Molecular Phylogenetics and Evolution
Vol. 18, No. 2, February, pp. 163–173, 2001
doi:10.1006/mpev.2000.0879, available online at http://www.idealibrary.com on IDEAL®

Molecular Systematics of Xenocyprinae (Teleostei: Cyprinidae): Taxonomy, Biogeography, and Coevolution of a Special Group Restricted in East Asia

Wuhan Xiao,*,†[1] Yaping Zhang,† and Huanzhang Liu*

*Institute of Hydrobiology, the Chinese Academy of Sciences, Wuhan, 430072, People's Republic of China; and †Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming, 650223, People's Republic of China

Received September 10, 1999; revised October 2, 2000; published online December 26, 2000

We surveyed mitochondrial DNA (mtDNA) sequence variation in the subfamily Xenocyprinae from China and used these data to estimate intraspecific, interspecific, and intergeneric phylogeny and assess biogeographic scenarios underlying the geographic structure of lineages. We sequenced 1140 bp of cytochrome *b* from 30 individuals of Xenocyprinae and one putative outgroup (*Myxocypris asiaticus*) and also sequenced 297 bp of ND4L, 1380 bp of ND4, 68 bp of $tRNA^{His}$, and 69 bp of $tRNA^{Ser}$ from 17 individuals of Xenocyprinae and the outgroup (*M. asiaticus*). We detected high levels of nucleotide variation among populations, species, and genera. The phylogenetic analysis suggested that *Distoechodon hupeinensis* might be transferred to the genus *Xenocypris*, the taxonomic status of the genus *Plagiognathops* might be preserved, and species of *Xenocypris* and *Plagiognathops* form a monophyletic group that is sister to the genus *Distoechodon* and *Pseudobrama*. The introgressive hybridization might occur among the populations of *X. argentea* and *X. davidi*, causing the two species to not be separated by mtDNA patterns according to their species identification, and the process and direction of hybridization are discussed. The spatial distributions of mtDNA lineages among populations of *Xenocypris* were compatible with the major geographic region, which indicated that the relationship between Hubei + Hunan and Fujian is closer than that between Hubei + Hunan and Sichuan. From a perspective of parasite investigation, our data suggested that the fauna of *Hexamita* in Xenocyprinae could be used to infer the phylogeny of their hosts. © 2001 Academic Press

Key Words: **phylogeny; cytochrome *b*; ND4; ND4L; $tRNA^{His}$; $tRNA^{Ser}$; phylogeography; coevolution; *Xenocypris; Plagiognathops; Distoechodon; Pseudobrama; Hexamita*.**

INTRODUCTION

More and more researchers are interested in exploring the degree to which phylogeographic patterns of plant and animal taxa are related to historical changes in the environment for discovering origins of biodiversity (Lydeard *et al.*, 1995a). The fishes, especially the freshwater fishes, have provided key insights into the relationships between the historical changes in the environment and the biotic responses to those (Lydeard *et al.*, 1995a; Dodson *et al.*, 1995; Murphy and Collier, 1996). The distribution of fish is strictly constrained by water and faunal exchanges among different river systems are usually obstructed; thus, a deeper differentiation may result from historical changes. Many authors have tried hard to explore the problems and have gotten many useful clues (Lydeard *et al.*, 1995b).

In East Asia, some of the most dramatic changes in speciation, distribution, and connectivity of Cyprinidae (Pisces: Teleostei) occurred 5–30 million years ago, primarily in response to regional climate change associated with mountain building and river system changing (Liu and Su, 1962; Zhou, 1990).

One fish taxon, the subfamily Xenocyprinae (Cypriniformes, Cyprinidae), offers an excellent opportunity to evaluate not only the level of congruence exhibited by phylogenetic hypotheses derived from independent data sets but also the level of congruence exhibited by area cladograms derived from the different taxon cladograms (Yang, 1964). In addition, the subfamily not only has a discrete geographic distribution (restricted to East Asia, especially in China) and a long history (dating to the Early Miocene), but also its extant gen-

Sequences have been submitted to GenBank under Accession Nos. AF036165, AF036173–AF036176, and AF036194–AF036208 for the cytochrome *b* gene and AF036177–AF036193 for ND4L, ND4, $tRNA^{His}$, and $tRNA^{Ser}$ genes.

[1] To whom correspondence should be addressed at present address: Department of Urology, Tarry 11-715, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611. Fax: (312) 908-7275. E-mail: w-xiao2@nwu.edu.

1055-7903/01 $35.00

TABLE 1

Species, Geographic Locality, Sample Size, and Given Number in This Study

Species	Locality	No.	Haplotypes of Cytb	Haplotypes of ND4[a]
Pseudobrama simoni	Dongtinghu Lake, Hunan	1	P. simoni	P. simoni
Pseudobrama simoni	Wuhan, Hubei	1	P. simoni	P. simoni
Xenocypris fangi	Langzhong, Sichuan	1	X. fangi	
Xenocypris fangi	Langzhong, Sichuan	1	X. fangi	X. fangi
Xenocypris microlepis	Jiayu, Hubei	1	X. microlepis 1	X. microlepis 1
Xenocypris microlepis	Nanping, Fujian	1	X. microlepis 1	X. microlepis 2
Xenocypris microlepis	Xiantao, Hubei	1	X. microlepis 1	
Xenocypris davidi	Jingshan, Hubei	1	X. davidi 1	X. davidi 1
Xenocypris davidi	Jingshan, Hubei	1	X. davidi 2	
Xenocypris davidi	Langzhong, Sichuan	1	X. davidi 3	
Xenocypris davidi	Langzhong, Sichuan	1	X. davidi 3	X. davidi 3
Xenocypris davidi	Nanping, Fujian	1	X. davidi 4	
Xenocypris davidi	Jianou, Fujian	1	X. davidi 4	X. davidi 4
Xenocypris argentea	Dongtinghu Lake, Hunan	1	X. argentea 1	X. argentea 1
Xenocypris argentea	Dongtinghu Lake, Hunan	1	X. argentea 2	
Xenocypris argentea	Diaochahu Lake, Hubei	1	X. argentea 3	X. argentea 3
Xenocypris argentea	Diaochahu Lake, Hubei	1	X. argentea 4	
Xenocypris argentea	Langzhong, Sichuan	1	X. argentea 5	
Xenocypris argentea	Langzhong, Sichuan	1	X. argentea 5	X. argentea 5
Xenocypris argentea	Langzhong, Sichuan	1	X. argentea 5	
Xenocypris argentea	Langzhong, Sichuan	1	X. argentea 5	
Xenocypris yunnanensis	Dianci, Yunnan	1	X. yunnanensis	X. yunnanensis
Distoechodon tumirostris	Luoshan, Sichuan	1	D. tumirostris 1	
Distoechodon tumirostris	Luoshan, Sichuan	1	D. tumirostris 1	D. tumirostris 1
Distoechodon tumirostris	Nanping, Fujian	1	D. tumirostris 2	D. tumirostris 2
Distoechodon tumirostris	Jianou, Fujian	1	D. tumirostris 3	
Distoechodon hupeiensis	Liangzihu Lake, Hubei	1	D. hupeiensis	
Distoechodon hupeinensis	Liangzihu Lake, Hubei	1	D. hupeiensis	D. hupeiensis
Xenocyprioides carinatus	Nongzhou, Guangxi	1	X. carinatus	X. carinatus
Xenocyprioides parvulus	Qingzhou, Guangxi	1	X. parvulus	X. parvulus
Myxocyprinus asiaticus	Shashi, Hubei	1	M. asiaticus	M. asiaticus

[a] The 18 individuals chosen for ND4L–ND4 region sequencing were identified on the line of haplotypes of ND4.

era and species become dominant and continued to develop beginning in the Early Pliocene (Chang *et al.*, 1996), making it possible to assess differentiation of population and species in response to both deep and more recent historical environmental changes in East Asia. Furthermore, previous parasitological studies within the subfamily Xenocyprinae indicate that the subfamily has a special fauna of parasitic Hexamitidae, offering the opportunity to explore host/parasite coevolution and investigate the possibility of host phylogeny based on the fauna of parasite (Li and Nie, 1995; Xiao, 1997).

The subfamily Xenocyprinae comprises 10 or 12 nominal species and three or four nominal genera. Tchang (1959), Yang (1964), and Liu and He (1998) successively proposed a classification of Xenocyprinae. Their classifications differ considerably in the delineation of species and in the assignment of species to genera. Chen (1982) established a new genus (*Xenocyprioides*) in the subfamily. Cao and Meng (1991) attempted to alleviate some of the confusion and propose a phylogeny of the subfamily based on cladistic analysis of morphological characters.

The purpose of this study was threefold. First, we wanted to elucidate the phylogenetic relationship of species and genera and to alleviate some of the confusion in the classifications of Xenocyprinae. Second, we wanted to describe the phylogenetic pattern of a widespread species (*Xenocypris argentea* and *X. davidi*). Third, we wanted to investigate the possibility of inferring host phylogeny from parasite fauna and tried to illustrate the mechanism of coevolution.

MATERIALS AND METHODS

Fish Specimens

A total of 30 specimens representing all species and genera of Xenocyprininae were included in this study. *Myxocyprinus asiaticus* was chosen as an outgroup due to its clear relationship with family Cyprinidae (Wu, 1964). All of the specimens were chosen for complete cytochrome *b* (Cytb) sequencing and 18 of them were chosen for complete ND4L–ND4 region sequencing (according to the representative haplotypes of Cytb) (Table 1). *Xenocypris yunnanensis* collected in 1963 was

TABLE 2

Six Primers Used for Amplification and Sequencing of Fish Cytochrome *b* Genes

Name of primer	Sequence
L14724	5′-GACTTGAAAAACCACCGTTG-3′
L15138	5′-ATGATGACCGCCTTCGTGGGCTA-3′
L15519	5′-GGAGACCCAGAAAACTTTACCCC-3′
H15149	5′-CCTCAGAAGGATATTTGTCCTC-3′
H15560	5′-GCGTAGGCAAATAGGAAGTATC-3′
H15915	5′-CTCCGATCTCCGGATTACAAGAC-3′

Note. The position of 3′ end oligonucleotide of each primer is given relative to the published sequence of human mtDNA (Anderson *et al.,* 1981).

preserved in 10% formalin in the Museum of Freshwater Fishes, Kumming Institute of Zoology, the Chinese Academy of Sciences (CAS) (Collection No. 63025). *Xenocyprioides carinatus* and *X. parvulus* collected in 1981 were preserved in 10% formalin in the Museum of Freshwater Fishes, Institute of Hydrobiology, CAS (Collection No. 81-VIII-1118). The others were collected from their distributed regions during 1994–1995, temporarily preserved on ice in the field, and then taken to the lab and stored at −70°C. After the research, all specimens were preserved in 4–10% formalin and stored in the Museum of Freshwater Fishes, Kunming Institute of Zoology, CAS.

DNA Sequences

Total cellular DNA was isolated from frozen tissue samples by standard proteinase K digestion followed by phenol/chloroform extraction. A developed protocol was employed to isolate total cellular DNA from fixed tissues (Xiao *et al.,* 1997). Two segments of the mitochondrial genome were amplified via the polymerase chain reaction (PCR) using two pairs of primers and sequenced using 11 internal sequencing primers (Tables 2 and 3). For formalin-fixed specimens, internal primers were used to amplify and sequence the different short fragments of these genes (average 400 bp) (Fig. 1).

The sequenced regions correspond to 1140 bases of cytochrome b, 297 bases of ND4L, 1380 bases of ND4, 68 bases of histidine tRNA, and 69 bases of serine tRNA. Amplification conditions consisted of 40 thermal cycles of a 1-min denaturation at 94°C, a 1-min annealing at 50°C, and a 1- to 20-min extension at 72°C, followed by a 5- to 10-min extension at 72°C.

PCR products were purified in 1.5–2.0% SeaPlaque agarose (FMC) and sequenced using the FS-*Taq* dye deoxy terminator cycle sequencing kit (Applied Biosystems Inc.) with an automated DNA sequencer (Applied Biosystems 377) following manufacturer instructions. To ensure accuracy, strands were sequenced in both directions for each individual.

Sequence Alignment and Phylogenetic Analysis

Sequences were read from one strand and aligned using PC/GENE program Version 6.6 (IntelliGenetics) and checked by eye to each other and also to the published sequence of *Cyprinus carpio* (Chang *et al.,* 1994). All sequences have been deposited with EMBL/GenBank Data libraries under Accession Nos. AF036165 and AF036173–AF036208. Pairwise sequence comparisons to determine the distribution and amount of variation and the degree of saturation by codon position were performed using MEGA 1.0 (Kumar *et al.,* 1993). Seventeen cytochrome *b* gene sequences were chosen for combined with ND4L–ND4 region sequences (Table 1). Phylogenetic analyses of the subfamily were performed using aligned sequences for all genes combined. For phylogeographical analyses, we separately used the sequence data of Cytb and ND4 region of *X. argentea* and *X. davidi* in the phylogenetic reconstruction.

Three methods for phylogenetic reconstruction were used, maximum parsimony (MP) (Swofford, 1993), maximum likelihood (ML) (Felsenstein, 1993), and neighbor joining (NJ) (Saitou and Nei, 1987), in combination with various character weighting schemes. Each base position was treated as an unordered character with four alternative states. In all analyses, trees were rooted using outgroup comparisons (Watrous and Wheeler, 1981), with the most distant species (*M. asiaticus*) as the designated outgroup (Wu, 1964).

Maximum-parsimony phylogenies were estimated using PAUP3.1.1 (Swofford, 1993). In order to search for most-parsimonious trees, three weighting schemes were used: (1) equal weighting for all characters, (2) transitions and transversions weighted separately, and (3) only transversion position weighting. In the first scheme, all substitutions were weighted equally, regardless of codon position or substitution type. In the second scheme, transitions (ts) were downweighted relative to transversions (tv) by a factor of 5 (ts:tv = 1:5)

TABLE 3

Nine Primers Used for Amplification and Sequencing of Fish ND4 Genes

Name of primer	Sequence
L10962	5′-AAGACCTCTGATTTCGGCTC-3′
L11282	5′-CAATGCTAAAAGTTCTAATCCC-3′
L11961	5′-GCAGGGTCCATAGTACTAGC-3′
L12363	5′-CCACTAACAGCAGTCTGATG-3′
H11263	5′-GTAGGAGATTAAGGTTTG-3′
H11709	5′-TAGGTTCCTGCGTTGAGGCG-3′
H12003	5′-TTCGTATTATTCCGTATCC-3′
H12414	5′-GTAATGATTATTAGTTCTCC-3′
H12849	5′-ACTTGGATTTGCACCAAGAG-3′

Note. The position of 3′ end oligonucleotide of each primer is given relative to the published sequence of *Cyprinus carpio* mtDNA (Chang *et al.,* 1994).

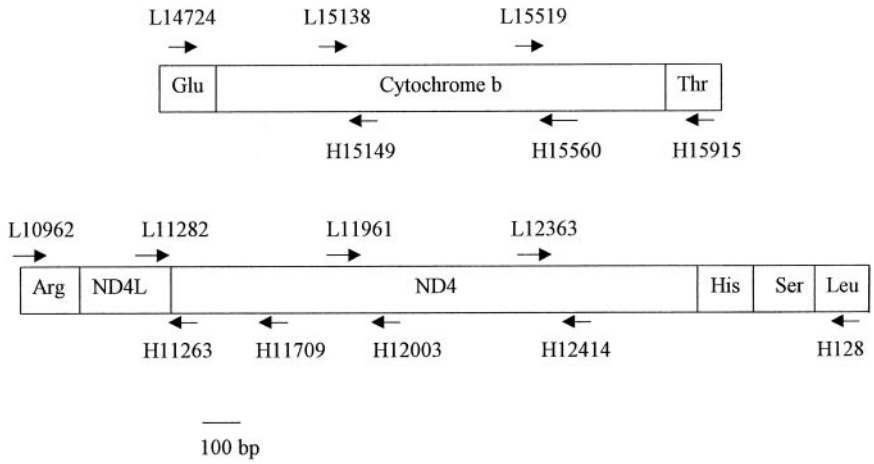

FIG. 1. Strategy for amplification and sequencing of the cytochrome *b* gene and the ND4L–ND4 gene. Arrows denote primers. Numbering is according to the human mtDNA sequences (Cytb) (Anderson *et al.*, 1981) and the carp mtDNA sequences (ND4L–ND4) (Chang *et al.*, 1991).

or 10 (ts:tv = 1:10). The DNAML program in PHYLIP 3.5 (Felsenstein, 1993) was used to reconstruct and evaluate ML trees. Neighbor-joining phylogenies were estimated using MEGA 1.0 (Kumar *et al.*, 1993). The reconstruction were based on Kimura's two-parameter distance (Kimura, 1981). Different character weightings and inclusion schemes were also used with the program according to the analysis of the nucleotide substitution patterns. Bootstrapping (Felsenstein, 1985) was used to estimate relative support for clades (1000 pseudoreplications for both parsimony and neighbor joining).

Investigation of Parasites

The investigations of the fauna of Hexamita parasitic in Xenocyprinae were taken during 1991–1995. All species in the subfamily were investigated except *X. yunnanensis*, *X. fangi*, *X. carinatus*, and *X. parvulus*, because fresh specimens were difficult to collect or their parasites were difficult to observe in the field. To ensure accuracy of the parasite fauna, 10 to 30 individuals of one population and one to five different populations of one species were chosen for observation. The methods of protargol dying, light and electron microscopic observations, and classification of parasitic *Hexamita* were described (Li and Nie, 1995).

We defined the occurrence of each species of parasitic *Hexamita* as a host character (absent = 0, present = 1) (see Table 4). The phylogeny of hosts was reconstructed using PAUP3.1.1 (Swofford, 1993) based on this character matrix. Bootstrap analyses were based on 100 replicates to estimate confidence in the results.

RESULTS

Sequence Variation

For 31 individuals, we sequenced the complete cytochrome *b* gene and identified 20 haplotypes (see Table 1). Eighteen individuals were chosen (according to the representative haplotypes of Cytb) for ND4 regions sequencing and 17 haplotypes were identified.

TABLE 4

Matrix of Distribution of 17 Species of Parasitic *Hexamita* in 6 Species of Xenocyprinae

	Parasites[a]																
Host	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
X. davidi	1	0	0	1	1	1	1	0	1	1	0	0	1	0	0	0	0
X. argentea	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1
X. microlepis	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
D. hupeinensis	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1
D. tumirostris	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P. simoni	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note. Present = 1; absent = 0.

[a] 1, *H. nobilis;* 2, *H. variformis;* 3, *H. polymorphola;* 4, *H. longiformis;* 5, *H. xenocyprii;* 6, *H. rodiformis;* 7, *H. honghuensis;* 8, *H. oviformis;* 9, *H. globulus;* 10, *H. capsularis;* 11, *H. axostyles;* 12, *H. giganti;* 13, *H. liangzihuensis;* 14, *H. guanqiaoensis;* 15, *H. transparenta;* 16, *H. wuchangensis;* 17, *H. vesiformis.*

TABLE 5

Percentage of Informative Sites by Gene and Codon Positions across the 17 Haplotypes, A + T Percentages, and Bias for the Same Regions

Region	% Informative	No. sites	% A + T	Bias[a]
Cytochrome *b*				
First codon	7.89	380	48.8	0.016
Second codon	1.05	380	61.6	0.219
Third codon	50.79	380	60.3	0.368
All	19.9	1140	56.9	0.137
ND4L				
First codon	5.05	99	44.8	0.069
Second codon	2.02	99	54.7	0.257
Third codon	36.36	99	62.3	0.403
All	14.48	297	54.0	0.123
ND4				
First codon	10.43	460	51	0.116
Second codon	0.65	460	57.1	0.246
Third codon	54.13	460	64.8	0.339
All	21.74	1380	57.6	0.145

[a] Bias in base composition is calculated as $C = (2/3)\Sigma (C_i - 0.25)$, where C is the compositional bias and C_i is the frequency of the ith base (Irwin *et al.*, 1991).

Combining the cytochrome *b* gene and ND4L–ND4 region, 2952 characters were obtained excluding insertions and deletions. Of these, 638 were phylogenetically informative. Percentages of A + T composition, informative sites, and nucleotide composition bias in the Cytb, ND4L, and ND4 gene fragments are shown in Table 5.

Levels of sequence divergence (uncorrected) between outgroup and ingroup lineage ranged from 19.0% (between *M. asiaticus* and *X. microlepis*1) to 20.1% (between *M. asiaticus* and *X. parvulus*). Percentage of sequence divergence among ingroup taxa ranged from 0.17% (between *X. davidi*1 and *X. davidi*4) to 14.26% (between *Pseudobrama simoni* and *X. yunnanensis*). The average transition/transversion ratio across all pairwise sequence comparisons in the data set is 5.35. This level of transition bias is within the range of biases previously reported for other vertebrates and serves as a basis for the transition/transversion weighting ratios used in phylogenetic reconstruction. Among the three genes, third codon positions of ND4 contain the most informative sites (54.13%) and the most A + T biased (64.8%).

The number of nucleotide substitutions per site was calculated for the combined data using Kimura two-parameter distance. The distances for the combined data set were from 0.003 to 0.247. The smallest distances always occurred between the two haplotypes within each of two species (*X. argentea* and *X. davidi*). The largest distances included the majority of the comparisons of *M. asiaticus* with all other taxa (data not shown).

Phylogenetic Analysis

Initially, we reconstructed the phylogenetic trees using Cytb, ND4L, and ND4 respectively by parsimony (Swofford, 1993). Cytb yielded 12 trees, whereas ND4L and ND4 yielded 12 and 1 tree (results not shown). The topology of the strict consensus tree of 12 Cytb gene trees was the same as that of ND4 gene tree, but in the strict consensus tree of 12 ND4L gene trees, many taxa could not be separated clearly (results not shown). Finally, we combined the data for further phylogenetic tree reconstruction (Zhang, 1996).

Phylogenetic relationships among Xenocyprinae. The phylogenetic trees reconstructed by three methods (MP, ML, NJ) were not significantly different and almost had a same topology except a little difference existing in the relationships among *X. carinatus, X. parvulus* and *X. yunnanensi* (Fig. 2). Monophylies of *X. argentea* + *X. davidi, X. argentea* + *X. davidi* + *Distoechodon hupeinensis, X. microlepis* + *X. fangi,* and *X. argentea* + *X. davidi* + *D. hupeinensis* + *X. microlepis* + *X. fangi* were well supported. *X. argentea* and *X. davidi* were most closely related, and two of their haplotypes were paraphyletic. It is difficult to separate them into two groups according their species identification except the populations distributed in Hubei, Hunan, and Sichuan. The clade of *D. hupeinensis,* representing a species of genus *Distoechodon,* grouped with the clade formed by the six haplotypes of *X. argentea* and *X. davidi,* rather than with another species of genus *Distoechodon, Distoechodon tumirostris. X. microlepis* was most closely related to *X. fangi* and together constituted a sister group of the clade formed by *X. argentea, X. davidi,* and *D. hupeinensis.* Although the relationship between *D. tumirostris* and *P. simoni* was relatively close, the monophyly of *D. tumirostris* and *P. simoni* was relatively little supported both by parsimony and by NJ with relatively lower bootstrap values (69 and 76, respectively). The monophyly of the clade formed by *D. tumirostris* and *P. simoni* and the clade formed by *X. davidi, X. argentea, D. hupeinensis, X. fangi,* and *X. microlepis* was well supported by NJ with high bootstrap values (96), but had a relatively lower bootstrap value (74).

The positions of *X. parvulus* and *X. carinatus* in MP tree were the same as those of ML tree, but were different from that of NJ tree (bootstrap values were <50). As suggested by the bootstrap analysis in MP and NJ, the relationships among *X. parvulus, X. carinatus,* and *X. yunnanensis,* and even among these three species with other seven species in the subfamily Xenocyprinae, were not resolved in the study.

Phylogenetic relationships among populations of X. davidi and X. argentea. For eight individuals of *X. argentea* collected from three locations, we identified five haplotypes by Cytb sequences and three haplotypes for three individuals by ND4L–ND4 region se-

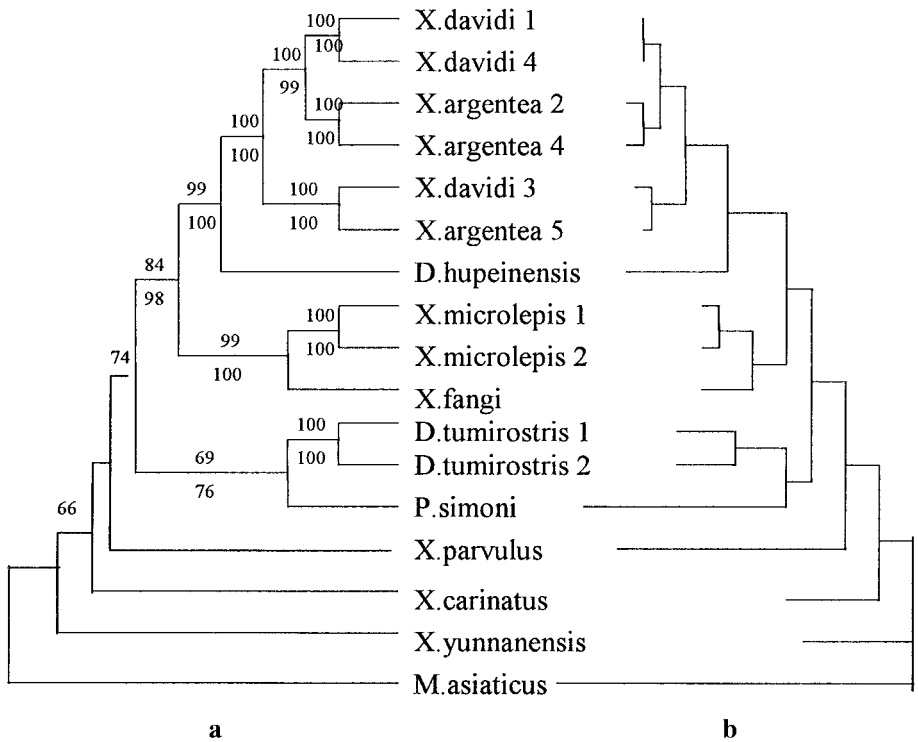

FIG. 2. Phylogenetic trees based on combined Cytb, ND4L, ND4, $tRNA^{Ser}$, and $tRNA^{His}$ data sets. *Myxocyprinus asiaticus* was treated as outgroup. (a) Maximum-parsimony tree and neighbor-joining tree. L = 2034, CI = 0.497 (excluding uninformative characters), and RI = 0.573. Bootstrap support is shown only for those branches where values were > 50 (parsimony values above branches; neighbor-joining bootstrap values below branches). (b) Maximum-likelihood tree. All codon positions weighted equally and transition:transversion weighting of 1:5 and 1:10. All significant branches are drawn proportional to branch lengths reported in the ML analysis.

quences. Four haplotypes were identified for six individuals of *X. davidi* collected from three locations by Cytb sequences and three haplotypes for three individuals by ND4L–ND4 region sequences (Table 1). Because the clades formed by the haplotypes of *X. argentea* and *X. davidi* were indistinguishable, thus, using the combination of all haplotypes of the two species we performed the analyses of phylogenetic relationships among populations. *D. hupeinensis* is the sister group of the clade formed by the haplotypes of *X. argentea* and *X. davidi*, so it was used as the outgroup for this data set. The NJ and UPGMA methods in MEGA 1.0 (Kumar *et al.*, 1993) were used for the phylogenetic reconstruction of population based on Kimura's two-parameter distance. The Cytb data and ND4L–ND4 region data were used respectively for analyses, generating NJ trees and UPGMA trees with bootstrap values based on 1000 replicates.

The topologies of the trees reconstructed by two methods and by using two data sets were congruent (Fig. 3), but differed in support by bootstrap values. The populations of *X. argentea* distributed in Dongtinghu Lake, Hunan Province, and Diaochahu Lake, Hubei Province, were most closely related. The haplotypes of *X. davidi* collected from Jingshan, Hubei, were clustered to the haplotypes of *X. davidi* distributed in Fujian and formed a clade to be a sister group of the clade formed by the populations of *X. argentea* distributed in Hunan and Hubei. The populations of *X. davidi* and *argentea* distributed in Jialingjing River, Sichuan, were most closely related and formed a clade to be a sister group of the clade formed by other populations of *X. argentea* and *X. davidi*.

Host Phylogenetic Relationships Inferred from Parasite Fauna

There were no parasitic *Hexamita* found in *P. simoni* and *D. tumirostris;* thus, we chose these two host species as outgroups for reconstruction of host phylogenetic tree. The tree length distribution for 10^3 randomly generated trees was significantly skewed to the left ($g_1 = -0.760551$), suggesting good phylogenetic signal in the data (Hillis and Huelsenbeck, 1992). Phylogenetic analysis of 17 characters weighted equally resulted in only one most parsimonious tree (CI = 0.895, HI = 0.105, RI = 0.696). Bootstrap values of each clade were higher than 50 (Fig. 4). The phylogenetic tree topology of Xenocyprinae, reconstructed by the analyses of fauna of parasitic *Hexamita* is similar to the phylogenetic trees constructed by the analyses of Cytb and ND4L–ND4 sequence data of the subfamily. The

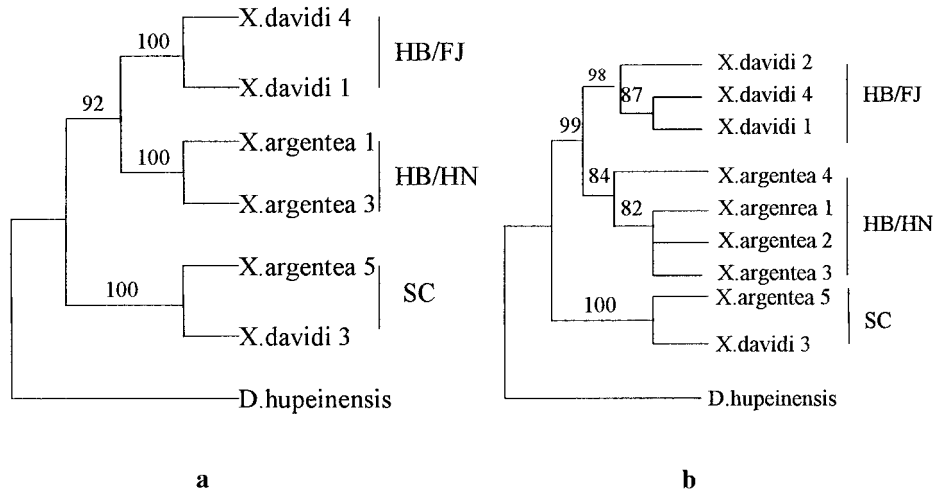

FIG. 3. NJ trees and UPGMA trees for populations of *X. davidi* and *X. argentea* constructed by MEGA 1.0. (a) Trees for the ND4L–ND4 region sequences. Neighbor-joining bootstrap values were shown on branches. (b) Trees for the Cytb sequences. Neighbor-joining bootstrap values were shown on branches.

monophylies of *X. argentea* + *X. davidi*, *X. argentea* + *X. davidi* + *D. hupeinensis*, and *X. argentea* + *X. davidi* + *D. hupeinensis* + *X. microlepis* were supported with bootstrap values exceeding 50.

DISCUSSION

Sequences Variation and Phylogenetic Information

The pattern of nucleotide composition among the cytochrome *b* sequences of Xenocyprinae is different from that observed in other studies of mammals, Perciformes, sharks, and newts (Irwin *et al.*, 1991; Cantarore *et al.*, 1994; Martin, 1995; Caccone *et al.*, 1997). At third codon positions, the nucleotide bias of Xenocyprinae (average 0.37) is higher than that of Perciformes (average 0.31) and sharks (average 0.33), but lower than that of mammals (average 0.40) and birds (average 0.45) (cited from Cantarore *et al.*, 1994). For each of the three genes, the nucleotide bias at different codon position was clearly different. At first positions, there was almost equal frequency of the four nucleotides. By contrast, at third positions, the frequencies A $\approx$ C > T >> G. The highest bias against G content occurred at third codon positions. It has been suggested that selection might restrict nucleotide frequencies at third codon positions (Martin, 1995).

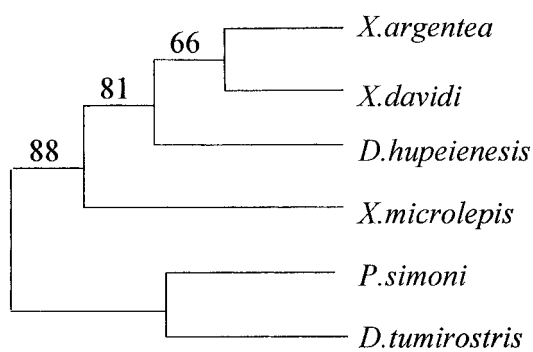

FIG. 4. Strict consensus parsimony tree of Xenocyprinae obtained using PAUP 3.1.1. Based on the fauna of parasitic *Hexamita*. Bootstrap values are shown above branches. CI = 0.895; RI = 0.696.

A high transition bias is well known in vertebrate mitochondrial DNA (mtDNA) (reviewed by Meyer, 1993). Because of this high bias, transitional substitutions reach saturation quickly and may be of limited use in phylogenetic reconstruction. In such instances, using only transversions (or weighting them more heavily than transitions) may be preferred because increased divergence transversions seem to accumulate more linearly over time (Irwin *et al.*, 1991). Sometimes, the transitions of third codon positions have been excluded from phylogenetic analyses. These notions have been recently challenged by several reports (Hedges *et al.*, 1992; Simon *et al.*, 1996). In this research, using only transversions and weighting them more heavily than transitions not only decreased the bootstrap values, but also collapsed phylogenetic relationships among the taxa (only using transversions) (results not shown). It would appear that the transitions of the data set in this research might contain some phylogenetic information. Because of these observations, we tend to favor equal weighting of transitions and transversions for closely related taxa.

Among the three gene fragments, the ND4 gene had the most informative sites and ND4L had the least. Combined with the results of phylogenetic analyses, we are in agreement with Zardoya and Meyer's opinions (Zardoya and Meyer, 1996).

Taxonomic and Phylogenetic Analyses and Introgression Consideration

In all cases, the haplotypes of two species, *X. argentea* and *X. davidi,* cannot be diagnosed by mtDNA patterns according to their species identification except the populations distributed in Hubei, Hunan, and Fujian. However, the discrimination of these two species could be accomplished readily on the basis of morphological criteria (Yang, 1964). A similar phenomenon also had been noted for subspecies or species of bluegill sunfish, walleye and saugar, bluefin tunas and yellowfin tuna (Avise and Saunders, 1984; Billington *et al.*, 1988; Chow and Kishino, 1995). One explanation of this phenomenon found in the two species of *Xenocypris* may be interspecific transfer of mtDNA similar to that of the above-mentioned species, with *X. argentea* mtDNA introduced by hybridization and having replaced most of original mitochondrial lineages in populations of *X. davidi.* The results also showed that the population of *X. argentea* in Jialing jiang River, Sichuan, had diverged from the populations distributed in Hubei and Hunan quite long ago and some geographical barriers prevented gene exchange among these populations recently. The genetic distance between *X. argentea* and *X. davidi* across Sichuan was very small; the two species cannot be separated by their mtDNA patterns. Moreover, recent gene flow or dispersal may be present among the populations of *X. davidi* in Hubei and Fujian; therefore, their relationships are very close and it is difficult to separate them by their distribution.

The systematic position of *D. hupeinensis* based on phylogenetic analysis of DNA sequence in this study is puzzling, as *D. hupeinensis* was established as a member of *Distoechodon* based on its two rows of pharyngeal teeth (Yang, 1964), being completely separated from genus *Xenocypris* with three rows of pharyngeal teeth. However, all of the other evidence, including our DNA sequence data, RFLP analysis of mtDNA, and fauna of parasitic *Hexamita* (Xiao, 1997), supports the view that *D. hupeinensis* is closely related to *X. argentea* and *X. davidi,* but distant from the species of *Distoechodon*. Although the distribution of *D. hupeinensis* is restricted and can coexist with *X. argentea* and *X. davidi* in one same lake (Yang, 1964), it was hardly interpreted as the result of introgressive hybridization because the populations of *X. argentea* and *X. davidi* distributed in different regions (Hubei, Hunan, and Sichuan) can cluster together to form one monophyletic clade. The *D. hupeinensis* clade was distinct from the clade formed by all haplotypes of *X. argentea* and *X. davidi,* despite *D. hupeinensis* and *X. argentea*'s populations of Hubei distributed in the same lake (Liangzihu Lake, Hubei).

Except the row number of pharyngeal teeth, the other morphological criteria of *D. hupeinensis,* such as the morphology and structure of mouth, the length and width of body, and the color of body lateral, are similar to that of *X. argentea* and *X. davidi.* The two rows of pharyngeal teeth of *D. hupeinensis* might be a secondary derived character and functionally linked, as a result of parallel evolution. Therefore, we suggest that *D. hupeinensis* be transferred to the genus *Xenocypris.*

In all cases, not only was the monophyly of *X. microlepis* + *X. fangi* + *X. argentea* + *X. davidi* + *D. hupeinensis* well supported with high bootstrap values, but also the monophylies of *X. microlepis* + *X. fangi* and *X. davidi* + *X. argentea* + *D. hupeinensis* were well supported too. These data lead us to question the taxonomic status of *X. microlepis.* With the complete ventral ridge from ventral fin to anus, it had been designated a member of the genus *Plagiognothops* (Tchang, 1959; Yang, 1964). However, doubt concerning its taxonomic status also has been raised by several morphological studies (He, 1987; Liu and He, 1998). It was suggested that the genus *Plagiognathops* must be combined with the genus *Xenocypris* because the ventral ridge was thought not to be an important taxonomic character and could be varied among the species of the subfamily. Our results based on the phylogenetic analysis of DNA sequences might provide an evidence for supporting the taxonomic status of *Plagiognathops,* agreeing with isoenzyme and skeleton characteristic evidence (Cao and Meng, 1992). This study also shows that *X. microlepis* clusters with *X. fangi,* constituting members of the genus *Plagiognathops,* despite close relationships between *Xenocypris* and *Plagiognathops.* On the other hand, returning to reexamine the morphological criteria of *X. microlepis* and *X. fangi,* some of their common characters such as the length of ventral ridge, the number of lateral scale, and the size of body, could be distinguished readily from that of the other species belonging to the genus *Xenocypris.*

The systematic positions of *X. parvulus, X. carinatus,* and *X. yunnanensis* within Xenocyprinae could not be resolved with confidence in present study, and no close relationship of these three species with other Xenocyprinae species is suggested. The morphological characters of *Xenocyprioides* are very different from that of other species belonging to Xenocyprinae except the morphology of the pharyngeal teeth, but quite similar to that of *Aphyocypris* (Chen, 1982; Luo *et al.*, 1985). Chang *et al.* (1996) suggested that either their affiliation to Xenocyprinae would broaden the scope of the subfamily or their phylogenetic position should be reconsidered together with the value of the subfamily. Our results indicated the monophyly of Xenocyprinae could not be supported if the genus *Xenocyprinoides* was assigned to the subfamily; thus, the study based on DNA sequences provided further evidence for re-evaluating the systematic position of the genus *Xenocyprinoides.*

In present study, the systematic position of *X. yun-*

nanensis is unusual. Based on morphological data, it is closely related with *X. fangi* and has been assigned a member of the genus *Xenocypris*. However, as a result of analyzing DNA sequences, it is distantly related to other species of *Xenocypris*. The distribution of *X. yunnanensis* is quite narrow, only inhabiting in Dianci, Yunnan, where no other species of Xenocyprinae exists (Yang, 1964; Chen and Li, 1989). It is impossible to interpret the phenomenon as a vicariant event, dispersal, and introgressive hybridization, etc. An alternative explanation is that the mtDNA pattern found in *X. yunnanensis* resulted from stochastic sorting and differential extinction of mtDNA lineages from polymorphic *Xenocypris* ancestral stock (Avise and Saunders, 1984; Avise *et al.*, 1987). Obviously, the mechanism creating special mtDNA lineages of *X. yunnanensis* is currently unknown. Clearly, additional morphological and molecular data are needed to resolve the problem.

Phylogenetic Pattern and Historical Biogeography

The mtDNA lineages of *X. argentea* exhibit clear geographic structuring. Considering the introgressive hybridization between *X. argentea* and *X. davidi*, the mtDNA lineages of these two species might be combined to infer area relationships. All tree topologies in our analyses depict similar patterns. The drainages of Hubei (Liangzihu Lake and Diaochahu Lake) are closely related to the drainages of Hunan (Dongtinghu Lake). The drainages of Hubei + Hunan, however, are more closely related to the Fujian drainages (Mingjiang River) than to Sichuan drainages (Jialingjiang River), despite the connection of river system provided by the drainages of Hubei + Hunan and Jialing River through Yangtizi River by now. On the contrary, there is no direct connection of river system between Hubei + Hunan and Fujian. Our findings support Liu's hypothesis, based on phylogenetic analyses of Sinipercine by morphological characters and fossil evidence, that the middle and lower reaches of Yangtizi River were more closely related to Mingjiang River than to the upper reaches of Yangtizi River (Liu, 1993).

Two clades identified in our analyses are not geographically separated; both cases involve populations of HB/HN and HB/FJ clades. The HB/HN clade might be explained as a result of recent gene flow between two drainages through Yangtizi River, as indicated by habitat and ecology of *X. argentea*. However, there is no direct connection between Hubei and Fujian. Thus, the HB/FJ clade might indicate that the river capture took place no longer ago in the Eastern plain of China, as suggested by Chen *et al.* (1986) that being after Late Tertiary. Another possible explanation is that this continuity is the result of historical overflow of rivers in the eastern plain of China, allowing exchange of the fish fauna among rivers, decreasing genetic divergence.

Perhaps the geographical barriers, such as the Three Gorges, existing between upper reaches and middle reaches of Yangtizi River, prevented the gene flow among populations of *X. argentea* in the regions. Thus, the mtDNA lineages were completely separated according their geographical distribution. Furthermore, the results also indicated the formation of barriers between upper reaches and middle reaches of Yangtizi River predates the river capture among the eastern plain of China, in agreement with a suggestion made by Chen *et al.* (1986).

Coevolution of Parasites

Up to now, few works were inferring the elucidation of host relationships from their parasite fauna. Perhaps this lack of success is due to the inherent difficulties in finding a special group of parasites within a group of hosts, which provides the ability to investigate host phylogeny from parasite fauna.

Fortunately, our previous parasitological investigation indicated that there were 17 species of *Hexamita* living in the intestine of the subfamily Xenocyprinae and 0–14 different parasite species could live in one species of host. However, they were not found in the intestine of other common freshwater fishes (e.g., *C. carpio, Carassius auratus, Aristichthys nobilis, Hypophthalmichthys molitrix, Pseudorasbora parva, Ctenopharyngodon idellus, Hemiculter leucisculus,* and *Cultrichthys erythropterus*) which often coexist with the members of Xenocyprinae. This indicated that these 17 species of *Hexamita* were special parasites of Xenocyprinae. Thus, it offered an excellent opportunity to explore host phylogeny from parasite fauna.

The host phylogeny based on the parasite fauna was completely congruent with that of based on mtDNA sequence data with high bootstrap values. So, it can be concluded that if a number of hosts have a special parasite fauna, the composition of parasite fauna might reflect the phylogeny of host. Although the mechanism for explaining the phenomenon is unclear, it may also represent an aspect of coevolution. A possible explanation is that two host species were closely related, and they might have similar gene structure, causing the biochemical composition in their intestine and their physiological characters to be also similar, which might permit similar speciation of parasites to take place in their intestine, so the hosts have similar parasite fauna. On the other hand, an alternative explanation is that closely related hosts permit closely related parasites to infect because of the parasites being suitable for similar living conditions. Consequently, the fauna of parasites could also reflect the relationship of host.

ACKNOWLEDGMENTS

We are grateful to Guihua Cui, Shunping He, Yungao Liu, Qixiang Deng, Zhengdong Zhu, Guizhi Ding, and Rong Zheng for assistance

in collecting specimens or making available tissue samples in their care. Bing Su and Wen Wang generously helped in the experiment and the computer analyses and also provided much helpful suggestion. We especially thank Michael R. J. Forstner for constructive commenting on original versions of the manuscript. Financial support was provided by grants from National Natural Science Foundation of China (No. 39870128) to W.X.; the Foundation of the Key Laboratory of Cellular and Molecular Evolution, Kumming Institute of Zoology, CAS, to W.X.; and a part of NSFC grant to Y.Z.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., Debruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290:** 475–465.

Avise, J. C., and Saunders, N. C. (1984). Hybridization and introgression among species of sunfish (*Lepomis*) analysis by mitochondrial DNA and allozyme markers. *Genetics* **108:** 237–255.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., and Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Syst.* **18:** 489–522.

Billington, N., Hebert, P. D. N., and Ward, R. D. (1988). Evidence of introgressive hybridization in the genus *Stizostedion:* Interspecific transfer of mitochondrial DNA between sauger and walleye. *Can. J. Fish. Aquat. Sci.* **45:** 2035–2041.

Caccone, A., Milinkovitch, M. C., Sbordoni, V., and Powell, J. R. (1997). Mitochondrial DNA rates and biogeography in European newts (genus *Euproctus*). *Syst. Biol.* **46:** 126–144.

Cantarore, P., Roberti, M., Pesole, G., Ludovico, A., Millella, F., Gadaleta, M. N., and Saccone, C. (1994). Evolutionary analysis of cytochrome b sequences in some Perciformes: Evidence for a slower rate of evolution than in mammals. *J. Mol. Evol.* **39:** 589–597.

Cao, L., and Meng, Q. (1992). Studies on isozymes, skeleton characters of the Xenocyprinae fishes of China and discussion on systematic evolution. *Acta Zootaxon. Sinica* **17:** 366–376.

Chang, Y.-S., Huang, F.-L., and Lo, T.-B. (1994). The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. *J. Mol. Evol.* **38:** 138–155.

Chang, M., Chen, Y., and Tong, H. (1996). A new Miocene Xenocyprininae (Cyprinidae) from Helongjiang Province, Northeast China and succession of late Cenozoic fish faunas of East Asia. *Vert. Palasiat.* **34:** 165–183.

Chen, Y. (1982). Description of a new genus and species of cyprinid fish. *Acta Zootaxon. Sinica* **7:** 425–427.

Chen, Y., Cao, W., and Zheng, C. (1986). Ichthyofauna of the Zhujiang River with a discussion on zoogeographical divisions for freshwater fishes. *Acta Hydrobiol. Sinica* **10:** 228–236.

Chen, Y., and Li, Z. (1989). Xenocyprinae. *In* "The Fishes of Yunnan, China. I. Cyprinidae" (X. Chu, Y. Chen *et al.*, Eds.), pp. 93–98. Science Press, Beijiang.

Chow, S., and Kishino, H. (1995). Phylogenetic relationships between tuna species of the genus *Thunnus* (Scombridae: Teleostei): Inconsistent implications from morphology, nuclear and mitochondrial genomes. *J. Mol. Evol.* **41:** 741–748.

Dodson, J. J., Colombani, F., and Ng, P. K. L. (1995). Phylogeographic structure in mitochondrial DNA of a Southeast Asian fresh water fish, *Hemibagrus nemurus* (Siluroidei; Bagridae) and Pleistocene sea-level changes on the Sunda shelf. *Mol. Ecol.* **4:** 331–346.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Felsenstein, J. (1993). "PHYLIP: Phylogenetic Inference Package," version 3.5. Department of Genetics, Univ. Washington, Seattle, WA.

He, M. (1987). Xenocyprinae. *In* "Systematic Synopsis of Chinese Fishes" (Q. Cheng, and B. Zheng, Eds.), pp. 136–137. Science Press, Beijing.

Hedges, S. B., Bogart, J. P., and Maxsona, L. R, (1992). Ancestry of unisexual salamanders. *Nature* **356:** 708–710.

Hillis, D. M., and Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analysis. *J. Hered.* **83:** 189–195.

Irwin, D. M., Kocher, T. D., and Wilson, A. C. (1991). Evolution of the cytochrome b gene in mammals. *J. Mol. Evol.* **32:** 128–144.

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. USA* **78:** 454–458.

Kumar, S., Tamura, K., and Nei, M. (1993). "MEGA: Molecular Evolutionary Genetics Analysis," version 1.0. Pennsylvania State Univ., University Park, PA.

Li, L., and Nie, D. (1995). Notes on the parasitic Hexamitids of fishes with descriptions of 18 new species (Zoomastigophores: Diplomonadida: Hexamitidae). *Acta Zootaxon. Sinica* **20:** 6–28.

Liu, H. T., and Su, T. T. (1962). Pliocene fishes from Yushe basin, Shanxi. *Vert. Palasiat.* **6:** 1–25.

Liu, H. (1993). "Studies on Skeleton Anatomy and Phylogeny of the Sinipercine Fishes," Ph.D. dissertation. Institute of Hydrobiology, the Chinese Academy of Sciences, Wuhan.

Liu, H., and He, M. (1998) Xenocyprinae. *In* " Fauna Sinica, Osteichthyes, Cypriniformes II" (Y. Chen *et al.*, Eds.), pp. 208–223. Science Press, Beijing.

Luo, Y., Chen, Y., and Huang, H. (1985). Description of two new species of cyprinid fish in Guangxi. *Acta Hydrobiol. Sinica* **9:** 280–284.

Lydeard, C., Wooten, M. C., and Meyer, A. (1995a). Molecules, morphology, and area cladograms: A cladistic and biogeographic analysis of *Gambusia* (Teleostei: Poeciliidae). *Syst. Biol.* **44:** 221–236.

Lydeard, C., Wooten, M. C., and Meyer, A. (1995b). Cytochrome b sequence variation and a molecular phylogeny of the live-bearing fish genus *Gambusia* (Cyprinodontiformes: Poeciliidae). *Can. J. Zool.* **73:** 213–227.

Martin, A. P. (1995). Mitochondrial DNA sequence evolution in sharks: Rates, patterns, and phylogenetic inferences. *Mol. Biol. Evol.* **12:** 1114–1123.

Meyer, A. (1993). Evolution of mitochondrial DNA in fishes. *In* "The Biochemistry and Molecular Biology of Fishes"(P. W. Hochachka and T. P. Mommsen, Eds.), Vol. 2, pp. 1–38. Elsevier, Amsterdam.

Murphy, W., and Collier, G. E. (1996). Phylogenetic relationships with the Aplocheiloid fish genus *Rivulus* (Cyprinodontiformes, Rivulidae): Implications for Caribbean and Central American biogeography. *Mol. Biol. Evol.* **13:** 642–649.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Simon, C., Nigero, L., Sullivan, J., Holsinger, K., Martin, A., Grapputo, A., Franke, A., and Mcintosh, C. (1996). Large differences in substitutional patterns and evolutionary rates of the 12s ribosomal RNA genes. *Mol. Biol. Evol.* **13:** 923–932.

Swofford, D. (1993). "PAUP: Phylogenetic Analysis Using Parsimony," version 3.1.1s. Illinois Natural History Survey, Champaign, IL.

Tchang, T. L. (1959). "The Fauna of Chinese Cyprinoid Fishes." Advanced Education Press, Beijing.

Watrous, L. E., and Wheeler, Q. D. (1981). The outgroup comparison method of character analysis. *Syst. Zool.* **30:** 1–11.

Wu, H.-W. (1964). "The Fauna of Cyprinidae in China." Shanghai Science and Technology Press, Shanghai.

Xiao, W., Wu, C., Su, B., Zhang, Y., and Cui, G. (1997). DNA extracted from formalin-fixed *Xenocypris yunnanensis* and sequence analysis of its cytochrome b gene. *Zool. Res.* **18:** 242.

Xiao, W. (1997). "Molecular Systematics of the Subfamily Xenocyprinae (Teleostei, Cyprinidae) and Some Other Related Groups with Consideration of Coevolution between Xenocyprinae and Hexamitidae (Fllagellate, Diplomonadida), Ph.D. dissertation. Institute of Hydrobiology, the Chinese Academy of Sciences, Wuhan.

Yang, G. (1964). Xenocyprininae. *In* "The Fauna of Chinese Cyprinidae Fishes" (H.-W. Wu, Ed.), pp. 121–136. Science Press, Beijing.

Zardoya, R., and Meyer, A. (1996). Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol. Biol. Evol.* **13:** 933–942.

Zhang, Y.-P. (1996). DNA sequence and species tree. *Zool. Res.* **17:** 247–252.

Zhou, J. (1990). The Cyprinidae fossils from middle Miocene of Shanwang. *Vert. Palasiat.* **28:** 95–127.