Journal of Fish Biology (2008) **72,** 2011–2025
doi:10.1111/j.1095-8649.2008.01836.x, available online at http://www.blackwell-synergy.com

Population structure and phylogeography of *Aphyocypris kikuchii* (Oshima) based on mitochondrial DNA variation

H.-D. Lin*, K.-C. Hsu*†, K.-T. Shao†, Y.-C. Chang‡, J.-P. Wang*, C.-J. Lin* and T.-Y. Chiang*§

**Department of Life Sciences, Cheng-Kung University, Tainan, Taiwan 701, †Research Center for Biodiversity, Academia Sinica, Taipei, Taiwan 115 and ‡General Education Center, Tzu-Chi University, Hualien, Taiwan 970*

(*Received 2 October 2007, Accepted 30 January 2008*)

Aphyocypris kikuchii is a cyprinid species endemic to northern and eastern Taiwan and is the only primary freshwater fish native east of the Coastal Mountain Range. In total, 92 individuals of *A. kikuchii* from seven populations in three regions of the island were surveyed for mitochondrial DNA (mtDNA) variation. High haplotype diversity ($h = 0{\cdot}989$) and low nucleotide diversity ($\pi = 0{\cdot}009$) of mtDNA were detected. Negative values of Tajima's *D* and unimodal mismatch distributions probably reflect a history of recent demographic expansions from small populations. Three major haplotype clusters displayed geographically non-overlapping distributions, indicating a long-term isolation between regions. Hierarchical analysis of molecular variance showed significant genetic structuring among populations ($\Phi_{ST} = 0{\cdot}66$). Significant haplotype heterogeneity was also detected among populations within regions ($\Phi_{SC} = 0{\cdot}41$, $P < 0{\cdot}001$) and among regions ($\Phi_{CT} = 0{\cdot}43$, $P < 0{\cdot}05$). Molecular clock estimates of coalescence in the three major mtDNA lineages indicated coalescence in the most recent common ancestor *c.* 0·11–0·39 million years ago. Haplotypes of cluster B nested as interior nodes in the haplotype network, indicating that migrations from Shueilian (SL) populations to the northern region (cluster A) and to the eastern region (cluster C) may have occurred independently. Lineages A and B + C should be managed as two distinct evolutionarily significant units, while the northern, SL and southern groups should be managed as separate management units.

Key words: demographic expansion; freshwater fish; geographical structuring; most recent common ancestor; Taiwan; vicariance.

INTRODUCTION

Given the island-like character of freshwater habitats, the phylogeographical patterns of freshwater fish populations often reflect the hydrographic and geological history of a region because of the development of hydrographic basins and their isolation and interconnection processes (Bermingham & Martin,

§Author to whom correspondence should be addressed. Tel.: +886 6 2757575; fax: +886 62742583; email: tychiang@mail.ncku.edu.tw

1998). Mountain ranges often represent impassible barriers to aquatic organisms. As a result, areas bounded by mountain ridges become isolated zoogeographic units for the aquatic fauna (Slechtova *et al.*, 2004). The historical patterns of isolation among drainages usually result in strong zoogeographical differences among populations (Meffe & Vrijenhoek, 1988). Therefore, phylogeographical analyses of freshwater fishes can be used to infer the relationship between the biotic and the geological evolution of a region (Lundberg, 1993).

The subtropical island of Taiwan is 395 km long and 144 km wide and provides an excellent opportunity for comparing contemporary phylogeographical patterns with biogeographical hypotheses. Geological evidence indicates that Taiwan was uplifted by the collision of the Luzon volcanic arc of the Philippine Sea Plate with the Eurasian Plate about 4 million years ago (Lin, 1966). The steep Central Mountain Range runs along the northeast-southwest longitudinal axis of the island with the highest peak nearly 4000 m above sea level. These mountains, therefore, present a major barrier to dispersal across the island. Many streams have etched deeply into the mountainous terrains, further isolating populations on each side of the mountain range. In addition, the volcanic islands north of Lutao and Lanhsu collided with accreted Asian continental margins to form the Coastal Range in eastern Taiwan (Yu & Song, 1994).

Aphyocypris kikuchii (Oshima, 1919) is a small cyprinid fish endemic to northeastern Taiwan (Fig. 1). During the past decades, its populations have declined drastically because of habitat degradation, and this species has been recently assigned endangered status (Liao *et al.*, 2005). The species has a low dispersal capability (Tzeng, 1986), thereby reducing gene flow between populations and likely leading to genetic differentiation. Because of the low commercial value of the fish, the distribution of *A. kikuchii* reflects the natural population history, undisturbed by artificial stocking or transplantation.

Unlike other widespread species, the distribution of this species is restricted to the northern and eastern zoogeographical districts of the island (*cf.* Tzeng, 1986). It is the only primary freshwater fish distributed eastward of the Coastal Mountain Range and, therefore, may play a key role in the understanding of the geographical history and the formation effect of the Coastal Mountain Range.

The complete control region (CR), partial cytochrome *b* (cyt *b*) and 12s rRNA mitochondrial DNA (mtDNA) genes were sequenced in this study to estimate the phylogeographical patterns among populations. The aims of this study were (1) to examine the population genetic structure with mtDNA, (2) to identify the colonization routes of *A. kikuchii* in Taiwan, (3) to assess whether the Central Mountain and Coastal Mountain ranges act as geographical barriers to dispersal, (4) to estimate coalescence times of the major mtDNA lineages to the most recent common ancestor (MRCA) and (5) to identify conservation units.

MATERIALS AND METHODS

FISH SAMPLES

In total, 92 individuals of *A. kikuchii* were collected from seven populations in three regions in Taiwan. A majority of collecting sites in the northern region were located north of the Central Mountain Range, including the Lanyang River (LY) and the

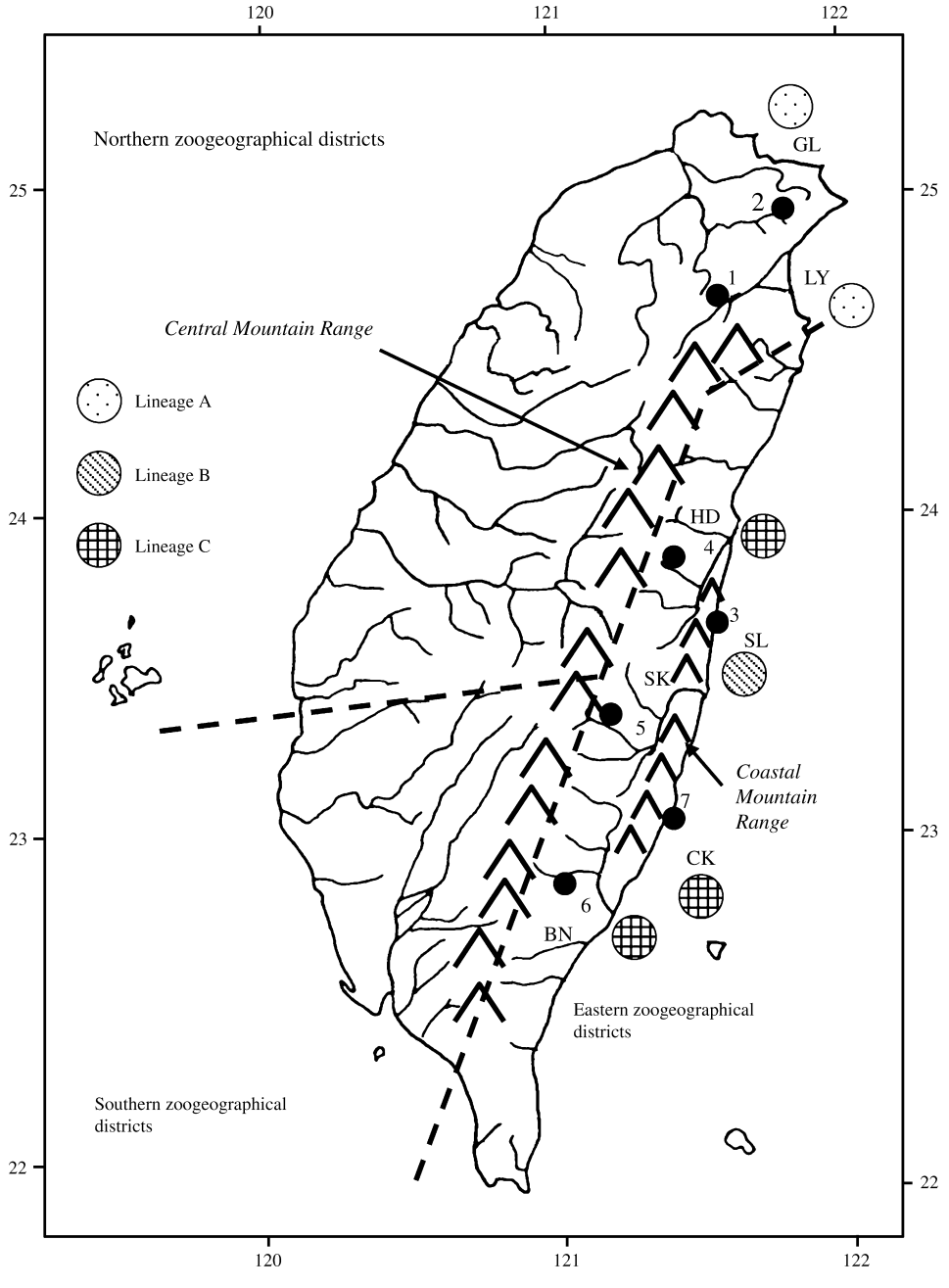

FIG. 1. Map of Taiwan showing the seven sampling localities. 1. Lanyang River (LY), 2. Keelung River (GL), 3. Shueilian River (SL), 4. Hualien River (HD), 5. Hsiukuluan River (SK), 6. Beinan River (BN) and 7. Chungkung River (CK).

Keelung River (GL). The longitudinal valley collecting sites were located between the Central Mountain Range and the Coastal Mountain Range and included the Hualien (HD), Hsiukuluan (SK) and Beinan (BN) rivers. The east coast mountain sites were

located east of the Coastal Mountain and included the Shueilian (SL) and Chungkung (CK) Rivers (Table I and Fig. 1).

DNA SEQUENCING

Specimens were preserved in 95% ethanol. Genomic DNA was extracted from muscle tissue following a standard phenol–chloroform protocol (Blin & Stafford, 1976). A *c.* 1879 bp DNA fragment of mtDNA, including the 3′ end of the cyt *b*, threonine tRNA, proline tRNA genes, the CR region, the phenylalanine tRNA gene and the 5′ end of 12S rRNA, was amplified by the polymerase chain reaction (PCR) with the primers Loop1 (F) (5′-AGG AGG TGT TCT TGC ACT AC-3′) and Loop2 (R) (5′-AGC TAC TTT CGT GTA TTG GG-3′) (Wang *et al.*, 2000, 2004). Each 100 μl PCR solution contained 10 ng template DNA, 10 μl 10× reaction buffer, 10 μl $MgCl_2$ (25 mM), 10 μl dNTP mix (8 mM), 10 pmole of each primer and 4 U of *Taq* polymerase (Promega, Madison, WI, U.S.A.). The PCR reaction was programmed on an MJ Thermal Cycler (MJ Research, Waltham, MA, U.S.A.) as one cycle of denaturation at 95° C for 4 min, then 30 cycles of denaturation at 94° C for 45 s, annealing at 48° C for 1 min 15 s and extension at 72° C for 1 min 30 s, followed by 72° C extension for 10 min and 4° C for storage. PCR products were purified by electrophoresis on a 1·0% agarose gel using 1× tris-acetate-EDTA (TAE) buffer. The gel was stained with ethidium bromide, and the desired DNA band was excised and eluted using agarose gel purification kit (Qiagen, Valencia, CA, U.S.A.). Purified DNAs were ligated to a pGEM-T easy vector (Promega). Plasmid DNAs were selected randomly with five clones and purified using a plasmid mini kit (Qiagen). Purified plasmid DNAs were sequenced in both directions with a *Taq* dye deoxy terminator cycle sequencing kit (Perkin Elmer, Wellesley, MA, U.S.A.). Primers for sequence determination were T7-promoter and SP6-promoter, located on p-GEM-T easy Vector termination site. Internal primers, MT-2 (R) (5′-ACATTTGAGCCTG-CACTCTG-3′) and MT-1 (F) (5′-CCCACCAAGCCGAGCGTTGT-3′), were also used for sequencing (Wang *et al.*, 2000).

DATA ANALYSIS

Sequences were aligned with CLUSTALX 1.81 (Thompson *et al.*, 1997). MODEL-TEST 3.06 (Posada & Crandall, 1998) was used to determine the best fitting mutation model of nucleotide substitution for each dataset. Searches for all gene portions were conducted by the K80 model (Hasegawa *et al.*, 1985). A haplotype neighbour-joining tree was generated with MEGA 2 (Kumar *et al.*, 2001). For all methods, confidence values for nodes in the trees were calculated with 1000 bootstrap replicates (Felsenstein, 1985). Nodes with bootstrap values greater then 0·70 were considered as significant (Hillis & Bull, 1993). The number of mutations between DNA haplotypes in pair-wise comparisons was calculated with MEGA 2 and used to construct a minimum spanning network with MINSPNET (Excoffier & Smouse, 1994). The rules of Templeton *et al.* (1987) were used to define an evolutionary clade hierarchy of haplotypes.

mtDNA polymorphism was estimated as nucleotide diversity (θ and π) (Jukes & Cantor, 1969) (Nei, 1987) and as haplotype diversity (h) (Nei & Tajima, 1983) using DNASP 3.95 (Rozas & Rozas, 1999). Analysis of molecular variance (AMOVA) was used to assess geographical patterns of population subdivision with ARLEQUIN 2000 (Schneider *et al.*, 2000) to determine the relative partitioning of variation within and between populations. Samples were grouped according to the mtDNA lineages recovered in the phylogenetic analyses or according to geography. A Mantel correlation of genetic and geographical distances was carried out with 10 000 permutations.

Sequence mismatch distributions and tests of neutrality were used to make inferences about historical demographies of *A. kikuchii* populations. Tajima's D (Tajima, 1989) was calculated for untranslated DNA fragments as an indicator of fit to mutation-drift equilibrium (Gelas & Meester, 2005). Mismatch analysis and Tajima's D were performed with DNASP 3.95 (Rozas & Rozas, 1999). Negative values of D ($\pi < \theta$) may indicate population expansions after a population bottleneck or the effects of selection.

TABLE I. Samples of *Aphyocypris kikuchii* used for mtDNA analyses and nucleotide and haplotype diversities

Site no.	Populations (Abbreviation)	Sample size	Haplotype numbers	Haplotype diversity (h)	Nucleotide diversity (π)	Nucleotide diversity (θ)	Haplotypes
1	Lanyang River (LY)	15	12	0·971	0·0026	0·0051	H1–H12
2	Keelung River (GL)	10	9	1·000	0·0021	0·0035	H13–H21
3	Shueilian River (SL)	18	12	0·941	0·0026	0·0050	H22–H33
4	Hualien River (HD)	19	17	0·982	0·0050	0·0089	H34–H50
5	Hsiukuluan River (SK)	10	8	0·978	0·0019	0·0029	H47, H51–H57
6	Beinan River (BN)	10	9	0·978	0·0041	0·0063	H53, H58–H65
7	Chungkung River (CK)	10	9	0·978	0·0028	0·0033	H47, H53, H66–H72
8	Total	92	72	0·989	0·0094	0·0182	

This study used the Bayesian coalescent framework implemented in BEAST 1.3 (Drummond & Rambaut, 2003) to estimate the coalescent times of major lineages to the MRCA. Wang *et al.* (2004) provided a robust estimate of average mutation rates in the CR of *Varicorhinus barbatulus* (Pellegrin, 1908) in Taiwan. A calibration of 3·2–10·4% changes per million years in the CR region was applied to the coalescent analyses to calibrate trees and to obtain absolute values of MRCA. All analyses were performed using the HKY model of nucleotide substitution (Hasegawa *et al.*, 1985). Rate variation among sites was modelled using a gamma distribution with six rate categories. Burn-in was set to 10^6 generations. Bayesian analysis using Markov-Chain Monte Carlo (MCMC) integration was used to estimate MRCA for the A and B and B and C lineages and for the entire Taiwan population. Adequate sampling and convergence to the stationary distribution were checked using TRACER 1.3 (Rambaut & Drummond, 2004). Posterior estimates of parameters were all distinctly unimodal (although with wide 95% highest posterior densities), and all parameters could be estimated, despite the relatively low information content in the sequences and the small age range of the sequences.

RESULTS

GENETIC VARIABILITY AND GENE GENEALOGY OF MTDNA

Aligned sequences of 1879 bp of mtDNA, including 286 bp of a cyt *b*, 70 bp of threonine tRNA, 72 bp of proline tRNA, 1002 bp of the CR, 59 bp of phenylalanine tRNA and 390 bp of 12S rRNA partial sequence of *A. kikuchii* were used in this study. One hundred and sixty-four positions (8·7%) were polymorphic, of which 67 (3·5%) were parsimoniously informative.

A remarkably large number of haplotypes appeared in the samples: 72 haplotypes were identified in the 92 individuals examined (Table I). Sequences of all haplotypes were deposited in the GenBank (accession numbers AM712236–AM712238 and AM886465–AM886533). The overall level of haplotype diversity (h) was high, ranging between 0·941 (SL) and 1·00 (GL). Nucleotide diversity (θ) averaged 0·0182 among samples and ranged from 0·0029 (SK) to 0·0089 (HD) within samples (Table 1).

Most analyses recovered a single set of statistically well-supported groups corresponding to geographically cohesive sets of populations. Three major lineages were recovered, hereafter referred to as lineages A, B and C. Lineages B and C clustered together as a sister to the lineage A. Samples from north of the Central Mountain Range (GL and LY) were genetically distinguishable and clustered together into mtDNA lineage A. The sample in the Shueilian River (SL) represented another lineage (B). The HD, SK and BN samples from the longitudinal valley between the Central Mountain and eastern Coastal Mountain ranges and CK sample from the southern districts of the Coastal Mountains represented lineage C. These lineages showed non-overlapping geographical distributions. Lineage A was separated from lineages B and C by the Central Mountain Range, whereas the Coastal Mountain Range separated lineages B and C (Fig. 2).

POPULATION STRUCTURE AND AMOVA

For the hierarchical AMOVA, samples were grouped according to the mtDNA lineages recovered in the phylogeny or according to different geographical

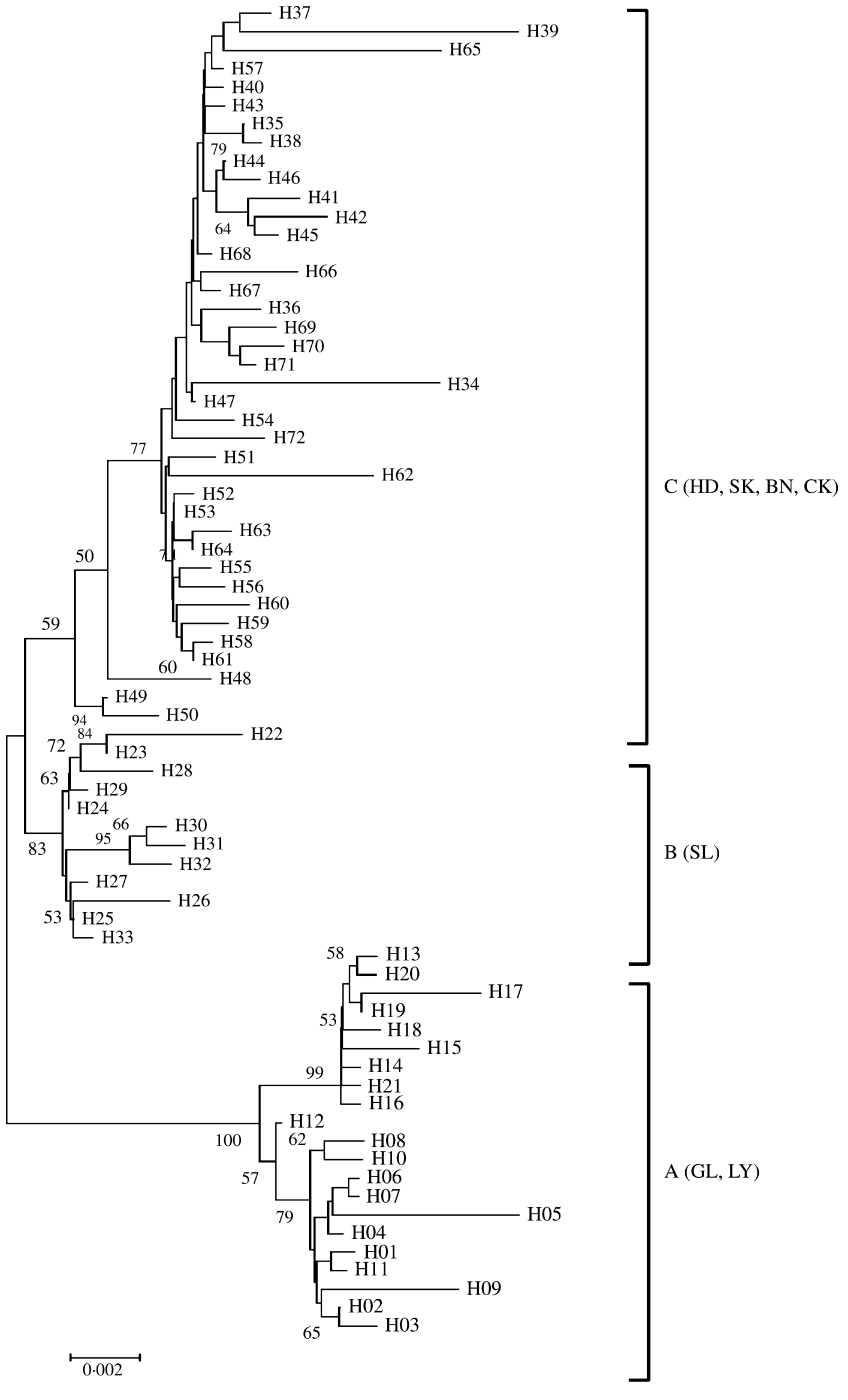

FIG. 2. Neighbour-joining tree based on haplotype sequence variation of the mtDNA in *Aphyocypris kikuchii*. Numbers at the nodes indicate bootstrap values (expressed as percentage) with 1000 replicates. BN, Beinan River; CK, Chungkung River; GL, Keelung River; HD, Hualien River; LY, Lanyang River; SK, Hsiukuluan River; SL, Shueilian River.

TABLE II. Analysis of molecular variance between populations of *Aphyocypris kikuchii* in Taiwan

Source of variation	Variance components	Percentage of variation
Among groups	6·18544**	43·83
Among populations within groups	3·23918**	22·95
Within populations	4·68687**	33·21
Overall F_{ST}	0·66**	

**$P < 0·05$.

hierarchies at three levels: (1) among lineages A, B and C; (2) among sites within lineages and (3) within sites. AMOVA identified significant genetic structure at each of these levels (Table II). Hierarchies based on the mtDNA phylogenetic groups maximized the among-group variation and were assumed to be the most probable geographical subdivision. Hierarchical AMOVA tests showed overall significant levels of genetic structuring among *Aphyocypris* collections ($\Phi_{ST} = 0·66$, $P < 0·001$). The relative contributions of differences among samples within regions to molecular variance ($\Phi_{SC} = 0·41$, $P < 0·001$) and among geographical regions ($\Phi_{CT} = 0·43$, $P < 0·05$) were also significant. The AMOVA indicated that 43·8% of the total variation existed among regions, 22·9% among samples within regions and 33·2% within samples. An overall $F_{ST} = 0·66$ also indicated significant structured among populations (Table III). Geographical subdivision was associated with high levels of genetic differentiation among populations within regions, except for lineage C, which had F_{ST} values ranging from 0·063 to 0·174. In contrast, higher levels of differentiation existed among populations from different regions, with F_{ST} values ranging from 0·587 (between SL and HD) to 0·856 (between GL and CK). No consistent patterns appeared between genetic distance (as pair-wise F_{ST}) and geographical distance (km), as indicated by a Mantel test of the pair-wise F_{ST} values and geographic distance with a correlation coefficient of 0·0024.

POPULATION EXPANSION AND MISMATCH ANALYSIS

Although negative, Tajima's D did not differ significantly from zero in the overall dataset ($D = -1·213$; $P > 0·10$). However, values were significant for lineages B and C but not for lineage A ($D = -1·746$, $0·10 > P > 0·05$) (Table IV). The three samples from the longitudinal valley (HD, SK and BN) displayed a significant deviation from neutrality for 39 sequences ($D = -2·354$, $P < 0·01$).

HAPLOTYPE NETWORK AND MOLECULAR DATING

A minimum spanning network showed seven 1-step and three 2-step clades. The entire network consisted of a three-step clade (Fig. 4). These three clades, 2-1, 2-2 and 2-3, corresponded to the three geographic groups. Clade 2-2 consisted of haplotypes H01–H21 in samples north of the Central Mountain Range, including samples GL (clade1-3) and LY (clade1-4). Clade 2-1 consisted of haplotypes H22–H33 and appeared in sample SL from east of the East

TABLE III. Pairwise Nm (above diagonal) and F_{ST} (below diagonal) values between samples of *Aphyocypris kikuchii* based on mtDNA

	GL	LY	SL	HD	SK	BN	CK
GL		0·31	0·10	0·12	0·31	0·11	0·08
LY	**0·619**		0·12	0·14	0·08	0·13	0·10
SL	**0·830**	**0·801**		0·35	0·21	0·34	0·25
HD	**0·805**	**0·779**	**0·587**		3·16	2·37	7·34
SK	**0·619**	**0·855**	**0·705**	**0·137**		13·77	3·20
BN	**0·816**	**0·796**	**0·598**	**0·174**	**0·035**		3·17
CK	**0·856**	**0·836**	**0·664**	**0·064**	**0·135**	**0·136**	

BN, Beinan River; CK, Chungkung River; GL, Keelung River; HD, Hualien River; LY, Lanyang River; SK, Hsiukuluan River; SL, Shueilian River. F_{ST} estimates in boldface indicate significance at $P < 0{\cdot}05$.

Coastal Mountain Range. Clade 2-3 (H34 to H72) occurred in the longitudinal valley group (HD, SK and BN) and in the sample CK from east of the Coastal Mountain Range. The haplotype network displayed a 'star-like' genealogy with several low-frequency haplotypes connected to the central haplotype H53 by one to a few mutations.

Molecular dating indicated that the three major lineages (A–C) coalesced to the MRCA $1{\cdot}1 \times 10^5$ to $3{\cdot}9 \times 10^5$ years before present, with a 95% posterior density (HPD) interval between $8{\cdot}3 \times 10^4$ to $2{\cdot}9 \times 10^5$ and $1{\cdot}4 \times 10^5$ to $4{\cdot}6 \times 10^5$ years, respectively. Likewise, lineages B and C coalesce between $7{\cdot}1 \times 10^4$ (with an interval of $5{\cdot}3 \times 10^4$ to $9{\cdot}4 \times 10^4$ years) and $2{\cdot}3 \times 10^5$ years ($1{\cdot}7 \times 10^5$ to $3{\cdot}0 \times 10^5$ years).

TABLE IV. Tajima's *D* in seven samples and five mtDNA lineages of *Aphyocypris kikuchii* in Taiwan

Area	Population	Number of populations	*n*	Singletons/ variable sites	Tajima's *D*	Probability
Total		7	92	97/164	*D*: −1·76576	$0{\cdot}10 > P > 0{\cdot}05$
Lineages A		2	25	38/50	*D*: −1·74597	$0{\cdot}10 > P > 0{\cdot}05$
	GL	1	10	18/18	*D*: −2·04996	$P < 0{\cdot}01$
	LY	1	15	27/30	*D*: −2·18385	$P < 0{\cdot}01$
Lineages B		1	18	24/31	*D*: −1·92616	$P < 0{\cdot}05$
	SL	1	18	24/31	*D*: −1·92616	$P < 0{\cdot}05$
Lineages C		4	49	67/96	*D*: −2·52732	$P < 0{\cdot}001$
	HD	1	19	38/56	*D*: −1·92252	$P < 0{\cdot}05$
	SK	1	10	14/15	*D*: −1·73719	$P < 0{\cdot}05$
	BN	1	10	28/32	*D*: −1·86365	$P < 0{\cdot}05$
	CK	1	10	10/17	*D*: −0·88576	$P > 0{\cdot}10$

BN, Beinan River; CK, Chungkung River; GL, Keelung River; HD, Hualien River; LY, Lanyang River; SK, Hsiukuluan River; SL, Shueilian River.

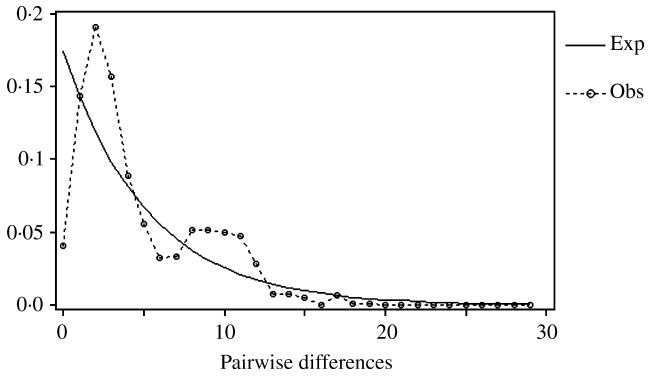

FIG. 3. Mismatch distribution analysis of *Aphyocypris kikuchii* mtDNA haplotype sequences in the longitudinal valley. A simulated Poisson distribution expected under constant population size is indicated by a dotted line.

DISCUSSION

GENETIC VARIATION IN *A. KIKUCHII*

The drainage system of Taiwan, like many other subtropical islands, is characterized by a steep topography, which results in short rivers (Han *et al.*, 2000). Water flow is high in summer because of typhoons but very low in winter. Fish in Taiwanese streams and rivers, therefore, experience a wide range of environmental conditions annually that influence population sizes. These population fluctuations can potentially lead to a loss of genetic diversity. The nucleotide

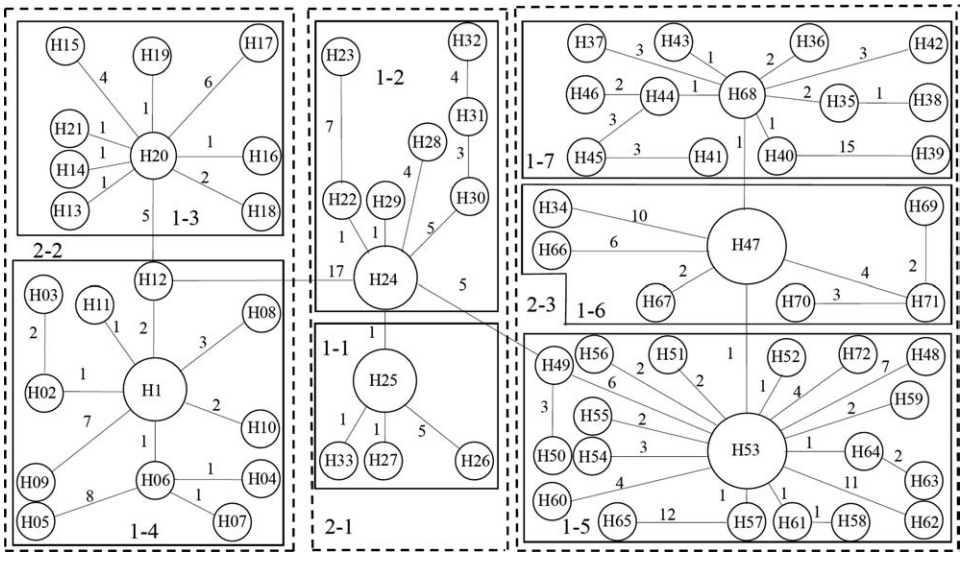

FIG. 4. Minimum spanning network based on mutations between haplotypes observed in seven populations of *Aphyocypris kikuchii*.

diversity observed in *A. kikuchii* is considered low and is similar to that in species suffering a severe historical bottleneck, *e.g. V. barbatulus* (Wang *et al.*, 2004) and *Rhinogobius maculafasciatus* Chen & Shao, 1996 (Cheng *et al.*, 2005). Diversities of *V. barbatulus* ($\theta = 0{\cdot}010$) (Wang *et al.*, 2004) and *Acrossocheilus paradoxus* (Günther, 1868) ($\theta = 0{\cdot}010$) in Taiwan (Wang *et al.*, 2000) and of *Zacco platypus* (Temminck & Schlegel, 1846) ($\theta = 0{\cdot}008$) in Mainland China (Perdices *et al.*, 2004), as well as *A. kikuchii*, are much lower than those in other Cyprinidae fishes, such as *Leuciscus souffia* Risso, 1827) ($\theta = 0{\cdot}056$) (Salzburger *et al.*, 2003). These comparisons indicate that populations of *A. kikuchii* may periodically experience bottlenecks in population size that lead to the loss of genetic diversity.

PHYLOGEOGRAPHY OF *A. KIKUCHII* AND MOLECULAR DATING

The characterization of phylogeographical patterns of *A. kikuchii* was used to test hypotheses of population structure in Taiwan. The haplotype network and results of AMOVA all indicated significant geographical structuring, with three major regions, northern, SL and southern population groups. Analysis of molecular variance (AMOVA) indicated that most of the genetic variation was found among these regions (43% of total variation), and the lack of shared haplotypes between regions indicates a lack of recent connection among geographical regions. The monophyly and relatively deep divergence of the three major clades suggest that gene flow between these clades is restricted and that they can be regarded as evolutionarily independent. Nm values between populations estimated from the F_{ST} values between populations are significantly lower than one fish per generation, except between populations in the longitudinal valley and CK. Restricted gene flow and significant genetic differentiation among populations of different freshwater systems indicate the presence of effective barriers to movement between populations. Hence, recent gene flow is unlikely to be an important factor in producing the patterns of deep genetic population structure observed in this study.

Historically, the Hualien, Hsiukuluan and Beinan Rivers in the longitudinal valley between the Central Mountain Range and the Coastal Mountain Range were connected in the same drainage in the past but have become separated from one another only recently (Lin, 1966). The low level of genetic differentiation between CK of the southern district of Coastal Mountains and populations of the Longitudinal Valley is consistent with a connection between these drainages at various times, possibly by head-water capture.

While recent gene flow may explain some patterns of divergence between populations within regions, divergences at larger geographical scales may be because of ancient colonizations or geological events. The separation of the major mtDNA clades by coalescent times less than *c.* 500 000 years indicates that events post-dating the formation of Taiwan have shaped the genetic population structure of this species. Geographically, the Central Mountain Range separates the northern region from SL and southern groups. SL population and the southern group are further divided by the Coastal Mountain Range. In the northern group, Shei Mountain Range must have also represented a vicariant barrier between LY and GL populations. Nevertheless, there was lack of

significant genetic differentiation between populations in the Longitudinal Valley, *i.e.* HD, SK, and BN, and CK of the southern district of East Coast Mountains.

HISTORICAL DEMOGRAPHY

This study examined haplotype variation using the approach of Grant & Bowen (1998), in which patterns of haplotype (h) and nucleotide (π) diversity reflect historical demographic events. Theoretically, low levels of nucleotide diversity and high levels of haplotype diversity are indicative of the loss of haplotype diversity during a population bottleneck and an accumulation of new haplotypes during population expansions (Alves *et al.*, 2001). A rapid population recovery after a bottleneck reduces the loss of haplotype diversity and can lead to an excess of low-frequency haplotypes, relative to the expected distribution of haplotype frequencies at mutation-drift equilibrium (Avise *et al.*, 1984; Grant & Bowen, 1998), as previously reported for mtDNA in other organisms (Riginos & Nachman, 2001).

The pattern mtDNA diversity in *A. kikuchii* displayed large values of h and low values of π ($h > 0{\cdot}5$, $\pi < 0{\cdot}5\%$). This pattern could be attributed to demographic expansions in some regions after a period of low effective population size. Negative values of Tajima's D in the samples in general and the significantly negative values in samples from the longitudinal valley were associated with a large number of rare variants. These results are consistent with an historical demographic perturbation and indicate a possible recent expansion in population size, as hypothesized for *V. barbatulus* and *A. paradoxus* (Wang *et al.*, 2000, 2004). The mismatch distribution among haplotypes in samples of *A. kikuchii* from the longitudinal valley is also consistent with a model of population expansion (Fig. 3) (Slatkin & Hudson, 1991; Rogers & Harpending, 1992). A population expansion also evidenced by the 'star-like' phylogeny of the haplotype network observed in the lineage C.

According to coalescence theory, the interior position in the minimum spanning network plus a high frequency suggest that lineage B (clade 2-1) in populations eastward of the Coastal Mountain Range represents an ancestral clade. This apparently older central lineage would have a greater probability of producing mutational derivatives (Crandall & Templeton, 1993), which can reveal historical migration routes. According to the network topology, there are at least two major migratory routes in Taiwan, one from the Shueilian River (SL, clade 2-1) through the Lanyang River (LY, clade1-4) to the Keelung River (GL, clade1-3), and the other from the Shueilian River to the longitudinal valley group, including the Hualien (HD), Hsiukuluan (SK) and Beinan (BN) rivers and the Chungkung River (CK) (clade 2-3).

Likelihood analyses of the clade's ages revealed that *A. kikuchii* coalesced to MRCA 0·11–0·39 million years ago in Taiwan, a time close to the divergence of *V. barbatulus* populations in the same region. Concordant results for the geographical divergence support a common history for these co-distributed freshwater taxa (Wang *et al.*, 2004). Geological evidence indicates that the formation of the Central Mountain Range via collision of the Philippine Sea and continental Asian Tectonic Plates (Page & Suppe, 1981) occurred about 1 million years before present (Lin, 1966). In the present study, molecular clock estimates did

not support such an ancient vicariance hypothesis. These estimates suggest that climatic oscillations during Pleistocene may have played an important role in the genetic subdivision of populations. Concordant phylogeographic patterns with other fishes and similar population age estimates suggest that similar events have shaped the distributions and genetic population structures of several Taiwanese freshwater fishes (*e.g. V. barbatulus*; Wang *et al.*, 2004).

CONSERVATION RECOMMENDATIONS

Moritz (1994) proposed that evolutionarily significant units (ESUs) should be reciprocally monophyletic for mtDNA haplotypes and should also differ significantly for the frequencies of alleles at nuclear loci. Populations that are not reciprocally monophyletic but that show significant divergence of allele frequencies at nuclear or mitochondrial loci constitute management units (MUs) that are appropriate for implementing short-term conservation measures.

The three geographically and genetically distinct groups identified in this study should be managed as distinct conservation units. At least two major ESUs, including the geographically isolated clades A and B + C, should be considered, and three lineages, including the northern, SL and southern groups, should be designated as distinct MUs. The SL population in particular warrants increased protection. However, the small SL river system will be difficult to conserve, as most of the catchment area is being developed for agriculture. Fish populations in rivers with small upstream drainage areas tend to experience strong annual contrasts because of the limited availability of dry season refuges and because of typhoon flooding in summer. These populations potentially experience high mortalities, reduced population sizes and losses of genetic diversity. In addition to these natural stressors, freshwater populations are also influenced by habitat degradation from continuing human development, introductions of alien species, overharvesting and water diversions. A history of stock transfers has increased ecological competition for resources in other cyprinid fishes, such as *Zacco pachycephalus* (Günther, 1868) and *A. paradoxus* in western Taiwan. The unique genetic diversity in populations of *A. kikuchii*, the restricted distributions of populations and the environmental variation in these rivers make it important to preserve habitats capable of supporting populations during droughts. Also important is the implementation of measures to prevent the dispersion of exotic species and the limitation of obstructions to population connectivity.

References

Alves, M. J., Coelho, H., Collares-Pereira, M. J. & Coelho, M. M. (2001). Mitochondrial DNA variation in the highly endangered cyprinid fish *Anaecypris hispanica*: importance for conservation. *Heredity* **87,** 463–473.

Avise, J. C., Neigel, J. E. & Arnold, J. (1984). Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution* **20,** 99–105.

Bermingham, E. & Martin, A. P. (1998). Comparative mtDNA phylogeography of neotropical freshwater fish: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology* **7,** 499–517.

Blin, N. & Stafford, D. W. (1976). A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Research* **3,** 2303–2308.
Cheng, H. L., Huang, S. & Lee, S. C. (2005). Morphological and molecular variation in *Rhinogobius rubromaculatus* (Pisces: Gobiidae) in Taiwan. *Zoological Studies* **44,** 119–129.
Crandall, K. A. & Templeton, A. R. (1993). Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134,** 959–969.
Excoffier, L. & Smouse, P. E. (1994). Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics* **136,** 343–359.
Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791.
Gelas, K. D. & Meester, L. D. (2005). Phylogeography of *Daphnia magna* in Europe. *Molecular Ecology* **14,** 753–764.
Grant, W. S. & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89,** 415–426.
Han, C. C., Tew, K. S., Chen, I. S., Su, L. Y. & Fang, L. S. (2000). Environmental biology of an endemic cyprinid, *Varicorhinus alticorpus*, in a subtropical mountain stream of Taiwan. *Environmental Biology of Fishes* **59,** 153–161.
Hasegawa, M., Kishino, H. & Yano, T. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **21,** 160–174.
Hillis, D. M. & Bull, J. J. (1993). An empirical test of bootstrapping as a method assessing confidence in phylogenetic analysis. *Systematic Biology* **42,** 182–192.
Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mammalian Protein Metabolism* (Munro, H. N., ed.), pp. 31–132. New York: Academic Press.
Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* **17,** 1244–1245.
Liao, T. Y., Lin, H. D., Tao, T. L. & Tzeng, C. S. (2005). The forgotten rare and endangered fishes in Taiwan. *Quarterlies Journal of Research* **50,** 52–58 (in Chinese).
Lin, C. C. (1966). An outline of Taiwan's quaternary geohistory with a special discussion of the relation between natural history and cultural history in Taiwan. *Bulletin of the Department of Archaeology and Anthropology* **23,** 7–44.
Lundberg, J. G. (1993). African–South American freshwater fish clades and continental drift, problems with a paradigm. In *Biotic Relationships Between Africa and South America* (Goldblatt, P., ed.), pp. 156–198. New Haven: Yale University Press.
Meffe, G. K. & Vrijenhoek, R. C. (1988). Conservation genetics in the management of desert fishes. *Conservation Biology* **2,** 157–169.
Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution* **9,** 373–375.
Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.
Nei, M. & Tajima, F. (1983). Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics* **105,** 207–217.
Page, B. M. & Suppe, J. (1981). The Pliocene lichi melange of Taiwan: its plate tectonic and olistostromal origin. *American Journal of Science* **281,** 193–227.
Perdices, A., Cunha, C. & Coelho, M. M. (2004). Phylogenetic structure of *Zacco platypus* (Teleostei, Cyprinidae) populations on the upper and middle Chang Jiang (= Yangtze) drainage inferred from cytochrome b sequences. *Molecular Phylogenetics and Evolution* **31,** 192–203.
Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14,** 817–818.
Rambaut, A. & Drummond, A. (2004). Tracer – MCMC Trace Analysis Tool. Oxford: Oxford University.
Riginos, C. & Nachman, M. W. (2001). Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous

habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology* **10,** 1439–1453.
Rogers, A. R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9,** 552–569.
Rozas, J. & Rozas, R. (1999). DnaSP version 3.0: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15,** 174–175.
Salzburger, W., Brandstatter, A., Gilles, A., Parson, W., Hempel, M., Sturmbauer, C. & Meyer, A. (2003). Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in Central Europe. *Molecular Ecology* **12,** 2371–2386.
Slatkin, M. & Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129,** 555–562.
Slechtova, V., Bohlena, J., Freyhofb, J., Persatc, H. & Delmastrod, G. B. (2004). The Alps as barrier to dispersal in cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. *Molecular Phylogenetics and Evolution* **33,** 225–239.
Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123,** 585–595.
Templeton, A. R., Boerwinkle, E. & Sing, C. F. (1987). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics* **117,** 343–351.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25,** 4876–4882.
Tzeng, C. S. (1986). Distribution of the freshwater fishes of Taiwan. *Journal of Taiwan Museum* **39,** 127–146.
Wang, J. P., Hsu, K. C. & Chiang, T. Y. (2000). Mitochondrial DNA phylogeography of *Acrossocheilus paradoxus* (Cyprinidae) in Taiwan. *Molecular Ecology* **9,** 1483–1494.
Wang, J. P., Lin, H. D., Huang, S., Pan, C. H., Chen, X. L. & Chiang, T. Y. (2004). Phylogeography of *Varicorhinus barbatulus* (Cyprinidae) in Taiwan based on nucleotide variation of mtDNA and allozymes. *Molecular Phylogenetics and Evolution* **31,** 1143–1156.
Yu, H. S. & Song, G. S. (1994). Oblique convergence between trench-arc system and continent-implications for the formation of the submarine physiography off southern Taiwan. *Acta Oceanographica Taiwanica* **32,** 45–53.

Electronic References

Drummond, A. J. & Rambaut, A. (2003). BEAST. Version 1.3. Oxford: Oxford University Press. Available at http://evolve.zoo.ox.ac.uk/beast/
Schneider, S., Roessli, J. D. & Excoffier, L. (2000). ARLEQUIN Version 2.000. *A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory*. Switzerland: University of Geneva. Available at http://anthro.unige.ch/arlequin.

