Journal of Fish Biology (2010) **77**, 2072–2092
doi:10.1111/j.1095-8649.2010.02792.x, available online at wileyonlinelibrary.com

Phylogeography and conservation genetics of Lake Qinghai scaleless carp *Gymnocypris przewalskii*

D. M. O'Bryan†, Z. Xie*, Y. Wang*‡, J. Du*§, C. J. Brauner¶, J. G. Richards¶, C. M. Wood**, X.-Q. Chen* and B. W. Murray*†‖

**Division of Neurobiology and Physiology, Zhejiang University, College of Life Science, Hangzhou, Zhejiang 310058, Zijingang Campus, China, †University of Northern British Columbia, Natural Resources and Environmental Studies Institute, 3333 University Way, Prince George, British Columbia V2N 4Z9, Canada, ‡Queen's University, Department of Biology, 99 University Ave., Kingston, Ontario K7L 3N6, Canada, §Zhejiang University, School of Medicine, Hangzhou, Zhejiang 310058, Zijingang Campus, China, ¶University of British Columbia, Department of Zoology, 6270 University Blvd., Vancouver, British Columbia V6T 1Z4, Canada and **McMaster University, Department of Biology, 1280 Main St. West, Hamilton, Ontario L8S 4K1, Canada*

(Received 18 January 2010, Accepted 1 September 2010)

The objective of this study was to examine the spatial genetic relationships of the Lake Qinghai scaleless carp *Gymnocypris przewalskii* within the Lake Qinghai system, determining whether genetic evidence supports the current taxonomy of *Gymnocypris przewalskii przewalskii* and *Gymnocypris przewalskii ganzihonensis* and whether *Gymnocypris przewalskii przewalskii* are returning to their natal rivers to spawn. Comparison of mitochondrial (control region) variation (42 haplotypes in 203 fish) of *G. przewalskii* with the postulated ancestral species found in the Yellow River, *Gymnocypris eckloni* (10 haplotypes in 23 fish), indicated no haplotype sharing, but incomplete lineage sorting. Consistent with the sub-species status, an AMOVA indicated that the Ganzi River population was significantly different from all other river populations ($F_{ST} = 0{\cdot}1671$, $P < 0{\cdot}001$). No genetic structure was found among the other rivers in the Lake Qinghai catchment. An AMOVA of amplified fragment length polymorphism (AFLP) loci, however, revealed significant genetic differences between most spawning populations ($F_{ST} = 0{\cdot}0721$, $P < 0{\cdot}001$). Both mitochondrial and AFLP data found significant differences among *G. p. przewalskii*, *G. p. ganzihonensis* and *G. eckloni* (F_{ST} values of 0·1959 and 0·1431, respectively, $P < 0{\cdot}001$). Consistent with the incomplete lineage sorting, Structure analysis of AFLP loci showed evidence of five clusters. One cluster is shared among all sample locations, one is unique to *G. p. ganzihonensis* and *G. eckloni*, and the others are mostly found in *G. p. przewalskii.* Genetic evidence therefore supports the current taxonomy, including the sub-species status of *G. p. ganzihonensis*, and is consistent with natal homing of most Lake Qinghai populations. These findings have significant implications for the conservation and management of this unique and threatened species. The evidence suggests that *G. p. przewalskii* should be treated as a single population for conservation purposes. Exchangeability of the populations, however, should not be used to promote homogenization of fish spawning in the different rivers. As some degree of genetic divergence was detected in this study, it is recommended that the spawning groups be treated as separate management units.

Key words: AFLP; conservation genetics; *Gymnocypris eckloni*; *Gymnocypris przewalskii*; Lake Qinghai scaleless carp; mtDNA control region variation.

‖Author to whom correspondence should be addressed. Tel.: + 1 250 960 5638; email: murrayb@unbc.ca

INTRODUCTION

Lake Qinghai is a saltwater lake (salinity 9–11) located within the Qiadam Basin of the Qinghai–Tibetan Plateau of western China at an elevation of 3200 m. The lake initially formed during the late Pliocene to mid-Pleistocene [*c*. 500 000 years before present (B.P.)] as a result of the downfaulting of the Tibetan Plateau (Bian *et al*., 2000). Although the ancestral lake originally drained into the Yellow River, the connection was severed *c*. 130 000–150 000 B.P. after an uplift of the plateau (termed the Gonghe Movement). This turned the drainage inwards, forming the endorheic basin that comprises the lake today (Zheng, 1997; Li & Fang, 1998).

Although it is the largest lake in China (surface area 4600 km^2), few fish species reside in Lake Qinghai, in part due to extreme conditions (high pH and high salinity) (Zhu & Wu, 1975; Zheng, 1997). The most economically important of the fish species is the endemic scaleless or naked carp, *Gymnocypris przewalskii* (Kessler) (Cyprinidae: Schizothoracinae). *Gymnocypris przewalskii* were fished commercially from the 1950s until the 1990s, when a moratorium was put in place. As the fish are very slow growing, taking *c*. 7 years to reach sexual maturity and 10 years to reach a marketable weight of 500 g, stocks did not easily recover from overexploitation (Walker *et al*., 1996). *Gymnocypris przewalskii* is listed as Grade B on a list of species of concern by the China Biodiversity Conservation Action Plan (Zhan, 1994; Yue & Chen, 1998). A small government fishery continues to operate and there have been reports of ongoing poaching.

Gymnocypris przewalskii is unique among stenohaline cyprinids in that they undergo an annual migration from the saline lake to adjoining freshwater rivers to spawn, after which they return to the lake (Walker *et al*., 1996). The transition from saline water to freshwater and back to saline water presents significant physiological challenges to the fish, to which they are acclimated in a unique manner. Direct transfer of river fish to lake water caused significant physiological changes while allowing most plasma ions (Na^+, Cl^- and K^+) to equilibrate with the saline lake water (within 12 h of transfer from river water), thereby reducing the energy necessary for osmoregulation (Wood *et al*., 2007).

In addition to the difficulty in recovering from previous overexploitation of stocks, *G. przewalskii* have been presented with two other challenges. First, freshwater tributaries have been dammed or diverted for agricultural purposes. In combination with decreased rainfall, this has resulted in the lake becoming progressively shallower. Water levels have fallen at a rate of *c*. 10–12 cm per year over the last 50 years, resulting in increasing salinity levels (Walker & Yang, 1999). Increases in salinity have occurred since 1978, and current data indicate that this increase is accelerating (Wang *et al*., 2003; Yang *et al*., 2005; Wood *et al*., 2007). Although *G. przewalskii* are currently physiologically able to adapt to the salinity of the lake, experiments have shown that a slight increase in salinity levels would be toxic (C. Brauner, unpubl. obs.). Second, the damming and diversion of rivers have resulted in a significant loss of spawning habitat. Irrigation weirs were constructed on the Shaliu and Haergai Rivers in the 1950s, restricting access to spawning locations upstream (Walker & Yang, 1999). The Quanji River was later dammed *c*. 3 km from the mouth of the lake, and a fish ladder was not installed until 2006.

On the basis of the morphological characters, primarily the shape and number of gill rakers, a sub-species of the scaleless carp, *Gymnocypris przewalskii ganzihonensis*, has been identified in the Ganzi River (Zhu & Wu, 1975). While ancient Chinese historical records indicate that the Ganzi River may have once flowed directly into Lake Qinghai, a 1964 survey found that it lacked direct connection to the lake, probably due to the shrinking of the lake shoreline and low flows in the lower reaches of the river (Zhu & Wu, 1975). The current situation, which is heavily affected by irrigation, is similar to the conditions in 1964. The lower and middle reaches consist of marshes with very low flows, and large dunes block direct surface connection to the lake. As the current status of the population is unknown, an important objective of this study was to determine the genetic relationship between *G. p. ganzihonensis* and *G. p. przewalskii* (fish from Lake Qinghai and all other rivers in the Lake Qinghai system). The Ganzi River has been heavily affected by anthropogenic forces and information on this isolated fish population is critical to initiation of conservation actions.

Several different genetic markers have been used in the study of fish conservation. This study uses the control region of mitochondrial (mt)DNA in combination with amplified fragment length polymorphisms (AFLP). There are an increasing number of fish studies that combine mitochondrial variation with nuclear markers, such as AFLPs (Young *et al*., 2001; Seehausen *et al*., 2003; Ishikawa *et al*., 2004; Mock & Miller, 2005). The mtDNA control region is mostly selectively neutral (non-coding) and provides information on non-selective processes, such as mutation and drift. Mitochondrial DNA is frequently used in population genetic structure studies (Brunner *et al*., 1998; Stepien & Faber, 1998; Taylor *et al*., 2001; Gharrett *et al*., 2005) and to assess sub-species (Tabata & Taniguchi, 2000; April & Turgeon, 2006). AFLPs, developed by Vos *et al*. (1995), allow analysis of multiple loci across the entire nuclear genome. These markers are capable of identifying relationships at the species or population level, providing the resolving power to detect the early divergence of lineages (Sullivan *et al*., 2004). AFLPs are, therefore, especially well suited to studies of organisms with very little prior genetic information, as in *G. przewalskii.*

The main objective of this study was to assess the spatial genetic structure of *G. przewalskii* within the Lake Qinghai system. Samples were collected from each of the major rivers that drain into Lake Qinghai and from the Ganzi River. To serve as an out-group, samples of *Gymnocypris eckloni* Herzenstein, the most closely related species, were collected from the Yellow River. This sampling design allowed genetic evidence for the existence of *G. p. ganzihonensis* (Ganzi River sub-species) to be placed into the context of genetic divergence from the closest out-group to *G. przewalskii*. It also allowed the assessment of the degree of structure found among fish caught in the rivers draining into the lake (*i.e. G. p. przewalskii*). Identification of a stock structure within the lake would indicate that *G. p. przewalskii* preferentially spawn in the same rivers in which they were born (natal homing). The ability to identify different stocks within Lake Qinghai is also extremely important for the conservation and management of *G. przewalskii* populations both within the lake and in each river. Some tributaries, and associated spawning habitat, may be lost due to falling lake levels, agricultural divergence and damming. Findings from this study will have significant implications for the conservation and management of this threatened species.

MATERIALS AND METHODS

Lake Qinghai was visited in June 2002, 2004 and 2005 during the spawning run of *G. przewalskii* (Fig. 1). Fish samples were collected under permits issued by local and national authorities and in accordance with the national animal care regulations. An animal care and use approval was obtained for this research from Queen's University, Ontario, Canada. Adult fish were collected from most rivers using a cast net; however, juvenile fish were collected from the Ganzi River as adult fish were not observed. Caudal fin clips [preserved in a dimethyl sulphoxide (DMSO) salt solution (Seutin *et al*., 1991)] were mostly taken from live fish, which were then released, although a small number of fish were lethally sampled. Samples were taken from one location each in the Heima, Quanji, Shaliu, Haergai and Ganzi Rivers, and from two locations in the Buha River (the largest tributary). Sample numbers and geographic co-ordinates are shown in Fig. 1. *Gymnocypris eckloni* samples were collected by a local fisherman from the upper reaches of the Yellow River near the town of Madoi in June 2004 ($n = 23$ and 38 for mtDNA and AFLP analyses, respectively).

DNA was extracted from tissue samples using a standard phenol–chloroform method followed by ethanol precipitation and resuspension in a standard Tris-EDTA solution (Sambrook & Russell, 2001).

MTDNA ANALYSIS

Approximately 50 ng of genomic DNA was added to 24 μl of polymerase chain reaction (PCR) master mix containing 1× PCR buffer, 250 μM of each deoxy-nucleoside triphosphate (dNTP), 2 mM $MgCl_2$, 0·2 μM of each primer (ProL1 and CRH1; Table I) and 1 U of *Taq*

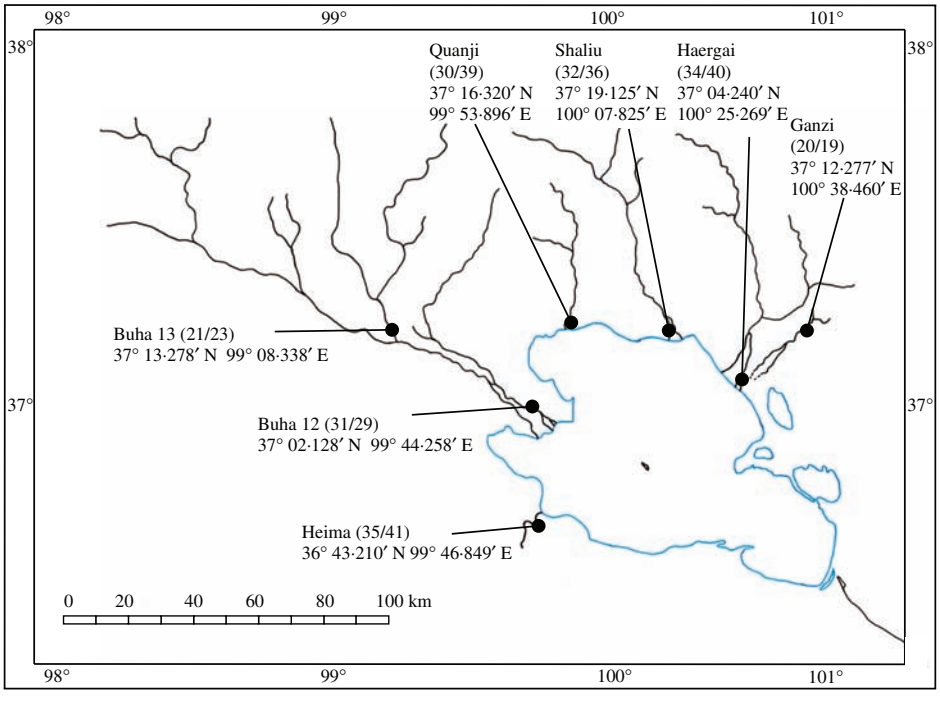

FIG. 1. Map of the Lake Qinghai catchment showing sampling locations of *Gymnocypris przewalskii przewalskii* (Heima, Buha 12, Buha 13, Quanji, Shaliu and Haergai) and *Gymnocypris przewalskii ganzihonensis* (Ganzi). For each location, the latitude and longitude and the sample size used for the mtDNA and AFLP analyses, respectively, are shown.

TABLE I. Oligonucleotides used in mitochondrial and amplified fragment length polymorphism (AFLP) analyses of *Gymnocypris przewalskii przewalskii* and *Gymnocypris przewalskii ganzihonensis*

Oligonucleotides	Sequence
mtDNA	
ProL1	5′ CCTAGCTCCCAAAGCCAGA 3′
CRH1	5′ TGTTGGCTGATACGTTCTTGG 3′
SProL1	5′ TTCTGAACTAAACTATTCTCTGG 3′
SH1	5′ CCATTTACTGAGCGTAGGG 3′
AFLP	
Adapters	
*Eco*RI-adapter-fwd	5′ CTCGTAGACTGCGTACC 3′
*Eco*RI-adapter-rev	5′ AATTGGTACGCAGTCTAC 3′
*Mse*I-adapter-fwd	5′ GACGATGAGTCCTGAG 3′
*Mse*I-adapter-rev	5′ TACTCAGGACTCAT 3′
Pre-selective primers	
*Eco*RI-A	5′ GACTGCGTACCAATTCA 3′
*Mse*I-C	5′ GATGAGTCCTGAGTAAC 3′
Selective primers	
E-AAG	5′ GACTGCGTACCAATTCAAG 3′
E-ACG	5′ GACTGCGTACCAATTCACG 3′
M-CAA	5′ GATGAGTCCTGAGTAACAA 3′
M-CAC	5′ GATGAGTCCTGAGTAACAC 3′
M-CAG	5′ GATGAGTCCTGAGTAACAG 3′

(Invitrogen; www.invitrogen.com). The amplification programme was 94° C for 4 min, 47° C for 1 min and 72° C for 2 min, followed by 39 cycles of 94° C for 30 s, 47° C for 30 s and 72° C for 90 s, ending with 72° C for 8 min. Amplifications were carried out on a DYAD or PTC-100 thermal cycler (MJ Research, Inc.; www.bio-rad.com). The resulting DNA was purified using the Wizard PCR Preps DNA Purification System (Promega; www.promega.com). DNA sequencing was performed using the DTCS Quick Start Kit (Beckman Coulter, Inc.; www.beckmancoulter.com) with 1·6 μM SProl or SH1 primers (Table I). Sequence fragments were separated on the CEQ 8000 Genetic Analysis System (Beckman Coulter) using the LFR-1 method.

Data analysis

Sequence chromatograms were viewed and edited in Sequencher 4.2 (Gene Codes; www.genecodes.com). To confirm variation, haplotypes were sequenced in both directions (with both SProL1 and SH1 primers) and, in many cases, multiple individuals (Fig. 2). Sequences reported here have been deposited in the Genbank database under accession numbers EU150135–EU150186.

Sequences were aligned with ClustalW implemented in MEGA 3.1 (Kumar *et al*., 2004). An analysis of molecular variance (AMOVA) was performed in Arlequin 3.11 [DNA data type; genotypic data set to 0; 1000 permutations (Excoffier *et al*., 2005)]. Three hierarchical analyses were conducted by first grouping the data in accordance with the currently accepted taxonomy and then sequentially removing the *G. eckloni* and *G. p. ganzihonensis* groups. Population statistics, including haplotype diversity, nucleotide diversity, and estimates of θ ($\theta = 2N_f\mu$, where N_f is the female effective population size and μ the mutation rate) were also calculated in Arlequin. Phylogenetic relationships among the haplotypes were studied using neighbour-joining (Kumar *et al*., 2004), parsimony networks (Clement *et al*., 2000) and maximum likelihood methods (Guindon & Gascuel, 2003). jModeltest 0.1.1 (Posada, 2008)

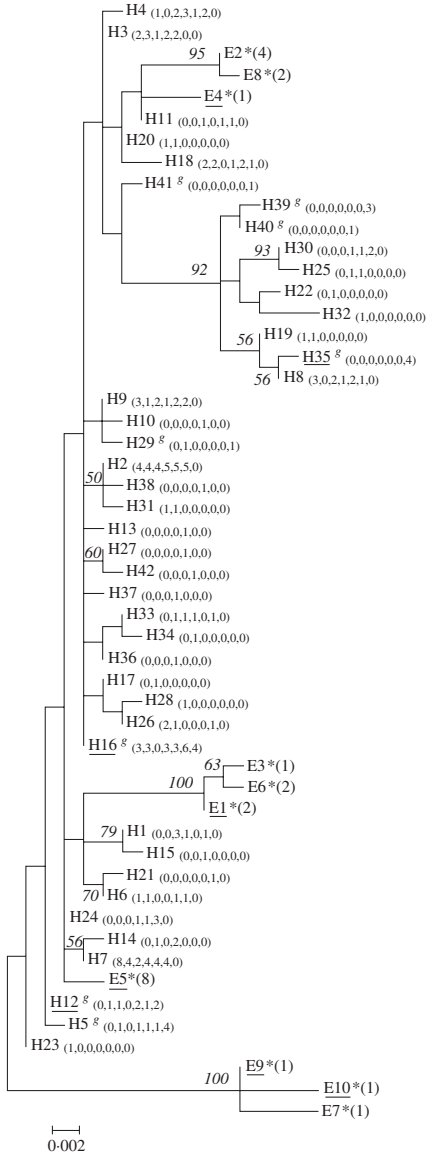

FIG. 2. Maximum likelihood tree of *Gymnocypris przewalskii* and *Gymnocypris eckloni* mitochondrial DNA haplotypes identified by sequencing the control region. The best-fit model of nucleotide substitutions identified using the Bayesian Information Criterion (BIC) implemented in jModeltest 0.1.1 (Posada, 2008), Hasegawa, Kishino and Yano (HKY) with unequal base frequencies and a proportion of invariable sites, was used by the programme PHYML (Guindon & Gascuel, 2003) to estimate the above phylogeny and calculate bootstrap values. Bootstrap values greater than 50% are shown. A model-averaged phylogeny (Posada, 2008) using relative model weights based on either the BIC or Akaike Information Criterion (AIC) is identical in topology to the tree shown with all nodes having a model support of 1·0. H indicates *G. przewalskii* and E (with asterisks) indicates *G. eckloni* haplotypes. Haplotypes found in the Ganzi River are indicated by superscript *g*. The sample size of each haplotype is shown in brackets. For the *G. przewalskii* haplotypes, the sample size in each sample location is shown in the order of Heima, Buha12, Buha13, Quanji, Shaliu, Haergai and Ganzi, respectively. Underlined haplotypes are found in samples that are members of an amplified fragment length polymorphism (AFLP) cluster found in the Yellow and Ganzi Rivers (Fig. 3).

was used to find the best-fit model of nucleotide substitution and to calculate a maximum likelihood model-averaged phylogeny.

AMPLIFIED FRAGMENT LENGTH POLYMORPHISM ANALYSIS

The AFLP protocol was based on Vos *et al.* (1995) and Hawkins *et al.* (2005) with some modifications. The locations and the numbers of samples used are shown in Fig. 1. All adapter and primer sequences are listed in Table I.

Two hundred nanograms of high-quality genomic DNA were digested (3 h at 37° C) with 10 U each of *Mse*I and *Eco*RI, 1× *Eco*RI restriction buffer and 2 μg of bovine serum albumin (BSA) (New England Biolabs; www.neb.com) in a 20 μl reaction. A 20 μl ligation master mix containing 3 U (NE Units) of T4 DNA ligase, 1× T4 ligase buffer (New England Biolabs) and 1·5 μl each of 50 μM adapters (*Mse*I and *Eco*RI) was added to the restriction digest and incubated at 16° C for at least 16 h.

Pre-selective amplifications were performed by combining 5 μl of diluted digestion–ligation product (1:5 in nuclease-free water) with 20 μl of a pre-selective master mix containing 1× PCR buffer, 1·5 mM $MgCl_2$, 200 μM of each dNTP, 0·8 μM each of the *Eco*RI-A and *Mse*I-C pre-selective primers and 1·25 U of *Taq* polymerase (Invitrogen). The amplification programme was 72° C for 2 min, followed by 20 cycles of 94° C for 30 s, 56° C for 1 min, and 72° C for 1 min, ending with 60° C for 30 min.

Selective amplification was subsequently performed by combining 5 μl of diluted pre-amplification product (1:10 in nuclease-free water) with 15 μl of a selective master mix containing 1× PCR buffer, 1·5 mM $MgCl_2$, 200 μM of each dNTP, 0·4 μM each of the *Eco*RI and *Mse*I selective primers and 1 U of Platinum *Taq* (Invitrogen). Amplifications were carried out with an initial activation of 94° C for 30 s, followed by 12 cycles of 94° C for 30 s, 65° C (decreased by 0·7° C per cycle) for 30 s and 72° C for 1 min. The preceding cycles were followed by 32 cycles of 94° C for 30 s, 56° C for 30 s and 72° C for 1 min, ending with a final extension of 72° C for 10 min. The selective primer combinations used in this study were E-ACG–M-CAA, E-ACG–M-CAG and E-AAG–M-CAC.

Repeatability was checked by randomly choosing samples to run in triplicate prior to the digestion–ligation procedure (one sample in 40 was run in triplicate). Repeatability was further assured by running each digestion–ligation product in duplicate through the pre-amplification and selective amplification procedures.

Data analysis

AFLP fragments were sized using the CEQ8000 Genetic Analysis system (Beckman Coulter) and viewed in the Fragment Analysis Module [600 base pairs (bp) size standard; quartic model; PA ver.1 dye mobility calibration]. Only reliable loci, which had strong intensities, were clear and crisp and were greater than 80 bp in size, were included in the analysis. Any fragments that were observed in blank samples were removed from analysis, although their presence does not necessarily indicate contamination, but probably results from primer dimers and interactions between adapters (Bensch & Åkesson, 2005). To further increase confidence in the data set, each sample's replicate was scored, and if a fragment did not occur in both replicates, it was also removed from the analysis, although this rarely occurred. Fragments were binned using the AFLP analysis feature within the fragment analysis module (1 nucleotide bin width and the *Y* threshold set to 0). Similar fragment sizes were assumed to represent homology. The data were represented as a binary matrix (1 indicates the presence of the fragment while 0 indicates its absence).

AFLP data were examined on an individual level using Structure 2.2 adapted to dominant markers (Falush *et al.*, 2007). Correlated allele frequencies and admixed population origins were assumed. Burn-in and run lengths were set to 10 000. The number of inferred groups (k) ranged from 2 to 8, and each possible k value was run 20 times. The best value of k was determined objectively using the method of Evanno *et al.* (2005). Hierarchical AMOVA and pair-wise F_{ST} comparisons were carried out using Arlequin 3.11 [restriction fragment length polymorphism (RFLP) data type; genotypic data and gametic phase set to 0; recessive data set to 1; 1000 permutations; assumed group structure based on current taxonomy] (Excoffier *et al.*, 2005). Principal co-ordinate analysis of pair-wise sample location genetic distance (Nei,

1972) was carried out in GenAlEx 6 (Peakall & Smouse, 2006). Per cent polymorphism was calculated in Tools for Population Genetic Analysis (TFPGA) 1.3 (Miller, 1997).

As adult fish were not observed in the Ganzi River at the time of sampling, only juvenile fish were collected. To test for the possible effects of sampling siblings (and thereby skewing the results), the average pair-wise genetic distance (Nei, 1972) for AFLP data was compared for individuals with shared mtDNA haplotypes (which are more likely to be siblings) *v*. those that do not share mtDNA haplotypes. If siblings were sampled, the expectation is that the average genetic distance will be less for individuals that share haplotypes compared to those that do not.

RESULTS

MTDNA ANALYSIS

Sequences reported in a preliminary analysis [Genbank accession numbers AY850398 and AY850397 (Xie *et al*., 2006)] were combined with sequences from this study, resulting in the identification of 42 distinct *G. przewalskii* haplotypes (Fig. 2). Ten *G. eckloni* haplotypes were also identified, none of which were shared with *G. przewalskii*. From a 701 bp alignment of 52 haplotypes, jModeltest was used to find the best-fit model of nucleotide evolution (Posada, 2008). The best-fit model was used to estimate a maximum likelihood tree (Fig. 2) and a rank weighting of models to estimate a model-averaged maximum likelihood tree. The model-averaged method (Posada, 2008) investigates the sensitivity of the tree topology to the model used. The results indicate that the tree is relatively insensitive to the model used. All nodes were supported by model-averaged weights of 1·0, indicating that the same topology was found in all the top-weighted best-fit models.

The maximum likelihood tree of the *G. przewalskii* and *G. eckloni* haplotypes showed incomplete lineage sorting (*i.e*. the *G. eckloni* haplotypes do not form a monophyletic group). Incomplete lineage sorting was also noted when neighbour-joining trees and parsimony networks were constructed (not shown). The topology of the maximum likelihood tree was very similar to the parsimony network in that many haplotypes differed by a single nucleotide change and polytomies were common. Basal haplotypes (inferred ancestral haplotypes from which the other branches radiate) are shown at nodes with no branch length. There are three major clusters and one singleton of *G. eckloni* haplotypes that occur among clusters of *G. przewalskii* haplotypes. There do not appear to be any clear geographical relationships among the *G. przewalskii* haplotypes. Of the 42 *G. przewalskii* haplotypes, 19 were unique to only one population (Table II). Of the eight haplotypes detected in the Ganzi River fish, four (50%) were unique to that river, even though the sample size was smaller than from the other rivers that had between 22 and 6% unique haplotypes. The four Ganzi haplotypes that were shared with the other populations generally represented frequently occurring haplotypes in the lake (Fig. 2).

TABLE II. Number of mtDNA haplotypes and number of unique haplotypes (per cent in brackets) found within *Gymnocypris przewalskii* sampling locations

	Heima	Buha 12	Buha 13	Quanji	Shaliu	Haergai	Ganzi
Haplotype number	16	20	12	17	18	17	8
Unique haplotypes	3 (19)	3 (15)	1 (8)	3 (18)	4 (22)	1 (6)	4 (50)

TABLE III. Results of hierarchical AMOVA for both mtDNA and amplified fragment length polymorphism (AFLP) markers with populations ($n = 8$) grouped according to the currently accepted taxonomy [three groups: *Gymnocypris przewalskii przewalskii* (6), *Gymnocypris przewalskii ganzihonensis* (1) and *Gymnocypris eckloni* (1)]. *F* statistics are generated in Arlequin 3.11 (Excoffier *et al.,* 2005). Degrees of freedom (d.f.), sum of squares (*S*), variance components (*V*), per cent variance (% *V*) and the *P* values (*P*) of % *V* are shown. F_{CT} is the variance among groups. F_{ST} with group structure is the fraction of total variance found among populations within groups plus the variance among groups. $F_{ST} = 0{\cdot}1959$, $P < 0{\cdot}001$ for mtDNA and $F_{ST} = 0{\cdot}1431$, $P < 0{\cdot}001$ for AFLP markers

	mtDNA					AFLP				
	d.f.	*S*	*V*	% *V*	*P**	d.f.	*S*	*V*	% *V*	*P**
Among groups	2	51·00	0·648	20·84	< 0·001 (0·0420)	2	79·62	0·480	7·76	< 0·05 (0·0420)
Among popolution within groups	5	6·62	0·039	−1·25	> 0·05 (0·0000)	5	96·23	0·405	6·55	< 0·001 (0·0000)
Within population	218	544·83	2·499	80·41	< 0·001 (0·0000)	257	1361·8	5·217	85·69	< 0·001 (0·0000)

*Given the specified group structure, there is a high probability of recreating the tested structure; therefore, *P*-values where the random values are greater than the observed values are reported with *P*-values of random values equal to the observed in brackets.

TABLE IV. Pair-wise F_{ST} values for *Gymnocypris przewalskii* and *Gymnocypris eckloni* sampling locations from mtDNA haplotypes (above the diagonal) and amplified fragment length polymorphism (AFLP) genotypes (below the diagonal). Only significant values are shown (all $P < 0{\cdot}001$, except those marked with an asterisk where $P < 0{\cdot}05$) and non-significant values are indicated by NS

	Heima	Buha 12	Buha 13	Quanji	Shaliu	Haergai	Ganzi	Yellow
Heima	—	NS	NS	NS	NS	NS	0·1251*	0·1640
Buha 12	0·1532	—	NS	NS	NS	NS	0·1565	0·1696
Buha 13	0·1165	0·1356	—	NS	NS	NS	0·1003*	0·1302
Quanji	0·0367	0·0970	0·0435	—	NS	NS	0·1923	0·1874
Shaliu	0·0303	0·1504	0·0547	NS	—	NS	0·1524	0·1756
Haergai	0·0656	0·1341	0·0588	0·0157*	0·0284	—	0·1691	0·1771
Ganzi	0·1654	0·1571	0·2484	0·1343	0·1884	0·1564	—	0·1801
Yellow	0·1150	0·1674	0·1940	0·1089	0·1199	0·1025	0·0956	—

AMOVA results indicated significant differences among *G. p. przewalskii*, *G. p. ganzihonensis* and *G. eckloni* ($F_{ST} = 0{\cdot}1959$, $F_{CT} = 0{\cdot}2083$) when populations were grouped according to the current taxonomy (Table III). Significant differences ($P < 0{\cdot}001$) were also noted when comparing *G. p. ganzihonensis* to *G. p. przewalskii* (*i.e.* two groups, results not shown: $F_{ST} = 0{\cdot}1671$, $F_{CT} = 0{\cdot}1781$). Removal of *G. p. ganzihonensis* to examine only differences among sample locations within the current lake drainage revealed no evidence for genetic structure (*i.e.* F_{ST} not significantly different from 0). These relationships are further illustrated by pair-wise comparisons of F_{ST} (among sample locations) values where *G. p. ganzihonensis* and *G. eckloni* were found to be significantly different from all *G. p. przewalskii* samples

(pair-wise F_{ST} from 0·1003 to 0·1923; Table IV). Pair-wise comparisons between the *G. p. przewalskii* samples from the other rivers, however, did not exhibit significant differences (*i.e.* pair-wise F_{ST} values ranged from −0·0244 to 0·0063; all P values >0·05).

Population statistics were calculated with samples grouped based on the current taxonomy (Table V). θ values ranged from 3·94 to 14·30 and gave estimates of female effective population sizes ranging from 19 700 to 71 500.

AFLP ANALYSIS

A total of 41 polymorphic loci were included in this analysis, with an average polymorphism across both species of 77·3%. All 41 loci were polymorphic in the *G. eckloni* samples, whereas 39 were polymorphic within the *G. przewalskii* samples. The Ganzi River sample location exhibited the smallest level of polymorphism (Table VI).

Structure analysis indicated that the greatest rate of change in likelihood occurs between $k = 3$ and $k = 4$. Applying the method of Evanno *et al*. (2005), an optimal k value of 5 was found. The results of the Structure analysis are best illustrated by comparing the results of $k = 3$ and $k = 5$ (Fig. 3). No direct association of sampling locations with inferred groups was observed. *Gymnocypris przewalskii ganzihonensis* and *G. eckloni* individuals were primarily contained within two shared clusters, one mostly unique to these locations (green) and one shared with all *G. p. przewalskii* sample locations (blue). In contrast, the six *G. p. przewalskii* sample locations each had a number of individuals belonging to a third cluster at $k = 3$ (red). Increasing k values to 4 or 5 split this third cluster into additional groups (red, yellow and pink) while primarily maintaining the integrity of the first two. At $k = 5$, two *G. eckloni* samples were included into the red and pink clusters commonly found in the *G. p. przewalskii* sample locations. Stock structure among *G. p. przewalskii* sample locations may also be illustrated by the frequency differences of cluster membership at each location. For example the percentage of the yellow cluster, unique to *G. p. przewalskii* samples, varies from 2 to 43% among these locations, with the greatest percentage being found in the two Buha River sites (24 and 43%).

One cluster (green) was unique to the *G. eckloni* and *G. p. ganzihonensis* samples at $k = 5$. The members of this cluster were compared to the results of the mtDNA analysis. Underlined haplotypes (Fig. 2) were found in individuals with AFLP genotypes that belong to this cluster. No correlation between mtDNA and AFLP genotypes was noted. In the three *G. p. ganzihonensis* samples belonging to this cluster, two had common, but divergent, *G. p. przewalskii* mtDNA haplotypes while one was unique to the Ganzi. Similarly, *G. eckloni* mtDNA haplotypes in this cluster were found in each of the four *G. eckloni* mtDNA lineages.

Similar to the mtDNA results, AMOVA of the AFLP data showed significant differences ($P < 0·05$) among all groups (Table III: $F_{ST} = 0·1431$, $F_{CT} = 0·0776$) and between *G. p. przewalskii* and *G. p. ganzihonensis* (results not shown: $F_{ST} = 0·1681$, $F_{CT} = 0·1031$) when grouped according to the current taxonomy. In contrast, significant differences among sample locations within *G. p. przewalskii* were also found ($F_{ST} = 0·0721$, $P < 0·001$). Pair-wise comparisons of F_{ST} values between all sampling locations found all but the Quanji and Shaliu samples to be significantly different from each other (Table IV). Pair-wise F_{ST} values for the Ganzi River fish

TABLE V. Population statistics generated using mtDNA

Species	$h \pm$ S.D.	$\pi \pm$ S.D.	$\theta_{(k)}$ (95% C.I.)	$\theta_{(H)} \pm$ S.D.	$\theta_S \pm$ S.D.	$\theta_{(\pi)} \pm$ S.D.	N_f
Gymnocypris przewalskii przewalskii	0·934 ± 0·008	0·006 ± 0·003	14·30 (9·73–20·69)	12·65 ± 1·80	6·22 ± 1·68	4·21 ± 2·32	71 500–21 050
Gymnocypris przewalskii ganzihonensis	0·884 ± 0·036	0·008 ± 0·005	4·44 (1·84–10·38)	6·43 ± 2·50	3·94 ± 1·65	5·86 ± 3·26	32 150–19 700
Gymnocypris eckloni	0·854 ± 0·057	0·015 ± 0·008	6·18 (2·75–13·62)	4·79 ± 2·40	10·57 ± 3·78	10·26 ± 5·42	52 850–23 950

h, haplotype diversity; π, nucleotide diversity; $\theta_{(k)}$, theta estimated from the expected number of alleles (k) and sample size (n) (Ewen, 1972); $\theta_{(H)}$, theta estimated based on the expected homozygosity (H) (Zouros, 1979; Chakroborty & Weiss, 1991); $\theta_{(S)}$, theta estimated from the number of segregating sites (s) and sample size (n) (Watterson, 1975); $\theta_{(\pi)}$, theta estimated from the mean number of pair-wise differences (Tajima, 1983); N_f, range of female effective population sizes calculated using previous estimates of θ.

TABLE VI. Per cent polymorphism of amplified fragment length polymorphism (AFLP) fragments in *Gymnocypris* spp. samples from each sampling location

Sampling location	% Polymorphic loci
Heima	66·0
Buha 12	62·3
Buha 13	58·5
Quanji	64·2
Shaliu	62·3
Haergai	71·7
Ganzi	56·6
Yellow	67·9

were generally higher (0·1343–0·2484) when compared to the values found between *G. p. przewalskii* sampling locations.

To further explore the differences among sample locations, a principal co-ordinate analysis of pair-wise genetic distances was conducted (Fig. 4). Most of the Lake Qinghai locations were found to group in the same quadrant, whereas the Buha sample locations were more distant and in a separate quadrant. The *G. p. ganzihonensis,* Ganzi, and *G. eckloni*, Yellow, sample locations were also each in separate quadrants and distant from all other locations.

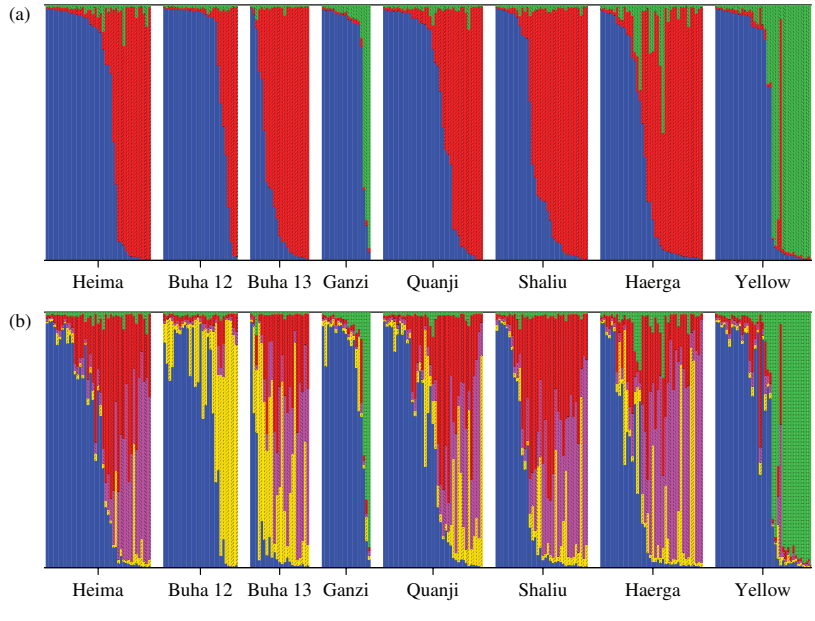

FIG. 3. Genetic structure estimated from amplified fragment length polymorphism (AFLP) markers for $k = 3$ (A) and $k = 5$ (B) using Structure 2.2 (Falush *et al.*, 2007). *Gymnocypris* individuals are grouped by sampling location; Heima, Buha 12, Buha 13, Quanji, Shaliu and Haergai for *Gymnocypris przewalskii przewalskii*; Ganzi for *Gymnocypris przewalskii ganzihonensis* and Yellow for *Gymnocypris eckloni.* Sample order is the same at both $k = 3$ and $k = 5$.

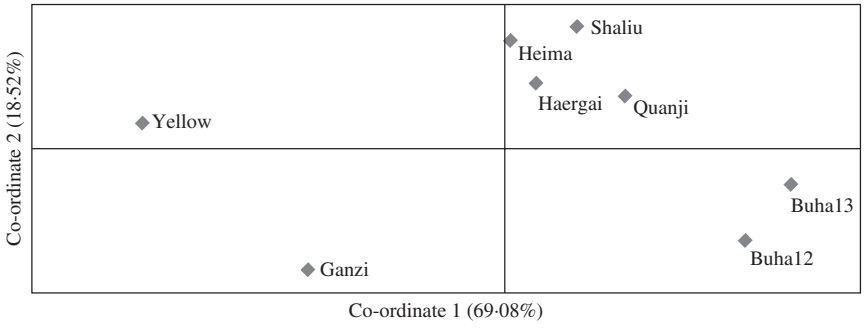

FIG. 4. Principal co-ordinate analysis of pair-wise genetic distance matrix (Nei's genetic distance) among sampling locations using GenAlEx 6 (Peakall & Smouse, 2006). The first two co-ordinates explain 87·61% of the variation. *Gymnocypris* sampling locations are Heima, Buha 12, Buha 13, Quanji, Shaliu and Haergai for *Gymnocypris przewalskii przewalskii*; Ganzi for *Gymnocypris przewalskii ganzihonensis* and Yellow for *Gymnocypris eckloni.*

TEST FOR THE EFFECTS OF SAMPLING JUVENILES IN THE GANZI RIVER

When sampling small populations of juveniles, collecting sibling groups is a possibility. If siblings were sampled in this case, it is expected that the average genetic distance (calculated using AFLP data, not shown) would be less for individuals that shared mtDNA haplotypes compared to those that did not. The average pair-wise genetic distance (calculated using AFLP data) for individuals that shared mtDNA haplotypes (0·011), however, was not significantly different from the average genetic distance for individuals that did not share mtDNA haplotypes (0·012) ($t = -1{\cdot}131$, $P > 0{\cdot}05$), indicating that sibling groups were probably not sampled in the Ganzi River.

DISCUSSION

The results of both the mtDNA and ALFP analyses indicated a population structure that is consistent with the current taxonomy. Among sample locations collected within the existing Lake Qinghai catchment, only the AFLP data indicated the existence of genetic structure. Crandall *et al.* (2000) stressed that both genetic and ecological data (both recent and historical in nature) should be used in defining evolutionary significant units (ESU). ESUs should lack both ecological and genetic exchangeability. Lack of ecological exchangeability refers to different selective pressures and adaptations between populations. These might include differences in life-history traits, habitats, predators or morphology. Lack of genetic exchangeability refers to lack of recent gene flow between populations and the observation of genetic differences consistent with geographic divisions.

RELATIONSHIP BETWEEN *GYMNOCYPRIS ECKLONI* AND *G. PRZEWALSKII*

To place the variation seen among *G. przewalskii* sampling locations into context, it is important to assess the variation found between *G. przewalskii* and its most

closely related species, *G. eckloni*. *Gymnocypris przewalskii* and *G. eckloni* do not share any mtDNA control region haplotypes; however, there is incomplete lineage sorting of the haplotypes (Fig. 2), a result that was also found in previous comparisons of the two species using cytochrome *b* sequences (Zhao *et al*., 2005). Using the high and low values for the *G. p. przewalskii* female effective population size (Table V) and a generation time of 7 years (Walker *et al*., 1996), a coalescence time for all haplotypes $[E(T)]$ can be estimated using the formula, $E(T) = 2N_f[(n-1)/n]G$ (N_f is the female effective population size, n the number of haplotypes and G the generation time). The estimated time for all haplotypes to coalesce was 41 097–139 595 generations, or 287 686–977 166 years. The estimated molecular divergence time reported using cytochrome *b* sequences between the fish of Lake Qinghai and the Yellow River also coincided with the estimated time of separation of the two water bodies during the Gonghe movement (Qi *et al*., 2007). Given that the lowest estimate for all haplotypes in Lake Qinghai or the Yellow River to coalesce was *c*. 287 686 years and that the minimum estimated time for separation of the two water bodies was 130 000 years, incomplete lineage sorting in Lake Qinghai is not unexpected. Indeed, this would be consistent with a relatively recent speciation event occurring sometime during the separation of the Yellow River and Lake Qinghai (Zhao *et al*., 2005; Qi *et al*., 2006).

The results of the Structure analysis of AFLP genotypes were also consistent with a recent ancestry of *G. przewalskii*. As no demographic exchange between the species is possible (and no anthropogenic exchange suspected), the AMOVA results also show a divergence of the nuclear genotypes of the species. Lineage sorting, however, is not complete, as noted by the high rate of sharing of polymorphic AFLP markers (39 of 41). Given the larger effective population size for nuclear genes *v*. mtDNA, lineage sorting of polymorphic AFLP markers is not expected. The sharing of clusters and markers is more than likely a result of a recent common ancestry. Furthermore, the *G. eckloni* samples were caught in the main stream of the Yellow River. The existence of multiple spawning runs within this (and possibly other, see below) sampling location would also further increase the complexity of the Structure results.

According to the definition of Moritz (1994), the lack of reciprocal monophyly at mtDNA markers would prevent the distinction of *G. eckloni* and *G. przewalskii* as separate ESUs (or species). This precludes the recognition of ESUs until completion of lineage sorting. Alternatively, Crandall *et al*. (2000) advocate a conservation management strategy that examines genetic and ecological data across two time scales, recent and historical. The examination of genetic or ecological exchangeability on an historical time scale can be problematic; however, the use of two different molecular markers may allow the partitioning of recent and historical phenomena. In this study, mitochondrial DNA may be more indicative of historical relationships, whereas AFLPs, due to the large number of loci examined, may be able to detect more recent population divergence (Egger *et al*., 2007). *Gymnocypris eckloni* and *G. przewalskii* were found to be significantly different at both mtDNA and AFLP markers. The two species are also morphologically distinct and occupy different habitats (*i.e. G. eckloni* permanently reside in the freshwater Yellow River drainage), an indicator of recent ecological non-exchangeability. Therefore, using the principles of genetic and ecological exchangeability (Crandall *et al*., 2000), *G. eckloni*

and *G. przewalskii* should be considered different ESUs, a result expected given the current taxonomy.

GANZI RIVER SUB-SPECIES

Both the control region and AFLP markers indicated that the Ganzi River fish are significantly different from all *G. p. przewalskii* populations, consistent with the current sub-species status. AMOVAs of both the control region and AFLP data (Tables III and IV) found the Ganzi fish to be significantly different from all sampling locations of *G. p. przewalskii*. As the average genetic distance (calculated with AFLP data) of individuals that shared mtDNA haplotypes was not significantly different from those that did not share haplotypes, it is unlikely that the results in this study were due to the sampling of sibling groups. Consistent with the results of the species level comparison and previous findings (Zhao *et al.,* 2005), the Ganzi samples were not found to be monophyletic when the control region was analysed. Fewer haplotypes, however, were shared with *G. p. przewalskii* than the degree of haplotype sharing found among *G. p. przewalskii* sampling locations (Fig. 2 and Table II). Additionally, Structure analysis revealed a unique AFLP cluster in *G. p. przewalskii* and the presence of genotypes within the Ganzi that were more similar to *G. eckloni* than to *G. p. przewalskii* (Fig. 3). These finding are intriguing as recent studies have shown that some AFLP markers may be linked to adaptive traits (*e.g.* Bonin *et al*., 2007). The strong selective pressure of Lake Qinghai (high salinity and pH) may indeed be driving adaptive divergence. As argued above, the AFLP data, however, were also consistent with the expectations of ongoing drift–lineage sorting at nuclear genes. Note, two *G. p. przewalskii*-like AFLP genotypes were also found within *G. eckloni.* Whether the observed differences were driven by selection or drift (or a combination of both), both marker sets clearly show genetic differences between the *G. p. ganzihonensis* and *G. p. przewalskii* sampling locations.

Two possible hypotheses as to the origin of *G. p. ganzihonensis* are presented. First, the sub-species may have evolved after a population of *G. p. przewalskii* was isolated in the Ganzi River when its lake connection was severed. Second, it is possible that the sub-species was always a distinct riverine population that evolved separately from the lake fish. A 1964 survey found that the Ganzi River lacked a direct surface water connection with Lake Qinghai (Zhu & Wu, 1975). Currently, the lower reaches of the river experience low flows and drain into a marshy area. During heavier flows, the marsh could periodically overflow into the adjacent Haergai River, but this would not allow a direct connection with the lake. It has been proposed that as the lake water levels fell, fish became trapped in the Ganzi River and were forced to adapt to the new environment (Zhu & Wu, 1975). The morphology of the Ganzi River is quite unlike the other rivers in that it originates from a hot spring and contains marshes in the middle and lower reaches. *Gymnocypris przewalskii ganzihonensis* has fewer gill rakers than *G. p. przewalskii*, a condition thought to have been an adaptation to a different food source in the Ganzi River (Zhu & Wu, 1975).

Although the lake has been regressing for several thousand years, it is not known at what time the direct connection with the Ganzi River was severed (Walker *et al*., 1996). Some reports suggest that the lake has receded *c*. 100 km since the late Pleistocene and historical data indicated that the lake has fallen *c*. 12 m from the late 1800s to the 1990s, with the regression being more rapid in the last 30 years

(*cf.* Walker *et al*., 1996). Intriguingly, historical records from the Tang Dynasty (608–904 A.D.) suggest that higher lake levels at the time may have flooded the current lower marsh area possibly allowing a direct connection with the Ganzi River (Zhu & Wu, 1975). If the connection was severed, as the lake receded within the last thousand years, then the morphological changes noted for this sub-species must have occurred quickly, as speculated by Zhu & Wu (1975).

Alternatively, it is possible that *G. p. ganzihonensis* did not evolve after being isolated from the lake in recent times, but were always a distinct riverine population that has not undergone genetic exchange with the lake fish. In addition to containing gravel areas necessary for spawning, the Ganzi River contains suitable overwintering habitats in the form of deep pools in the marshy areas (Zhu & Wu, 1975); therefore, it is possible that the fish were able to overwinter in the river without ever having to return to the lake. Polymorphisms that existed in the founding population may have been lost due to selection or bottlenecks, resulting in fewer shared haplotypes with the lake fish, *G. p. przewalskii* (Fig. 2). The relative high frequency of unique mtDNA haplotypes in the Ganzi, the clustering of some AFLP genotypes with *G. eckloni* genotypes and the lack of lake-specific genotypes (Fig. 3) support an older over a more recent ancestry. The presence of lake mtDNA haplotypes within the Ganzi, however, suggests that the split from the lake occurred some time after the separation of the Qinghai drainage from the Yellow River drainage.

Applying the criteria of Crandall *et al*. (2000), the two sub-species are likely to be genetically non-exchangeable, both historically and recently, as significant genetic divergence was detected using both mitochondrial and AFLP markers. The ecological exchangeability between the two sub-species is, however, difficult to assess. Even though the Ganzi River fish occupy a very distinct habitat relative to fish from the other rivers, it is currently unknown if they are able to tolerate the salinity of the lake. Therefore, there is a need for studies similar to that of Wood *et al*. (2007) to determine the physiological effects of moving *G. p. ganzihonensis* from river to lake water. If the Ganzi River fish are unable to acclimate to the salinity of the lake (an indicator of recent ecological non-exchangeability), then, according to Crandall's criteria, *G. p. przewalskii* and *G. p. ganzihonensis* should be considered as distinct ESUs. If, however, the Ganzi fish are able to reside in the lake and no other evidence exists which suggests a lack of ecological exchangeability, then *G. p. przewalskii* and *G. p. ganzihonensis* should be treated as a single ESU. In this case, however, for the purposes of management and conservation, it is recommended that they be treated as separate management units (MU) (Moritz, 1994).

LAKE QINGHAI STOCK STRUCTURE

AMOVA and Structure results of the AFLP data indicated the presence of a stock structure in Lake Qinghai (Tables III and IV, Fig. 3). This supports and expands previous work that found significant difference between the Buha and Shaliu Rivers (Heima, Quanji and Haergai not examined) with microsatellite markers (Zhang *et al*., 2009) and among the Buha, Heima and Shaliu Rivers (Quanji and Haergai not examined) with AFLP markers (Chen *et al*., 2005). Population structuring among four of the five lake rivers was also suggested in a separate study using isozymes (although no statistical analysis was performed) (Meng *et al*., 2007). Population structuring was not detected using mitochondrial analysis, a result that is similar to

other studies that combine mtDNA and AFLP data (Giannasi *et al*., 2001; Sullivan *et al*., 2004; Egger *et al*., 2007). It is thought that while mtDNA can detect larger-scale, historical relationships (such as the divergence between *G. p. przewalskii* and *G. p. ganzihonensis* where it is unlikely that gene flow exists), AFLPs are better at detecting finer-scale relationships, such as the population structure within the lake (where gene flow between populations may still exist due to straying) (Egger *et al*., 2007). The formation of distinct populations in anadromous species is primarily due to natal homing (Stewart *et al*., 2003). Although AFLP analysis indicated that most *G. p. przewalskii* populations of Lake Qinghai are genetically distinct, the degree of straying or dispersal present among the populations and the effects of straying on genetic divergence are unknown. It is also unknown if the populations are diverging towards an equilibrium state where the genetically isolating effects of natal homing are balanced by the homogenizing effects of a low level of migration (straying), or if they are currently at an equilibrium state. Fish species with a strong propensity for natal homing still rely on the mechanisms of dispersal to colonize new habitat or to avoid potentially adverse local conditions (Pascual *et al*., 1995; Milner *et al*., 2000; Stewart *et al*., 2003). Conditions in the Lake Qinghai watershed can be quite severe and large die-offs in spawning runs have been noted due to the rapid changes in water levels that have trapped fish in drying stream beds. A possible propensity for straying was also noted by the observation of fish swimming up an ephemeral stream that appeared after a heavy rainfall in June 2004, but was dry 2 days later (it is unknown if these individuals were sexually mature at the time). Sexually immature fish have also been observed in the Heima River (Wood *et al*., 2007), although the reason for their presence is unclear.

Despite the detection of a stock structure within the lake, samples collected from the Quanji and Shaliu Rivers were not found to be significantly different from each other. There may be a significant amount of gene flow between fish in these two tributaries (*i.e*. more straying than natal homing). Unstable habitats (*i.e*. decreased flows due to the installation of weirs and dams) may promote straying (Quinn & Dittman, 1990). Both rivers have had recent (since the 1960s) weir and water diversion projects that have impeded access to upstream spawning habitat and have affected water flow at downstream spawning sites. It is also possible that fish from one river may have been extirpated but repopulated by fish from the other river. Again, the isozyme study conducted by Meng *et al*. (2007) suggests that these two populations (from the Quanji and Shaliu Rivers) clustered together.

It is particularly interesting that the fish sampled from two different locations on the Buha River (*c*. 50 km from each other) were found to be significantly different, possibly indicating the presence of distinct spawning populations within the river. The evolutionary benefits of natal homing, namely increasing the chances of locating mates or habitats suitable for early life-history survival, may result in selective forces that develop within river subpopulations (Garant *et al*., 2000). The Buha River is the largest tributary of Lake Qinghai (Fig. 1), contributing approximately half of the total inflow (Walker *et al*., 1996). It has not been affected by weirs as in the other tributaries, possibly allowing for the continued existence of distinct within-river stocks. As there have been reports of juveniles overwintering in the upper reaches of the Buha River (Walker & Yang, 1999), it is also possible that a resident population exists in the upper reaches of the Buha. Because very little is known about the life-history of *G. p. przewalskii* (*e.g.* how long they reside in the river as

juveniles before travelling to the lake), however, this theory is speculative and would need to be confirmed with further studies. The detection of subpopulations within the Buha River also has implications for the management of *G. p. przewalskii*. It may be possible that similar within-river structuring exists in other spawning tributaries, but went undetected as only one location was sampled in each of these rivers.

The lack of mitochondrial divergence between the different sampling locations indicates genetic exchangeability on a historical level, whereas significant divergence at AFLP markers may indicate recent genetic non-exchangeability or natal homing with low levels of straying among populations. Ecological exchangeability refers to the ability of individuals to occupy the same niche, due to shared life-history traits and morphology. It seems likely that the fish of Lake Qinghai would be ecologically exchangeable, although further studies are required for this to be accurately determined. Applying the criteria of Crandall *et al*. (2000), *G. p. przewalskii* should be treated as a single ESU for conservation purposes. Exchangeability of the populations, however, should not be used to promote homogenization of fish spawning in the different rivers (Crandall *et al*., 2000). As some degree of genetic divergence was detected in this study, it is recommended that the spawning groups be treated as separate MUs.

It is clear from this research that much more information is needed if *G. przewalskii* stocks are to be managed properly. Although the genetic data in this study indicate a significant level of homing, the degree of straying is unknown as are the exact spawning locations. Unfortunately, the lack of clustering of genetic groups with species or populations precludes the use of Structure-type approaches with the AFLP marker systems to detect migrants. Future studies should therefore include tagging so that the fish may be accurately tracked to their spawning locations. Genetic analysis should focus on spawning locations. Straying rates can then be determined by combining temporal tagging data with ongoing genetic analysis so that this unique and possibly threatened population can be protected and managed.

This research was supported by a National Natural Science Foundation of China Excellent Young Investigator award (30128016) to Y.X.W. and J.Z.D., by the Zhejiang University. Overseas Researcher Program to B.W.M. and Y.X.W., a Tang Yongqian Foundation Fellowship at Zhejiang University to C.M.W. and by Natural Sciences and Engineering Research Council Canada Discovery Grants awarded to B.W.M., C.J.B., J.G.R. and C.M.W. The Fishery and Aquaculture Bureau of Qinghai Province is thanked for providing the technical and logistical assistance required for this research expedition. A thank you is also extended to Y. Liu of the Northwest High Plateau Institute of Biology for providing excellent logistical assistance; graduate students S. Wang and Y. Cao, Zhejiang University, J. Rogers, McMaster University, B. Madison, Queen's University and J. Semple, University of Waterloo, for their field assistance; M. Thompson of the University of Northern British Columbia Genetics Facility for laboratory assistance and M. Shrimpton and anonymous reviewers for helpful criticism.

References

April, J. & Turgeon, J. (2006). Phylogeography of the banded killifish (*Fundulus diaphanus*): glacial races and secondary contact. *Journal of Fish Biology* **69,** 212–228. doi: 10.1111/j.1095-8649.2006.01233.x

Bensch, S. & Åkesson, M. (2005). Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* **14,** 2899–2914. doi: 10.1111/j.1356-294X.2005.02655.x

Bian, Q. T., Liu, J. Q., Liu, X. Q. & Xiao, J. L. (2000). Geotectonic setting, formation, and evolution of the Qinghai Lake. *Seismology and Geology* **22,** 20–26.

Bonin, A., Nicole F., Pompanon, F., Miaud, C. & Taberlet, P. (2007). Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology* **21,** 697–708. doi: 10.1111/j.1523-1739.2007.00685.x

Brunner, P. C., Douglas, M. R. & Bernatchez, L. (1998). Microsatellite and mitochondrial DNA assessment of population structure and stocking effects in Arctic charr *Salvelinus alpinus* (Teleostei: Salmonidae) from central Alpine lakes. *Molecular Ecology* **7,** 209–223. doi: 10.1046/j.1365-294x.1998.00341.x

Chakraborty, R. & Weiss, K. M. (1991). Genetic variation of the mitochondrial DNA genome in American Indians is at mutation-drift equilibrium. *American Journal of Physical Anthropology* **86,** 497–506. doi: 10.1002/ajpa.1330860405

Chen, D., Zhang, C., Lu, C., Chang, Y. & Chang, J. (2005). Amplified fragment length polymorphism analysis to identify the genetic structure of the *Gymnocypris przewalskii* (Kessler, 1876) populations from the Qinghai Basin, China. *Journal of Applied Ichthyology* **21,** 178–183. doi: 10.1111/j.1439-0426.2005.00631.x

Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9,** 1657–1660. doi: 10.1046/j.1356-294X.2000.01020.x

Crandall, K. A., Bininda-Edmonds, O. R. P., Mace, G. M. & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution* **15,** 290–295. doi: 10.1016/S0169-5347(00)01876-0

Egger, B., Koblmüller, S., Sturmbauer, C. & Sefc, K. M. (2007). Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. *BMC Evolutionary Biology* **7,** 137–151. doi: 10.1186/1471-2148-7-137

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14,** 2611–2620. doi: 10.1111/j.1356-294X.2005.02553.x

Ewen, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoretical Population Biology* **3,** 87–112.

Excoffier, L. G., Laval, G. & Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1,** 47–50.

Falush, D., Stephens, M. & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7,** 574–578. doi: 10.1111/j.1471-8286.2007.01758.x

Garant, D., Dodson, J. J. & Bernatchez, L. (2000). Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar* L.). *Molecular Ecology* **9,** 615–628. doi: 10.1046/j.1365-294X.2000.00909.x

Gharrett, A. J., Matala, A. P., Peterson, E. L., Gray, A. K. & Li, Z. (2005). Two genetically distinct forms of rougheye rockfish are different species. *Transactions of the American Fisheries Society* **134,** 242–260. doi: 10.1577/T04-055.1

Giannasi, N., Thorpe, R. S. & Malhotra, A. (2001). The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, *Trimeresurus albolabris*. *Molecular Ecology* **10,** 419–426. doi: 10.1046/j.1365-294X.2001.01220.x

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52,** 696–704. doi: 10.1080/10635150390235520

Hawkins, J. S., Pleasants, J. & Wendel, J. F. (2005). Identification of AFLP markers that discriminate between cultivated cotton and the Hawaiian islands endemic, *Gossypium tomentosum* Nuttall ex Seeman. *Genetic Resources and Crop Evolution* **52,** 1069–1078. doi: 10.1007/s10722-004-6115-z

Ishikawa, S., Tsukamoto, K. & Nishida, M. (2004). Genetic evidence for multiple geographic populations of the giant mottled eel *Anguilla marmorata* in the Pacific and Indian oceans. *Ichthyological Research* **51,** 343–353. doi: 10.1007/s10228-004-0241-7

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5,** 150–163. doi: 10.1093/bib/5.2.150

Li, J. J. & Fang, X. M. (1998). Uplift of the Tibetan Plateau and environmental changes. *Chinese Science Bulletin* **44,** 2117–2125.

Meng, P., Wang, W.-J., Kong, J., Li, X.-J. & Shi, J.-Q. (2007). Isozyme variation of Lake Qinghai naked carp *Gymnocypris przewalskii* (Kessler) from five rivers. *Acta Zoologica Sinica* **53,** 892–898.

Miller, M. P. (1997). Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Milner, A. M., Knudsen, E. E., Soiseth, C., Robertson, A. L., Schell, D., Phillips, I. T. & Magnusson, K. (2000). Colonization and development of stream communities across a 200-year gradient in Glacier Bay National Park, Alaska, U.S.A. *Canadian Journal of Fisheries and Aquatic Science* **57,** 2319–2335. doi: 10.1139/cjfas-57-11-2319

Mock, K. E. & Miller, M. P. (2005). Patterns of molecular diversity in naturally occurring and refugial populations of the least chub. *Transactions of the American Fisheries Society* **134,** 267–278. doi: 10.1577/T04-037.1

Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution* **9,** 373–375. doi: 10.1016/0169-5347(94)90057-4

Nei, M. (1972). Genetic distance between populations. *American Naturalist* **106,** 283–292.

Pascual, M. A., Quinn, T. P. & Fuss, H. (1995). Factors affecting the natal homing of fall chinook salmon from the Columbia River hatcheries. *Transactions of the American Fisheries Society* **124,** 308–320. doi: 10.1577/1548-8659(1995)124<0308:FATHOF>2.3.CO;2

Peakall, R. & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6,** 288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Posada, D. (2008). jModeltest: phylogenetic model averaging. *Molecular Biology and Evolution* **25,** 1253–1256. doi: 10.1093/molbev/msn083

Qi, D., Li, T., Zhao, X., Guo, S. & Li, J. (2006). Mitochondrial cytochrome b sequence variation and phylogenetics of the highly specialized Schizothoracine fishes (Teleosteii: Cyprinidae) in the Qinghai–Tibet Plateau. *Biochemical Genetics* **44,** 270–285.

Qi, D., Guo, S., Tang, J., Zhao, X. & Liu, J. (2007). Mitochondrial DNA phylogeny of two morphologically enigmatic fishes in the subfamily Schizothoracinae (Teleostei: Cyprinidae) in the Qinghai-Tibetan plateau. *Journal of Fish Biology* **70,** 60–74. doi: 10.1111/j.1095-8649.2007.01366.x

Quinn, T. P. & Dittman, A. H. (1990). Pacific salmon migrations and homing: mechanism and adaptive significance. *Trends in Ecology & Evolution* **5,** 174–177. doi: 10.1016/0169-5347(90)90205-R

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Seehausen, O., Koetsier, E., Schneider, M. V., Chapman, L. J., Chapman, C. A., Knight, M. E., Turner, G. F., van Alphen, J. J. M. & Bills, R. (2003). Nuclear markers reveal unexpected genetic variation and a Congolese-Nilotic origin of the Lake Victoria cichlid species flock. *Proceedings of the Royal Society of London B* **270,** 129–137. doi: 10.1098/rspb.2002.2153

Seutin, G., White, B. N. & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69,** 82–90. doi: 10.1139/z91-013

Stepien, C. A. & Faber, J. E. (1998). Population genetic structure, phylogeography and spawning philopatry in walleye (*Stizostedion vitreum*) from mitochondrial DNA control region sequences. *Molecular Ecology* **7,** 1757–1769. doi: 10.1046/j.1365-294X.1998.00512.x

Stewart, I. J., Quinn, T. P. & Bentzen, P. (2003). Evidence for fine-scale natal homing among island beach spawning sockeye salmon, *Oncorhynchus nerka*. *Environmental Biology of Fishes* **67,** 77–85. doi: 10.1023/A:1024436632183

Sullivan, J. P., Lavoué, S., Arnegard, M. E. & Hopkins, C. D. (2004). AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric

fish (Mormyroidea: Teleostei). *Evolution* **58,** 825–841. doi: 10.1111/j.0014.3820.2004.tb00415.x

Tabata, K. & Taniguchi, N. (2000). Differences between *Pagrus major* and *Pagrus auratus* through mainly mtDNA control region analysis. *Fisheries Science* **66,** 9–18. doi: 10.1046/j.1444-2096.2000.00032x

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105,** 437–460.

Taylor, E. B., Rendenbach, Z., Costello, A. B., Pollard, S. & Pacas, C. J. (2001). Nested analysis of genetic diversity in northwestern North American char, Dolly Varden (*Salvelinus malma*) and bull trout (*Salvelinus confluentus*). *Canadian Journal of Fisheries and Aquatic Science* **58,** 406–420. doi: 10.1139/cjfas-58-2-406

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23,** 4407–4414.

Walker, K. F. & Yang, H. Z. (1999). Fish and fisheries of western China. In *Fish and Fisheries at Higher Altitudes: Asia*, p. 304. *FAO Fisheries Technical Paper. No.* **385**. Rome: FAO.

Walker, K. F., Dunn, I. G., Edwards, D., Petr, T. & Yang, H. Z. (1996). A fishery in a changing lake environment: the naked carp *Gymnocephalus przewalskii* (Kessler) (Cyprinidae: Schizothoracinae) in Qinghai Hu, China. *International Journal of Salt Lake Research* **4,** 169.

Wang, Y. S., Gonzalez, R. J., Patrick, M. L., Grosell, M., Zhang, C., Feng, Q., Du, J., Walsh, P. J. & Wood, C. M. (2003). Unusual physiology of scale-less carp, *Gymnocypris przewalskii*, in Lake Qinghai: a high altitude alkaline saline lake. *Comparative Biochemistry and Physiology A* **134,** 409–421. doi: 10.1016/S1095-6433(02)00317-3

Watterson, G. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* **7,** 256–276.

Wood, C. M., Du, J., Rogers, J., Brauner, C. J., Richards, J. G., Semples, J. W., Murray, B. W., Chen, X. Q. & Wang, Y. (2007). Przewalski's Naked Carp (*Gymnocypris przewalskii*): an endangered species taking a metabolic holiday in Lake Qinghai, China. *Physiological and Biochemical Zoology* **80,** 59–77. doi: 10.1086/509212

Xie, Z., Du, J., Chen, X., Wang, Y. & Murray, B. (2006). The significance of mitochondria control region (D-Loop) in intraspecific genetic differentiation of fish. *Hereditas (Beijing)* **28**, 362–368.

Yang, J., Qi, H., Shi, J., Chen, D., Huang, F. & Tang, H. (2005). Hydrochemistry properties and water quality analysis of Qinghai hu lake. *Freshwater Fisheries* **35,** 28–32.

Young, W. P., Ostberg, C. O., Keim, P. & Thorgaard, G. H. (2001). Genetic characterization of hybridization and introgression between anadromous rainbow trout (*Onchorhyncus mykiss irideus*) and coastal cuthroat trout (*O. clarki clarki*). *Molecular Ecology* **10,** 921–930. doi: 10.1046/j.1365-294X.2001.01247.x

Yue, P. & Chen, Y. (1998). *China Red Data Book of Endangered Animals – Pisces.* Beijing: Science Press.

Zhan, T. (1994). Annex 2: priority species in China. In *China Biodiversity Conservation Action Plan* (Maxey, C. & Lutz, J., eds). Beijing: China Environmental Science Press.

Zhao, K., Li, J., Yang, G., Duan, Z., He, S. & Chen, Y. (2005). Molecular phylogenetics of *Gymnocypris* (Teleostei: Cyprinidae) in Lake Qinghai and adjacent drainages. *Chinese Science Bulletin* **50,** 1325–1333. doi: 10.1360/982005-223

Zhang, Y., Jiang, P., Gang, H., Shi, J.-Q. & Sun, X. (2009) Development and characterization of microsatellite markers to study population genetic diversity of Przewalski's naked carp *Gymnocypris przewalskii* in China. *Journal of Fish Biology* **75,** 2352–2356. doi: 10.1111/j.1095-8649.2009.02437.x

Zheng, M. (1997). An introduction to saline lakes on the Qinghai–Tibet plateau. In *Monographiae Biologicae* (Dumont, H. J. & Werger, M. J. A., eds), pp. 1–289. Beijing: Scientific Publisher.

Zhu, S. & Wu, F. (1975). Study of fish fauna in Qinghai Lake. In *Fish Fauna and Biology of Naked Carp in Qinghai Lake*, pp. 9–26. Beijing: Scientific Publisher.

Zouros, E. (1979). Mutation rates, population sizes and amounts of electrophoretic variation of enzyme loci in natural populations. *Genetics* **92,** 623–646.