Pritchard et al. 2000). **a** Results of the assignment for $K = 3$ (second best). Each vertical bar represents an individual partitioned into the three clusters defined by STRUCTURE. Different colors represent three genetic elements; blue “Japanese chinensis,” green “continental chinensis,” and red “kikuchii” (see text). In the results with $K = 2$, the “Japanese chinensis” and “continental chinensis” elements were merged. Circles below the bars indicate mtDNA types (see Fig. 2); blue JP, green C1, violet C2 (*A. chinensis*), red TW (*A. kikuchii*), and dash missing. For sample abbreviations, see Fig. 1 and Table 1.

b ΔK following Evanno et al. (2005) as a function of the number of assumed populations ($K = 1-6$)

Fig. 4 Frequency distribution of the average expected heterozygosity in MIG-seq data for the suspected genuine Japanese *A. chinensis* populations (black), hybrid populations (gray), and genuine *A. kikuchii* populations (open)

Fig. 5 Scatter plot of the first and second principal components for 240 SNPs from the MIG-seq analysis. The percent of variation explained by each axis is shown in parentheses

Fig. 6 Genetic features of wild and captive population samples of Japanese *Aphyocypris chinensis* inferred by mtDNA and genome-wide SNP analyses. The analyzed samples are mapped on the chronological scheme of population management. Blue circles genuine

Japanese population, *yellow triangles* hybrids between *A. chinensis* and *A. kikuchii*, *red triangle* (almost) genuine *A. kikuchii* population. For sample ID (shown in symbols) and abbreviations, see Fig. 1 and Table 1

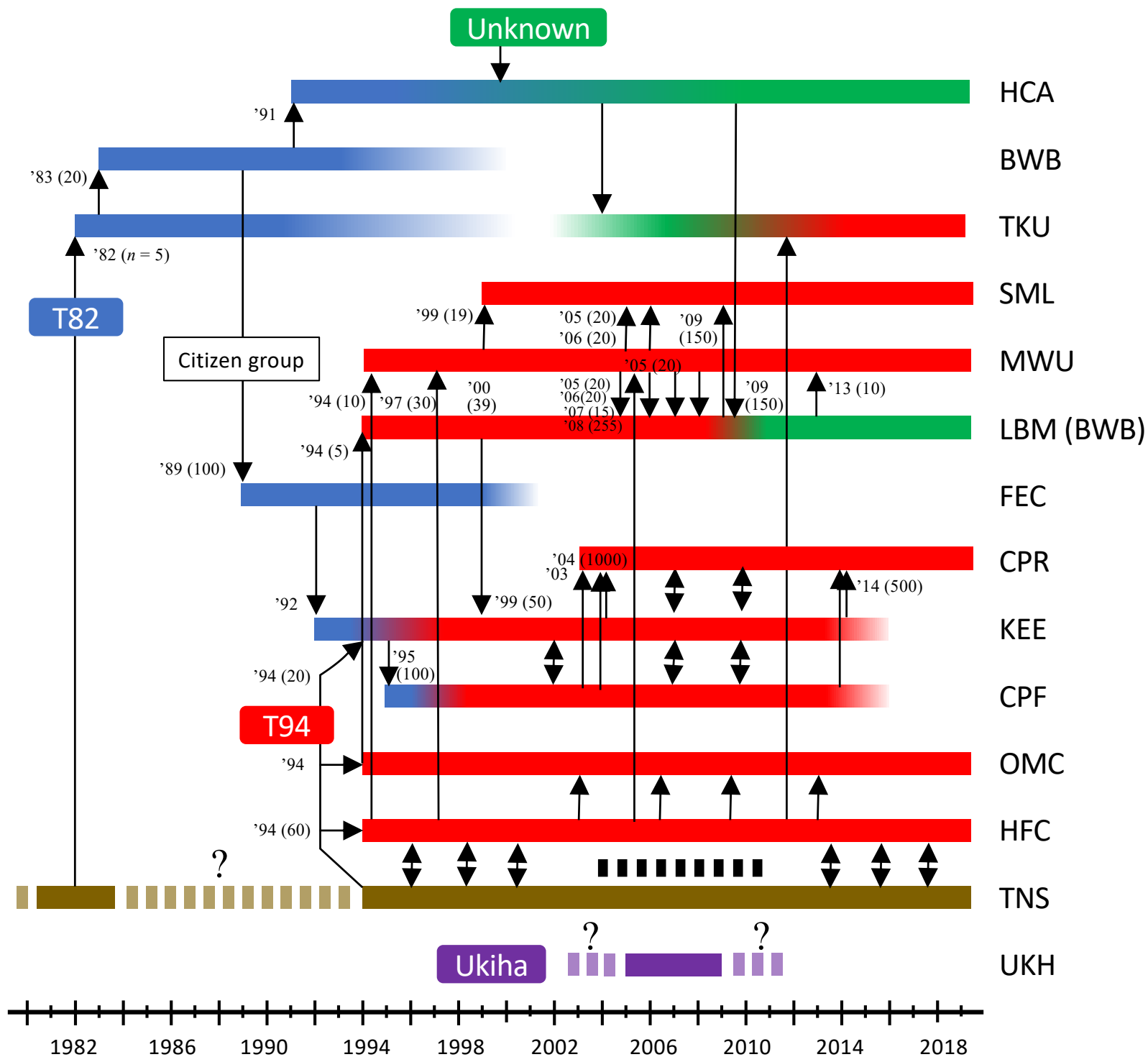


Fig. 1

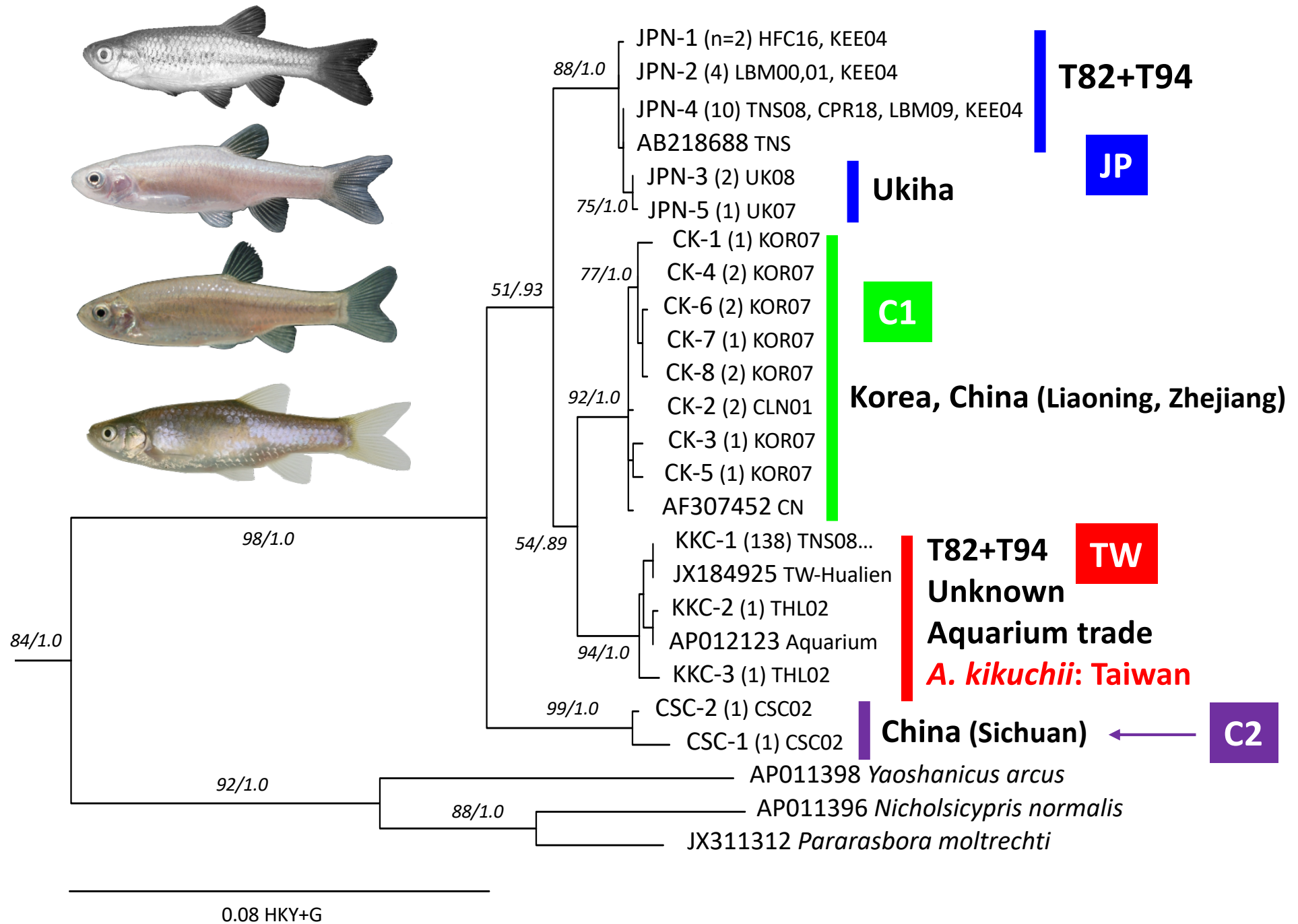


Fig. 2

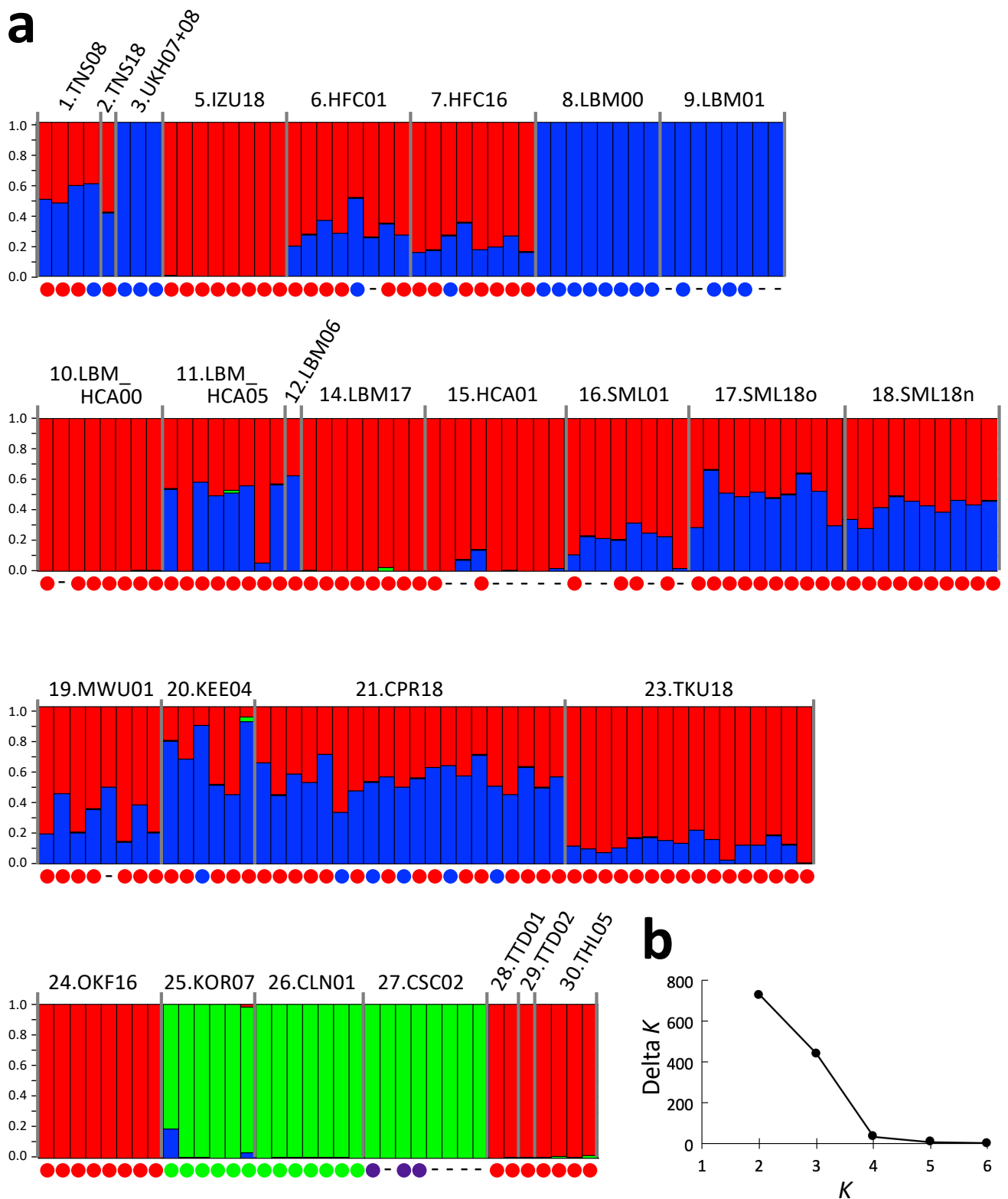


Fig. 3

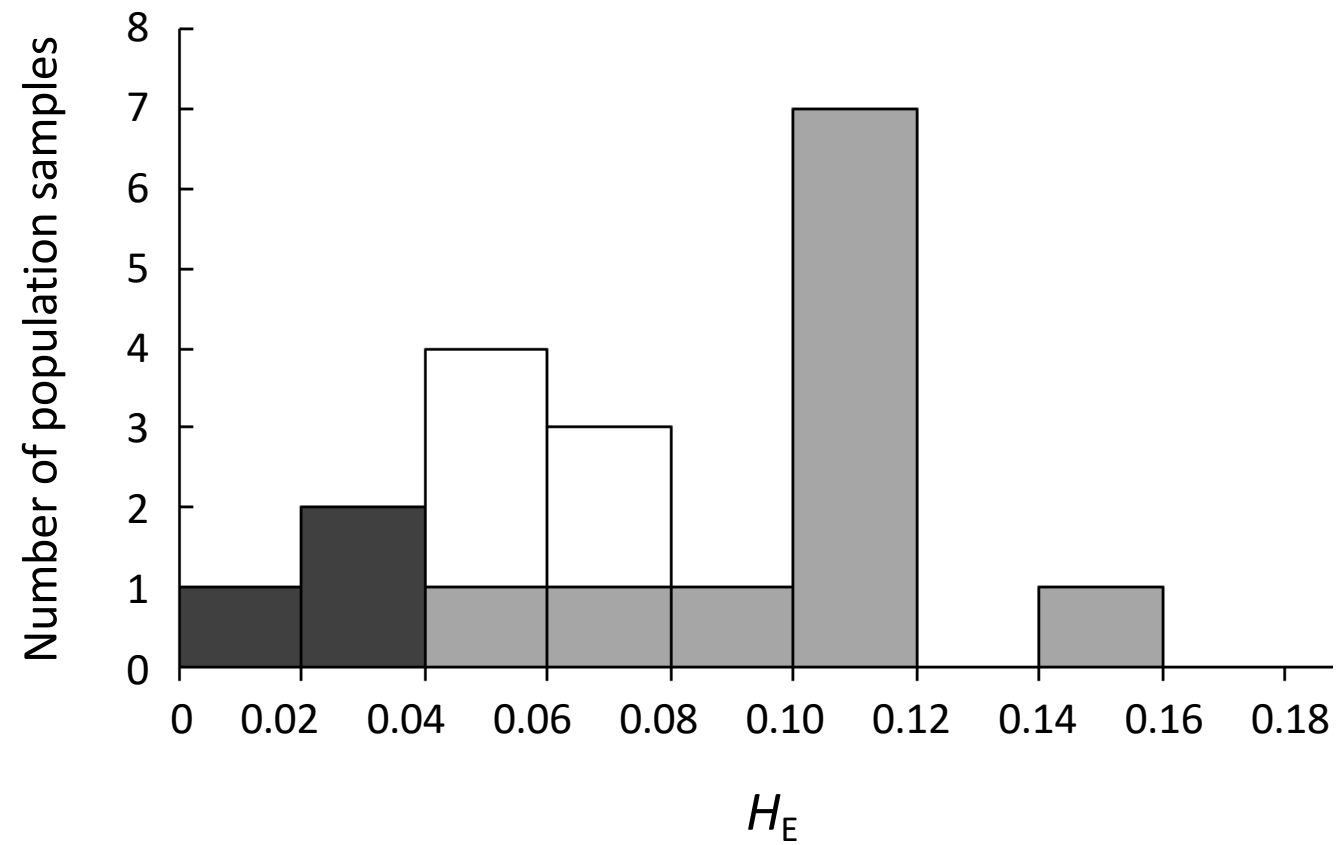


Fig. 4

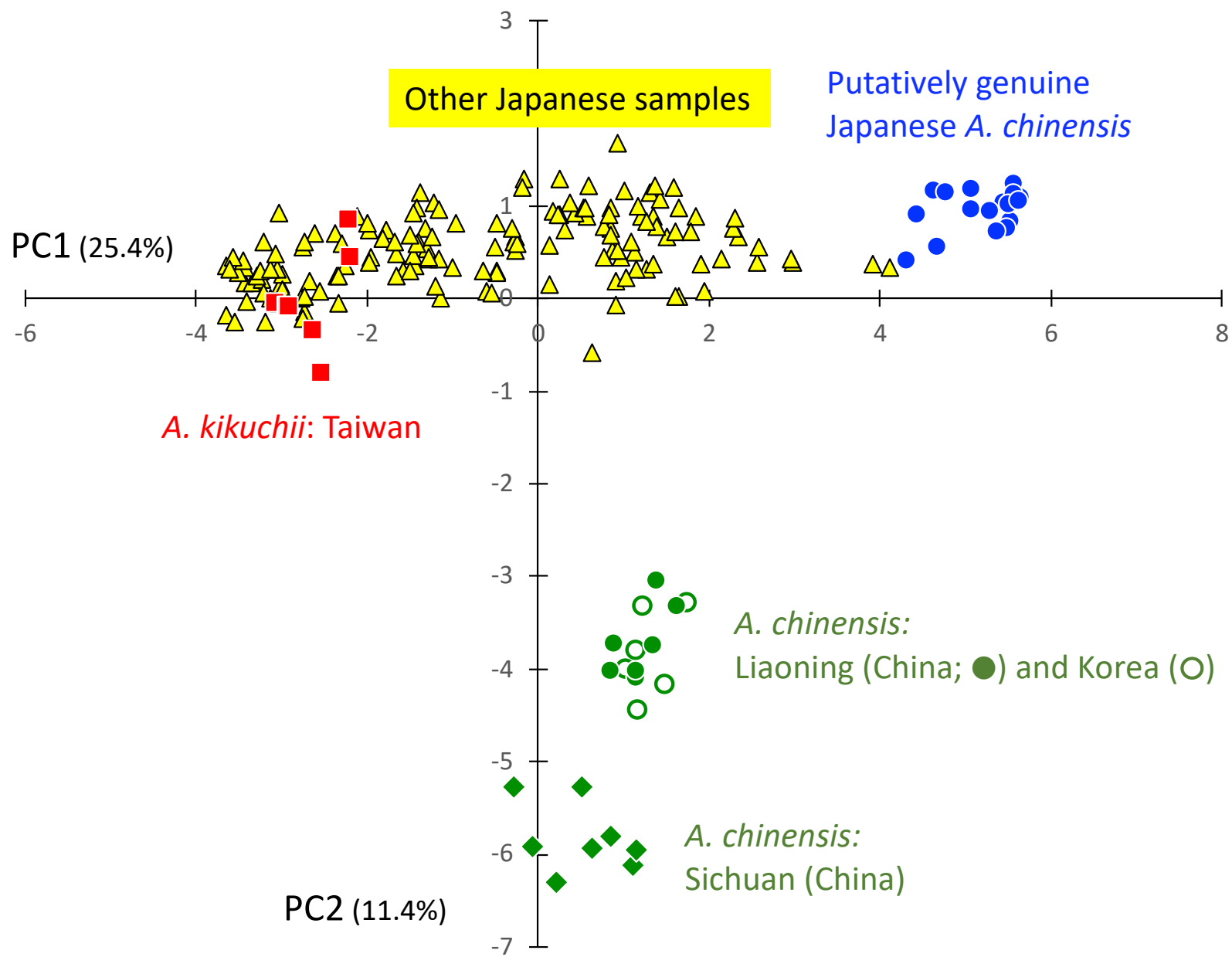


Fig. 5

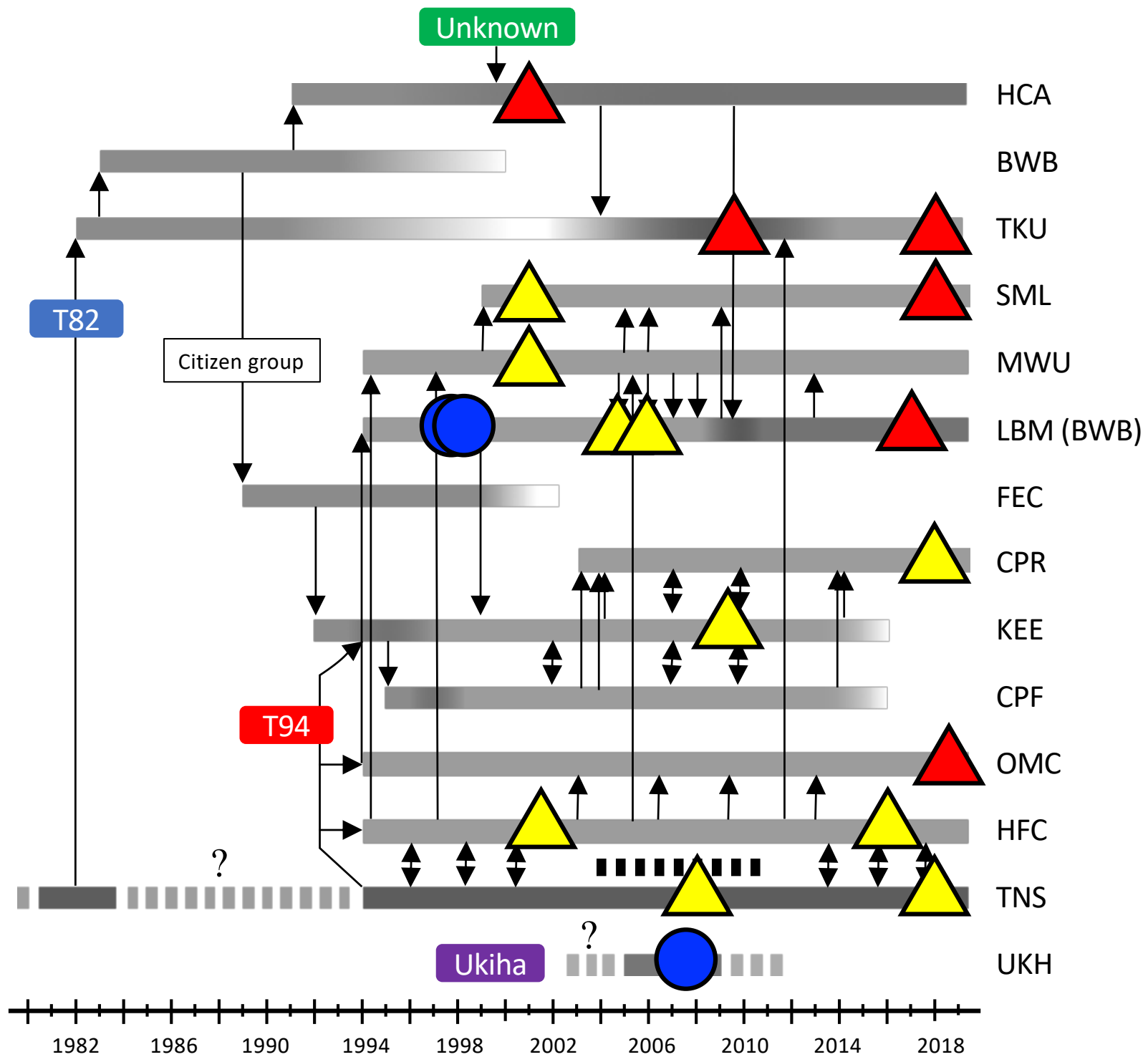


Fig. 6

Table 1 Wild and captive populations of *Aphyocypris chinensis* and comparative materials analyzed in this study

Sample #	Code	Locality or institution	Known origin	Collection	Catalog No. ¹ (GEDIMAP ID)	Number of specimens				Average H_E of MIG-seq
						mtDN A	(KKC) 2	MIG-seq q	(KKC) 3	
Wild										
1	TNS08	Tanushimaru, Kurume, Fukuoka, Japan	T94-strain	2008	uncataloged (P2354)	6	(3)	4	(46%)	0.124 ± 0.221
2	TNS18	Tanushimaru, Kurume, Fukuoka, Japan	T94-strain	2018	uncataloged (P2355)	1	(1)	1	(59%)	0.092 ± 0.289
3	UKH07+08	Ukiha, Fukuoka, Japan	Native	2007, 2008	uncataloged (P2356)	3	(0)	3	(0%)	0.019 ± 0.108
4	IZU11	Matsuzaki, Shizuoka, Japan	Introduced from unknow origin	2011	uncataloged (P2357)	2	(2)	—	—	—
5	IZU18	Matsuzaki, Shizuoka, Japan	Introduced from unknow origin	2018	uncataloged (P2358)	16	(16)	8	(100%)	0.042 ± 0.139
Captive										
6	HFC01	Hinamoroko Foster-parent Club; Fukuoka, Japan	T94-strain; used in Ohara et al. (2003)	2001	LBM1210034576–1210034578 (P2359)	7	(6)	8	(69%)	0.111 ± 0.200
7	HFC16	Hinamoroko Foster-parent Club; Fukuoka, Japan	T94-strain	2016	uncataloged (P2360)	8	(7)	8	(78%)	0.099 ± 0.177
8	LBM00	Lake Biwa Museum; Shiga, Japan; 1998 year-class of T94-strain	Genuine T94-strain, 1998-year class	2000	LBM1210029751–1210029753 (P2361)	8	(0)	8	(0%)	0.028 ± 0.104
9	LBM01	Lake Biwa Museum; Shiga, Japan	Genuine T94-strain; used in Ohara et al. (2003)	2001	LBM1210032516, 1210034461–1210034463 (P2362)	4	(0)	8	(0%)	0.027 ± 0.101
10	LBM_HCA00	Lake Biwa Museum; Shiga, Japan; From HCA	T94-strain from HCA (patially lacking information)	2000	LBM1210029751–1210029753 (P2363)	7	(7)	8	(100%)	0.063 ± 0.152
11	LBM_HCA05	Lake Biwa Museum; Shiga, Japan	T94-strain mixed with HCA	2005	LBM1210044783–1210044786 (P2364)	8	(8)	8	(58%)	0.145 ± 0.220
12	LBM06	Lake Biwa Museum; Shiga, Japan	T94-strain mixed with HCA?	2006	LBM1210046456 (P2365)	1	(1)	1	(37%)	0.138 ± 0.345
13	LBM09_FS	Lake Biwa Museum; Shiga, Japan	T94-strain; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2009	NIES ID 1881P–11891P (P2366)	11	(10)	—	—	—
14	LBM17	Lake Biwa Museum; Shiga, Japan	TKU from HCA	2017	uncataloged (P2367)	8	(8)	8	(99%)	0.041 ± 0.128

15	HCA01	Himeji City Aquarium; Hyogo, Japan	T82-strain from Biwako-bunkakan probably followed by introduction of an unknown strain; used in Ohara et al. (2003)	2001	LBM1210034617–1210034621 (P2368)	2	(2)	9	(97%)	0.069 ± 0.151
16	SML01	Shima Marineland; Mie, Japan	T94-strain from MWU in 1999; used in Ohara et al. (2003)	2001	LBM1210034504–1210034506 (P2369)	4	(4)	8	(80%)	0.062 ± 0.148
17	SML18o	Shima Marineland; Mie, Japan	T94-strain probably from LBM	2018	uncataloged (P2370)	10	(10)	10	(51%)	0.112 ± 0.197
18	SML18n	Shima Marineland; Mie, Japan	T94-strain from MWU and LBM	2018	uncataloged (P2371)	10	(10)	10	(58%)	0.101 ± 0.189
19	MWU01	Marine World Uminonakamichi; Fukuoka, Japan	T94-strain from HFC in 1995, 1997 and 2000; used in Ohara et al. (2003)	2001	LBM1210034554–1210034557 (P2372)	7	(7)	8	(70%)	0.101 ± 0.184
20	KEE04	Kyushu Environmental Evaluation Association; Fukuoka, Japan	T82 and T94-strains; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2004	NIES ID 1618P–1634P (P2373)	17	(14)	6	(30%)	0.103 ± 0.189
21	CPR18	Clean Park Rinkai; Fukuoka, Japan	T94-strain and probably T82 from KEE and a conservation pond in Fukuoka (CPF)	2018	uncataloged (P2374)	20	(15)	20	(46%)	0.113 ± 0.187
22	OMC	Omuta; Fukuoka, Japan	T94-strain probably mixed with HFC	2019	uncataloged (P2375)	13	(13)	—	—	—
23	TKU10_FS	Tokai University; Shizuoka, Japan; via Research Institute of Environment, Agriculture and Fisheries, Osaka Prefecture	T94-strain from HFC and HCA?; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2010	NIES ID 1816P, 1817P, 1819P, 1820P (P2376)	4	(4)	—	—	—
24	TKU18	Tokai University; Shizuoka, Japan	T94-strain from HFC and HCA?	2018	uncataloged (P2377)	16	(16)	16	(88%)	0.047 ± 0.130
25	OKF16	OK Fish Farm; Saitama, Japan	Aquarium trade in Tokyo in 1990s	2016	uncataloged (P2378)	4	(4)	8	(100%)	0.047 ± 0.138
Outside of Japan										
26	KOR07	Gunsan, Korea	Native	2007	uncataloged (P2379)	10	(0)	6	(1%)	0.078 ± 0.161
27	CLN01	Liaoning, China	Native; used in Ohara et al. (2003)	2001	LBM1210034659–1210034661 (P2380)	8	(0)	7	(0%)	0.058 ± 0.154
28	CSC02	Sichuan, China	Native?; Donated from Institute of Hydrobiology, Chinese Academy of Sciences	2002	LBM1210036127–1210036129 (P2381)	3	(0)	8	(0%)	0.050 ± 0.135
29	TTD01	Taidong, Taiwan	Native (<i>A. kikuchii</i>)	2001	LBM1210034720–1210034721	—	—	2	(100%)	0.055 ± 0.173
30	TTD02	Taidong, Taiwan	Native (<i>A. kikuchii</i>)	2002	uncataloged	—	—	1	(99%)	0.042 ± 0.200

31	THL05	Hualian, Taiwan	Native (<i>A. kikuchii</i>)	2005–2006	NMNSF01373-3, NMNSF1501-1, uncataloged (2) (P2382)	2	(2)	4	(99%)	0.070 ± 0.197
Total						220		196		

¹ *LBM* Lake Biwa Museum, Shiga, *NIES* National Institute for Environmental Studies, Tsukuba; *NMNSF* National Museum of Natural Science (Fish), Taichung

² Number of specimens with a haplotype of *A. kikuchii*

³ Proportion of the genomic element of *A. kikuchii* (Structure, $K = 3$; 'populations' parameter, $r = 0.7$, $m = 15$ in stacks)

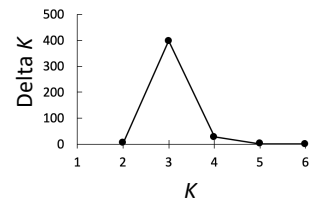
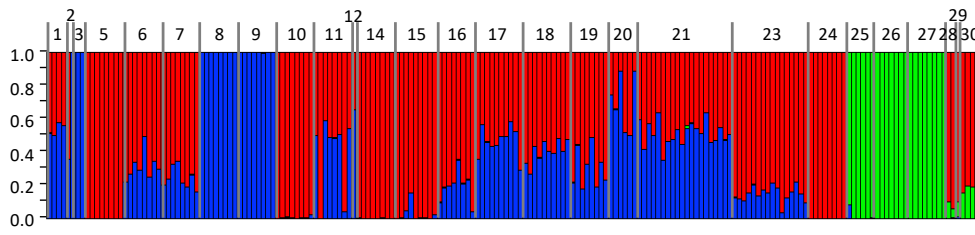
Table 2 Average (\pm SD) distances of mtDNA *cytb* sequences between four clades of *Aphyocypris chinensis* and *A. kikuchii*

	JP	C1	C2	TW
JP	<u>0.0023</u> \pm 0.0013	0.0311 \pm 0.0020	0.0645 \pm 0.0044	0.0344 \pm 0.0020
C1	0.0263 \pm 0.0014	<u>0.0039</u> \pm 0.0017	0.0607 \pm 0.0026	0.0286 \pm 0.0023
C2	0.0472 \pm 0.0023	0.0452 \pm 0.0015	<u>0.0079</u> \pm —	0.0680 \pm 0.0029
TW	0.0289 \pm 0.0014	0.0248 \pm 0.0017	0.0497 \pm 0.0015	<u>0.0048</u> \pm 0.0023

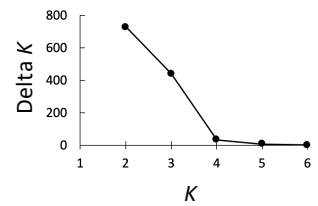
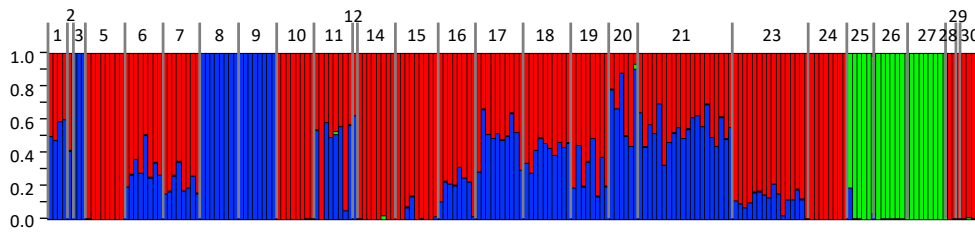
Underlined values along the diagonal are average uncorrected *p*-distances between haplotypes within clades

The uncorrected *p*-distance and HKY + G distance values lie below and above the diagonal, respectively

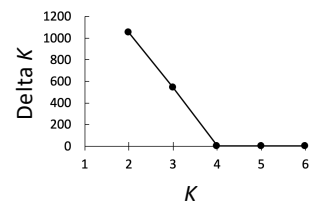
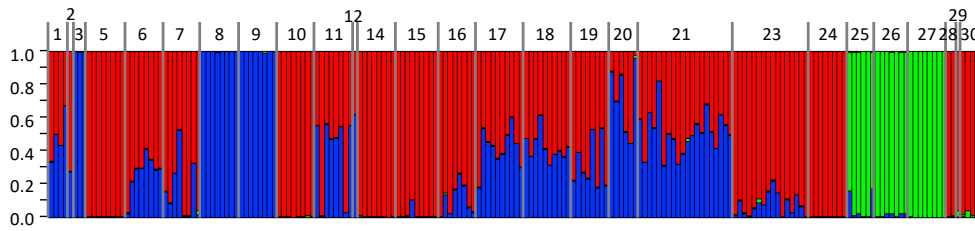
$r = 0.7, m = 10$ (359 SNPs)



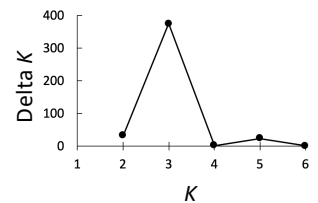
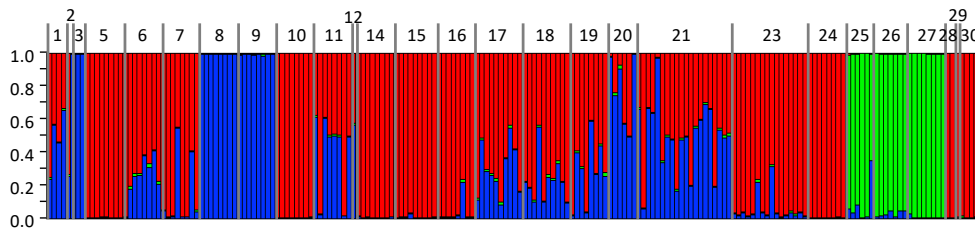
$r = 0.7, m = 15$ (240 SNPs)



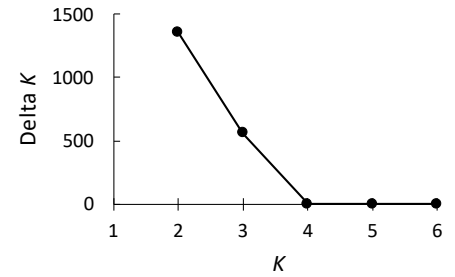
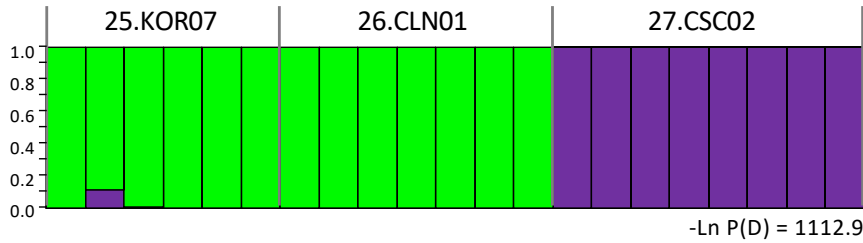
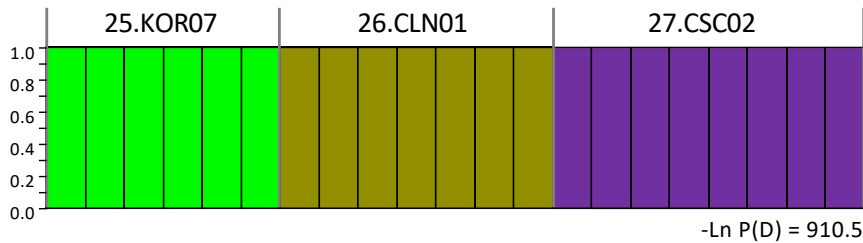
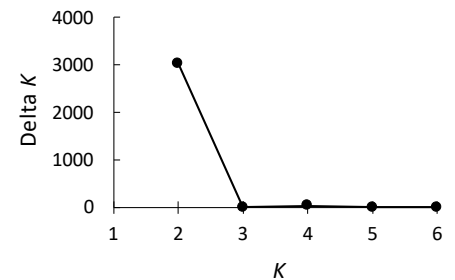
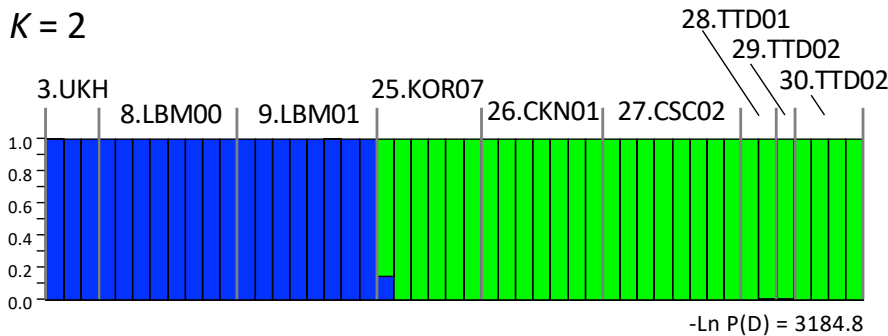
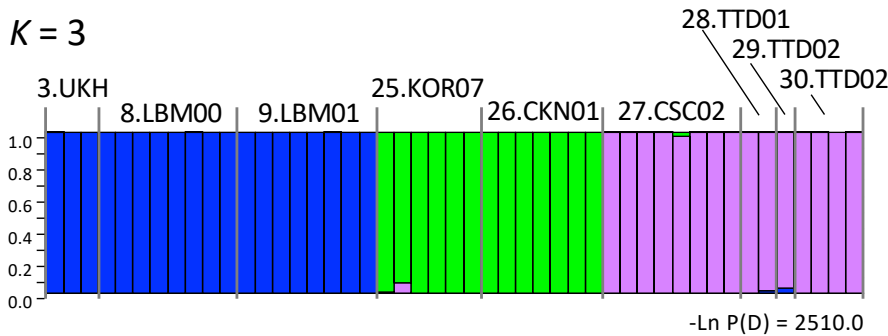
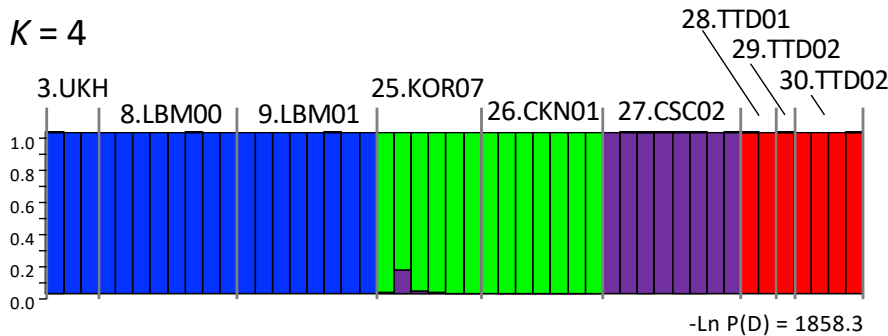
$r = 0.8, m = 20$ (102 SNPs)



$r = 0.9, m = 20$ (47 SNPs)



ESM Fig. S1 Results of admixture analysis for captive and wild populations of *Aphyocypris chinensis* and *A. kikuchii* based on selected parameter sets of the program ‘populations’ obtained using STRUCTURE. Results of the assignment for $K = 3$ and ΔK as a function of the number of assumed populations ($K = 1-6$) are shown. Each vertical bar represents an individual partitioned into the three clusters defined by STRUCTURE. Different colors represent three genetic elements; blue “Japanese chinensis,” green “continental chinensis,” and red “kikuchii” (see text). In the results with $K = 2$, the “Japanese chinensis” and “continental chinensis” elements were merged. For sample codes (above bars), see Fig. 1 and Table 1

a $K = 2$  $K = 3$ **b** $K = 2$  $K = 3$  $K = 4$ 

ESM Fig. S2 Results of admixture analysis for two subsets of *Aphyocypris chinensis* and *A. kikuchii* samples using STRUCTURE. **a** Results for three continental population samples ($n = 21$ specimens in total). **b** Results for samples except for the Japanese samples with the “kikuchii” element ($n = 47$ specimens in total). The assignments for $K = 2$ and 3 (a) and $K = 2-4$ (b), and ΔK as a function of the number of assumed populations ($K = 1-6$) are shown. The parameter setting of the program ‘populations’ is $r = 0.7$ and $m = 15$ (240 SNPs). Each vertical bar represents an individual partitioned into 2–4 clusters defined by STRUCTURE. Different colors represent different genetic elements. For the details of samples (above bars), see Fig. 1 and Table 1

ESM Table S1 mtDNA cytochrome *b* haplotype composition of wild and captive populations of *Aphyocypris chinensis* and comparative materials analyzed in this study

Sample #	Code	Locality or institution		Collection year	<i>n</i>	Haplotypes (<i>n</i>)				GEDIMAP ID
						JP	C1	C2	TW	
Wild										
	1	TNS08	Tanushimaru, Kurume, Fukuoka, Japan	T94-strain	2008	6	JPN-4(3)			KKC-1(3)
	2	TNS18	Tanushimaru, Kurume, Fukuoka, Japan	T94-strain	2018	1				KKC-1(1)
	3	UKH07+08	Ukiha, Fukuoka, Japan	Native	2007, 2008	3	JPN-3(2), JPN-5(1)			
	4	IZU11	Matsuzaki, Shizuoka, Japan	Introduced from unknow origin	2011	2				KKC-1(2)
	5	IZU18	Matsuzaki, Shizuoka, Japan	Introduced from unknow origin	2018	16				KKC-1(16)
Captive										
	6	HFC01	Hinamoroko Foster-parent Club; Fukuoka, Japan	T94-strain; used in Ohara et al. (2003)	2001	7	JPN-4(1)			KKC-1(6)
	7	HFC16	Hinamoroko Foster-parent Club; Fukuoka, Japan	T94-strain	2016	8	JPN-1(1)			KKC-1(7)
	8	LBM00	Lake Biwa Museum; Shiga, Japan; 1998 year-class of T94-strain	Genuine T94-strain, 1998-year class	2000	8	JPN-2(8)			
	9	LBM01	Lake Biwa Museum; Shiga, Japan	Genuine T94-strain; used in Ohara et al. (2003)	2001	4	JPN-2(4)			
	10	LBM_HCA00	Lake Biwa Museum; Shiga, Japan; From HCA	T94-strain from HCA (patially lacking information)	2000	7				KKC-1(7)
	11	LBM_HCA05	Lake Biwa Museum; Shiga, Japan	T94-strain mixed with HCA	2005	8				KKC-1(4), CR-1(4) ²
	12	LBM06	Lake Biwa Museum; Shiga, Japan	T94-strain mixed with HCA?	2006	1				KKC-1(1)
	13	LBM09_FS	Lake Biwa Museum; Shiga, Japan	T94-strain; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2009	11	JPN-4(1)			KKC-1(10)
	14	LBM17	Lake Biwa Museum; Shiga, Japan	TKU from HCA	2017	8				KKC-1(8)
	15	HCA01	Himeji City Aquarium; Hyogo, Japan	T82-strain from Biwako-bunkakan probably followed by introduction of an unknown strain; used in Ohara et al. (2003)	2001	2				KKC-1(2)
	16	SML01	Shima Marineland; Mie, Japan	T94-strain from MWU in 1999; used in Ohara et al. (2003)	2001	4				KKC-1(4)
	17	SML18o	Shima Marineland; Mie, Japan	T94-strain probably from LBM	2018	10				KKC-1(10)
	18	SML18n	Shima Marineland; Mie, Japan	T94-strain from MWU and LBM	2018	10				KKC-1(10)
	19	MWU01	Marine World Uminonakamichi; Fukuoka, Japan	T94-strain from HFC in 1995, 1997 and 2000; used in Ohara et al. (2003)	2001	7				KKC-1(8)
	20	KEE04	Kyushu Environmental Evaluation Association; Fukuoka, Japan	T82 and T94-strains; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2004	17	JPN-1(1), JPN-2(1), JPN-4(1)			KKC-1(14)
	21	CPR18	Clean Park Rinkai; Fukuoka, Japan	T94-strain and probably T82 from KEE and a conservation pond in Fukuoka (CPF)	2018	20	JPN-4(5)			KKC-1(15)
	22	OMC	Omura; Fukuoka, Japan	T94-strain probably mixed with HFC	2019	13				KKC-1(13)
	23	TKU10_FS	Tokai University; Shizuoka, Japan; via Research Institute of Environment, Agriculture and Fisheries, Osaka Prefecture	T94-strain from HFC and HCA?; Frozen sperm deposited in National Institute for Environmental Studies, Japan	2010	4				KKC-1(4)
	24	TKU18	Tokai University; Shizuoka, Japan	T94-strain from HFC and HCA?	2018	16				KKC-1(16)
	25	OKF16	OK Fish Farm; Saitama, Japan	Aquarium trade in Tokyo in 1990s	2016	4				KKC-1(4)
Outside of Japan										
	26	KOR07	Gunsan, Korea	Native	2007	10		CK-1(1), CK-3(1), CK-4(2), CK-5(1), CK6(2), CK-7(1), CK-8(2)		
	27	CLN01	Liaoning, China	Native; used in Ohara et al. (2003)	2001	8		CK-2(8)		
	28	CSC02	Sichuan, China	Native?; Donated from Institute of Hydrobiology, Chinese Academy of Sciences	2002	3			CSC-1(2), CSC-2(1)	
	29	TTD01	Taidong, Taiwan	Native (<i>A. kikuchii</i>)	2001	—				
	30	TTD02	Taidong, Taiwan	Native (<i>A. kikuchii</i>)	2002	—				
	31	THL02	Hualian, Taiwan	Native (<i>A. kikuchii</i>)	2002	2				KKC-2(1), KKC-3(1)
Total					220					

¹ LBM Lake Biwa Museum, Shiga, NIES National Institute for Environmental Studies, Tsukuba; NMNSF National Museum of Natural Science (Fish), Taichung

² CR Haplotype of the mtDNA control region

ESM Table S2 Percentage of alleles substituted between population pairs of *Aphyocypris chinensis* and *A. kikuchii*

	JPN_genuine	KOR+CLN	CSC	TWN	JPN_introduced
JPN_genuine (<i>n</i> = 38)	—				
KOR+CLN (<i>n</i> = 26)	7.8 (166)	—			
CSC (<i>n</i> = 16)	10.7 (169)	2.6 (153)	—		
TWN (<i>n</i> = 14)	16.1 (199)	6.7 (165)	10.8 (166)	—	
JPN_introduced (<i>n</i> = 32)	15.5 (200)	5.4 (166)	10.1 (169)	0.0 (201)	—
KOR+CLN+CSC+TWN (<i>n</i> = 56)	5.5 (183)	—	—	—	—

n Number of specimens

The percentage was calculated using loci with the missing rate < 0.5 in each population (the total number of loci is shown in parentheses)

JPN_genuine: 3_UKH07&08, 8_LBM00, 9_LBM01; *KOR+CLN*: 25_KOR07, 26_CLN01; *CSC*: 27_CSC02; *TWN*: 28_TTD01, 30_THL02; *JPN_introduced*: 05_IZU18, 24_OKF16

For the details of samples, see Fig. 1 and Table 1