



Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



The phylogenetic placement of sinipercid fishes (“Perciformes”) revealed by 11 nuclear loci

Chenhong Li^a, Guillermo Ortí^b, Jinliang Zhao^{c,*}

^a School of Biological Sciences, University of Nebraska, Lincoln, NE 68588-0118, USA

^b Department of Biological Sciences, The George Washington University, Washington, DC 200052, USA

^c College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China

ARTICLE INFO

Article history:

Received 26 February 2010

Revised 11 May 2010

Accepted 17 May 2010

Available online 28 May 2010

Keywords:

Phylogeny

Siniperca

Sinipercidae

Percomorpha

Nuclear markers

Species-tree

ABSTRACT

Sinipercids are freshwater fishes from East Asia assigned to the genera *Siniperca* (8–9 species) and *Coreoperca* (3 species). Although supported as a monophyletic group, the phylogenetic placement of sinipercids among percomorph fishes and the identification of their closest relatives have perplexed generations of taxonomists. Here, we collect and analyze new molecular evidence from 11 nuclear genes – six of them novel markers developed for this study – sampled from seven sinipercid species and several closely-related species that have been proposed as their putative sister taxa in the past. The concatenated alignment of the 11 loci included 10,560 bp that were split into 18 blocks for a partitioned analysis using Bayesian and Maximum Likelihood methods. Both approaches resulted in nearly identical topologies, in which the monophyly of the sinipercids was strongly supported (Bootstrap Support = 99; Posterior Probability = 1.0) with Centrarchidae as their sister taxon (BS = 87; PP = 1.0). Percichthyidae formed a sister group to the sinipercid–centrarchid clade (BS = 100; PP = 1.0). Tests for alternative hypotheses of relationships significantly rejected the placement of sinipercids within the Serranidae, Moronidae, or Centropomidae, or as sister taxon to *Lateolabrax japonicus*, *Perca flavescens*, or *Pristigenys alta* (SH test, $p < 0.001$). The contribution of individual gene genealogies to support this result was explored further using a multispecies coalescent approach (BEST) to infer the species phylogeny. The sister-group relationship with Centrarchidae also is supported by this approach but relationships among species within the genus *Siniperca* showed a different pattern than the results obtained with the concatenated data set. Our study supports the designation of a distinct family, Sinipercidae, for this group of fishes (*Siniperca* and *Coreoperca*).

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Chinese perches (sinipercids) are freshwater fishes endemic to East Asia, mainly distributed in China (hence their name) with a few species also found in Korea, Japan, and Northern Vietnam (Li, 1991; Zhou et al., 1988). It is now generally accepted that they constitute a monophyletic group (Liu, 1997) and have been assigned to a distinct family, Sinipercidae (Roberts, 1993; Nelson, 2006). Two or three genera have been included in this group: *Siniperca* (8–9 species), *Coreoperca* (3 species; Liu and Chen, 1994), and sometimes *Coreosiniperca* (Fang and Chong, 1932; Zhou et al., 1988), before the latter was considered a junior synonym of *Siniperca* (Liu and Chen, 1994). In early classifications, sinipercids were placed with the Serranidae (Jordan, 1923; Nichols, 1943; Zhou et al., 1988). Gosline (1966), however, restricted the family Serranidae to the subfamilies Serraninae, Epinephelinae, and Anthiinae, and

he erected a new family Percichthyidae for sinipercids and other excluded taxa. Among these, Gosline (1966) assigned sinipercids to a group of estuarine and freshwater percichthyids, also including *Morone* (using the name *Roccus*), *Lateolabrax*, and *Percichthys*. Several years later, Johnson (1984) pointed out the lack of synapomorphic characters for Gosline's Percichthyidae and therefore restricted the Percichthyidae to freshwater genera of Australia and South America, leaving Moronidae, *Lateolabrax*, and the sinipercids as *incertae sedis*. Later morphological studies supported this view (Liu and Chen, 1994; Waldman, 1986) and sinipercids were assigned either to the Centropomidae (Waldman, 1986) or in a polytomy with the centropomids, *Lateolabrax*, and centrarchids (Liu, 1997). The likely relationship between centrarchids and “percichthyids” (including sinipercids), also was addressed by earlier morphological studies (Chang, 1988; McCully, 1962). More recent studies using molecular sequence data also produced conflicting results, recovering either a clade consisting of *Perca* and *Pristigenys* (Zhao et al., 2005), or a clade of percichthyids and centrarchids (Smith and Craig, 2007) as the sister taxon of sinipercids. There-

* Corresponding author. Fax: +86 21 61900435.

E-mail address: jlzhao@shou.edu.cn (J. Zhao).

fore, the phylogenetic placement of the siniperids remains uncertain in spite of their critical position to understanding higher-level interrelationships among percomorph fishes. A summary of these hypotheses is presented in Fig. 1 and Table 1.

In addition to higher classification, interrelationships among species of siniperids also are controversial. A recent study by Chen et al. (2007) suggested that siniperids are non-monophyletic based on an analysis of cytochrome *b* sequence data. The proposed relationships among species vary considerably among studies using morphological and molecular data, or even among different molecular phylogenetic studies (Fig. 1) (Chen et al., 2007, 2010; Liu, 1997; Shirai et al., 2003; Zhao et al., 2008, 2006a,b).

In the present contribution, we aim to assess these unsettled questions by implementing a strategy of extensive taxon and character sampling. We present analyses of new sequence data for 11 unlinked nuclear loci sampled from seven siniperids and representatives of all putative sister taxa that have been proposed so far. Specifically, our objectives are (1) to determine the phylogenetic placement of the siniperids within the percomorph fishes and identify their sister taxon and (2) to resolve the species phylogeny within the family. Our study will add important knowledge for understanding the interrelationships of the percomorph fishes, and also provide insights into historical biogeography of East Asia by recovering the species phylogeny of the siniperids.

2. Materials and methods

2.1. Taxon sampling

Specimens from seven siniperid species, including one *Coreoperca* and six *Siniperca* were sampled for this study. Putative sister taxa or close relatives to the siniperids that have been proposed in previous studies also were examined, including samples from Serranidae, Percichthyidae, Centropomidae, Moronidae, Centrarchidae, *Lateolabrax japonicus*, *Perca flavescens*, *Pristiglenys alta*, and *Howella brodiei*. A few other taxa were used as outgroups for a total of 25 samples (Table 2).

2.2. Phylogenetic markers

Eleven nuclear protein-coding genes were used as the phylogenetic markers. Five of them have been published in phylogenetic studies of ray-finned fishes (Li et al., 2007, 2008; Li and Ortí, 2007), including recombination activating gene 1 (*rag1*), *zic* family member 1 (*zic1*), cardiac muscle myosin heavy chain 6 alpha (*myh6*), pleiomorphic adenoma protein-like 2 (*plagl2*), and brain super conserved receptor 2 (*sreb2*) gene.

Table 1

Alternative hypotheses about the placement of the siniperids and their putative sister taxon proposed in different studies.

Placement of the siniperids/their sister taxon	References
Serranidae	Jordan (1923), Nichols (1943), and Zhou et al. (1988)
Percichthyidae	Gosline (1966), Liu and Chen (1994), and Nelson (1976, 1984)
<i>incertae sedis</i> genera	Johnson (1984)
Centropomidae	Nelson (1994) and Waldman (1986)
Siniperidae	Nelson (2006) and Roberts (1993)
Sister to <i>Lateolabrax</i> , centropomids, and centrarchids	Liu (1997)
Close to the percichthyids and centrarchids	McCully (1962) and Smith and Craig (2007)
Sister to <i>Perca</i> and <i>Pristiglenys</i>	Zhao et al. (2005)
Non-monophyly, <i>Coreoperca</i> is sister to <i>Centrarchus</i>	Chen et al. (2007)

The remaining six loci are new markers that we chose from the candidate markers proposed by Li et al. (2007) and were optimized for the present study. These are *si:dkey-174m14.3* (*sidkey*), *zgc:85947* (*SLC10A3*), *zgc:92173* (*UBE3A*), Zinc finger protein 503 (*znf503*), *ENSDARG00000059744* (*KIAA2013*), and *ENSOGLG00000015759* (*UBE3A*-like), which has similar motif as the *UBE3A*, so we use *UBE3A*-like to name this marker. All these genes have been chosen because they are single-copy genes in the model organisms zebrafish and torafugu (defined operationally by a bioinformatic filter) used to develop the markers (Li et al., 2007); therefore, a single orthologous copy is expected to be amplified by PCR in most species with locus-specific primers. Primers were designed to amplify a large exon for each gene (Table 3).

2.3. DNA extraction, polymerase chain reaction, and sequencing

DNA samples were extracted from ethanol-preserved muscle or gill tissues using the DNeasy tissue kit (Qiagen). Nested-PCR was used to amplify the target genes if the first-round of PCR failed to generate clean product. In these cases, the products of the first-round PCR were diluted 20–100 times and used as template for a second PCR with a set of primers nested within the product of the first PCR. All PCRs were performed in a total volume of 20 µl, including 0.1 µl TaKaRa Taq™, 2.0 µl 10× PCR buffer (+MgCl₂), 1.6 µl dNTP Mixture (2.5 mM each dNTP), 0.64 µl (10 µM) of each forward and reverse primer, 0.8 µl DNA template, and 14.22 µl distilled water. The PCRs were carried out in an Eppendorf Mastercycler with silver block. Touchdown programs were used for all PCRs. The program consisted of a 95 °C initial

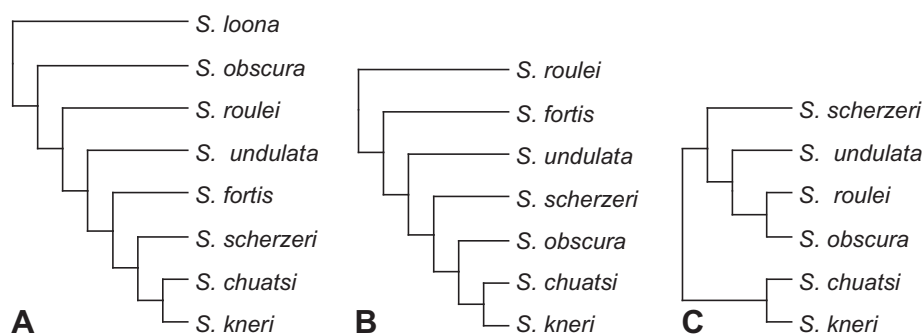


Fig. 1. Phylogenies of siniperid taxa proposed by different studies. (A) Cladogram based on 34 morphological characters (following Liu and Chen, 1994); (B) phylogeny obtained from codon-partitioned Bayesian analysis (link) of cytochrome *b* gene sequences (following Chen et al., 2007); (C) Bayesian phylogeny inferred from mitochondrial control region sequences (following Zhao et al., 2008).

Table 3

PCR primers and annealing temperatures used for the six new markers.

Marker ^a	Primer name	Primers Sequence	Temp.	PCR
sidkey	sidkey_F116	CGGATGARGYCTGCAGCAG	6260	1st
	sidkey_R1360	GCTGGGCYTTKGTGAGACTGT		
	sidkey_F247	GACCTSTACAGCAGYGACAC	6058	2nd
	sidkey_R1355	GCCTTKGTGAGACTGTCCTT		
SLC10A3	SLC10A3_F486	CCTCATYTTYGTCAACAAGTGTC	5856	1st
	SLC10A3_R1273	GCATCTCCRAVGTGCTGCT		
	SLC10A3_F520	AAGGTGGAGGTGGAGGTGCT	5856	2nd
	SLC10A3_R1256	CTSAGGGCCACGATGAAGG		
UBE3A	UBE3A_F354	GTCTACGACAGCCTATTGTCCAATGAGAARAT	6058	1st
	UBE3A_R1140	CGCTRTACATGCTGATCCKGTGT		
	UBE3A_F411	GTCTACTTCTCGCCCAACGTRGAGTG	5856	2nd
	UBE3A_R1081	TTGAGGATGAARGGRCAGGTCAT		
znf503	znf503_F12	GAAAAGTCCGYTGGCTCTKCT	6260	1st
	znf503_R1422	CGCCGAYGCTGTGGTSAGTCT		
	znf503_F40	ACCTGCTCTCARATCGGYAAA	6058	2nd
	znf503_R1415	GCTGTGGTCAGTCTCTGKCCRTA		
UBE3A-like	UBElke_F368	GCTGGCCATCGASCAGGTGGAG	6260	1st
	UBE3A_R1140	CGCTRTACATGCTGATCCKGTGT		
	UBElke_F423	CCGAAYATGGAGCTCGACCTGGA	6260	2nd
	UBElke_R1080	TGAGGATGAAGGGGAGSTCTG		
KIAA2013	KIAA2013_F41	CCAGYCGAACAGTSAACAACACCCT	5856	1st
	KIAA2013_R829	CGGGTCCRCAGTACTCRTTGTA		
	KIAA2013_F49	ACAGTSAACAACACCCTSTACTACAT	5856	2nd
	KIAA2013_R801	TTTGAAGAGGAASAARTGGAAGAG		

^a The markers are named with the ensembl locus tag name or the orthologous gene name of humans.

heating for 30 s, followed by 15 cycles of 98 °C for 10 s, the first touchdown annealing temperature for 30 s and 72 °C for 45 s, 15 cycles of 98 °C for 10 s, the second touchdown annealing temperature for 30 s and 72 °C for 45 s, followed by a final extension at 72 °C for 5 min. The touchdown annealing temperatures for the primer pairs are listed in Table 3. The PCR products were checked on agarose gels and amplified products were sequenced directly on an ABI 3730 platform.

2.4. Sequence alignment, data partitioning, and model selection

Sequences for each gene were edited using the software BioEdit (Hall, 1999). Because these nuclear loci are exons of protein-coding genes, the DNA sequences were translated into amino acid sequences to align each individual locus using ClustalW (Thompson et al., 1994) implemented in MEGA 4 (Tamura et al., 2007). The aligned protein sequences were then back-translated into nucleotide sequences for further analyses. For genes with regions difficult to align, the program, MAFFT (Katoh et al., 2009) was used and the results were compared with those from ClustalW. The base composition, number of variable sites, and average *p*-distance of each gene among all taxa were calculated using MEGA 4 (Tamura et al., 2007).

Data partitions were defined to account for heterogeneity in the concatenated data set and improve the model used for phylogenetic inference (Brandley et al., 2005; Brown and Lemmon, 2007; Caterino et al., 2001; Pupko et al., 2002). Li et al. (2008) proposed a strategy to partition concatenated data sets of protein-coding genes into groups of homogeneous blocks and showed that this method performs better than simply partitioning the data by gene and codon position. We applied this data partitioning strategy here to find the optimal number of partitions for this large concatenated data set.

Once the optimal partitioning scheme was found, the best model for each partition was identified by using the “propose model” routine in TreeFinder (Jobb, 2008). The resulting models were employed in the final partitioned ML analysis using TreeFinder. For

Bayesian analyses, the closest models to the TreeFinder models available in MrBayes 3.1.2 (Nylander et al., 2004) were used.

2.5. Phylogenetic analysis and topology test

Maximum likelihood tree searches were conducted with TreeFinder under the optimal partitioning scheme and the best models for each data partition. Bootstrap analysis with 1000 pseudo replicates was performed to assess the statistical support for each node in the resulting phylogeny. Bayesian analysis was performed using MrBayes 3.1.2. The same partitioning scheme and similar models were applied in Bayesian analysis. Two independent runs, with 20 Markov Chain Monte Carlo (MCMC) chains for each run were carried out for the Bayesian analysis. The heating parameter was set as “temp = 0.1” and “nswaps = 5” to improve the mixing of the MCMC chains. The runs were terminated after 10 million generations. Parameter states were sampled every 1000 generations (10,001 trees save for each run). Two approaches were taken to evaluate the convergence of the runs. First, the average standard deviation of split frequencies between two independent runs was monitored. A standard deviation below 0.01 was considered indicative of convergence of the two independent runs to the same stationary phase. Second, AWTY (Nylander et al., 2008) was used to evaluate the results from MCMC runs. Rather than tracking the likelihood of the model, AWTY focuses on the phylogenies themselves by comparing the posterior probability of each node among the tree samples. Following recommendations for AWTY, when independent runs reach stable and similar estimations for all of the nodes, we assumed that stationarity has been reached. The majority-rule 50% consensus tree was summarized from the sampled trees after discarding the burn-in samples preceding the stationary distribution.

To test alternative phylogenetic hypotheses, the preferred phylogeny obtained by this study was modified using Treeview (Page, 1996) to produce topological constraints consistent with alternative phylogenetic hypotheses (outlined in Table 1). The alternative topologies were evaluated a posteriori by one tailed Shimodaira–

Hasegawa (SH) tests (Shimodaira and Hasegawa, 1999) with 50,000 RELL bootstrap replicates as implemented in TreeFinder.

2.6. Signal from individual loci and the species-tree

In order to explore the contribution of each individual locus to the overall phylogenetic result, ML phylogenies for each locus also were reconstructed. The presence/absence of sinipercid monophyly and the sister taxon to the sinipercids were observed and compared among individual gene trees. The best ML tree inferred from the concatenated analysis (above) was used as a constraint for SH tests to estimate the level of consistency with this topology exhibited by each locus.

As an alternative to concatenation, a “species-tree” approach based on the multispecies coalescent also was applied to account for potentially confusing signal arising from lineage sorting events (Maddison, 1997; Kubatko and Degnan, 2007). Although lineages sorting may potentially affect any branching pattern obtained, our main concern was centered on the relationships among the sinipercid taxa, where short internal branches may be expected due to recent divergence times among closely-related species. Therefore, we applied the Bayesian estimation of species trees (BEST) under the coalescent model to jointly estimate gene trees and the species-tree from a multilocus sequence alignment (Liu, 2008). Due to computational limitations, to infer the species-tree of the sinipercids only a subset of the taxa was analyzed, including all sinipercids, *Micropterus*, *Percichthys*, and three outgroups (*Centropomus undecimalis*, *Dicentrarchus labrax*, and *Morone chrysops*). The likelihood model used was GTR + Gamma. Two independent runs, each with 20 chains were carried out. The temperature was set to 0.1 and number of swaps was set to 10 to improve the mixing of the chains. A MCMC chain with one 100 million generations was executed with a sampling frequency of 1 every 5000 (20,000 samples collected for each run). Convergence to stationarity of the runs was monitored as above and a number of samples before reaching convergence were discarded as “burn-in” to calculate the 50% majority-rule consensus species-tree.

3. Results

3.1. Phylogenetic markers, sequence variation, and data partitioning

Five nuclear loci were sequenced for all 25 taxa, including myh6, RAG1, UBE3A, zic1 and znf503. The other six loci had only a few (1–3) missing taxa (Table 2). All of the missing data were due to failed PCR amplifications. No multiple bands or noisy signal was detected in the resulting PCR products or sequences. However, the presence of paralogous fragments could be detected for the KIAA2013 locus in a study of ostariophysans (unpublished data); therefore we caution against using this marker for studying basal actinopterygian relationships. The use of the other new markers in a broader taxonomic range should also await more rigorous testing.

The aligned sequences for each of the 11 loci varied from 699 to 1527 bp. Recombination activating gene 1 and znf503 had the

Table 5

The optimal partitioning scheme. The ML model for each partition was selected by using “propose model” in TreeFinder (Jobb, 2008). The closest models to the ML models available in MrBayes were used for Bayesian analysis.

Partition	Data included	Model used in ML analysis	Model used in MrBayes
1	SLC10A3_1st, UBE3A_1st, sidkey_2nd	TN + G	GTR + G
2	sidkey_3rd, UBE3A-like_3rd, UBE3A_3rd, zic1_3rd, znf503_3rd, myh6_3rd, plagl2_3rd, rag1_3rd, sreb2_3rd	GTR + G	GTR + G
3	KIAA2013_3rd, SLC10A3_3rd	GTR + G	GTR + G
4	KIAA2013_2nd, UBE3A-like_2nd, rag1_2nd	TVM + G	GTR + G
5	rag1_1st, sidkey_1st, UBE3A-like_1st	TVM + G	GTR + G
6	KIAA2013_1st	TVM + G	GTR + G
7	myh6_1st	J3 + G	GTR + G
8	myh6_2nd	J2 + G	GTR + G
9	plagl2_1st	HKY + G	HKY + G
10	plagl2_2nd	J1 + G	GTR + G
11	SLC10A3_2nd	J1 + G	GTR + G
12	sreb2_1st	GTR + G	GTR + G
13	sreb2_2nd	HKY + G	HKY + G
14	UBE3A_2nd	HKY + G	HKY + G
15	zic1_1st	TVM + G	GTR + G
16	zic1_2nd	HKY + G	HKY + G
17	znf503_1st	J2 + G	GTR + G
18	znf503_2nd	TVM + G	GTR + G

highest number of variable sites and parsimony informative sites, partly because they were the longest fragments sequenced (Table 4). The average *p*-distance among taxa was close to 10% for most loci, with UBE3A-like and sidkey exhibiting the largest values, while zic1 had the smallest *p*-distance, indicating a slow evolutionary rate of this marker, which is consistent with previous results (Li et al., 2007, 2008). The base composition (GC%) also was variable among different loci, ranging from 51.5% in myh6 to 64.6% in SLC10A3 (Table 4).

For the concatenated analysis a model based on 18 partitions was chosen as the optimal partitioning scheme using the method of Li et al. (2008). The third codon positions of most loci were grouped together as one block in this partitioning scheme. The partition scheme and best model selected for each partition are listed in Table 5.

3.2. Phylogeny from the partitioned analysis

Concatenating the 11 loci produced a total length of 10,560 bp. Partitioned Bayesian analyses and the ML method using the concatenated data resulted in nearly the same phylogeny (Fig. 2). The only difference was the placement of *Pristigynys alta* and *Sciaenops ocellatus*, so their relationships are shown as a polytomy in Fig. 2, especially because neither method produced a well supported resolution of their relationships. The sinipercids were recovered as a monophyletic group with two genera, *Siniperca* and *Coreoperca*. The support for the monophyly of the sinipercids was very high, 99% for the bootstrap support (BS) and 1.0 for the Bayesian posterior probability (PP). *Siniperca roulei* was

Table 4

The characteristics of the 11 loci sequenced for the 25 taxa in this study.

Characteristics	KIAA-2013	myh6	plagl2	RAG1	sidkey	SLC10-A3	sreb2	UBE3A	UBE3A-like	zic1	znf503
Number of bp	762	771	915	1527	1176	765	1008	699	699	879	1359
Number of variable sites	277	296	331	698	530	299	272	253	326	223	451
Number of parsimony informative sites	162	209	181	453	312	194	158	172	214	121	274
Average genetic <i>p</i> -distance (%)	9.1	10.2	7.9	10.5	12.2	10.7	6.3	8.8	13	5	8.4
GC%	61.9	51.5	57.7	52.8	62.2	64.6	58	54.4	51.9	58.5	58.7

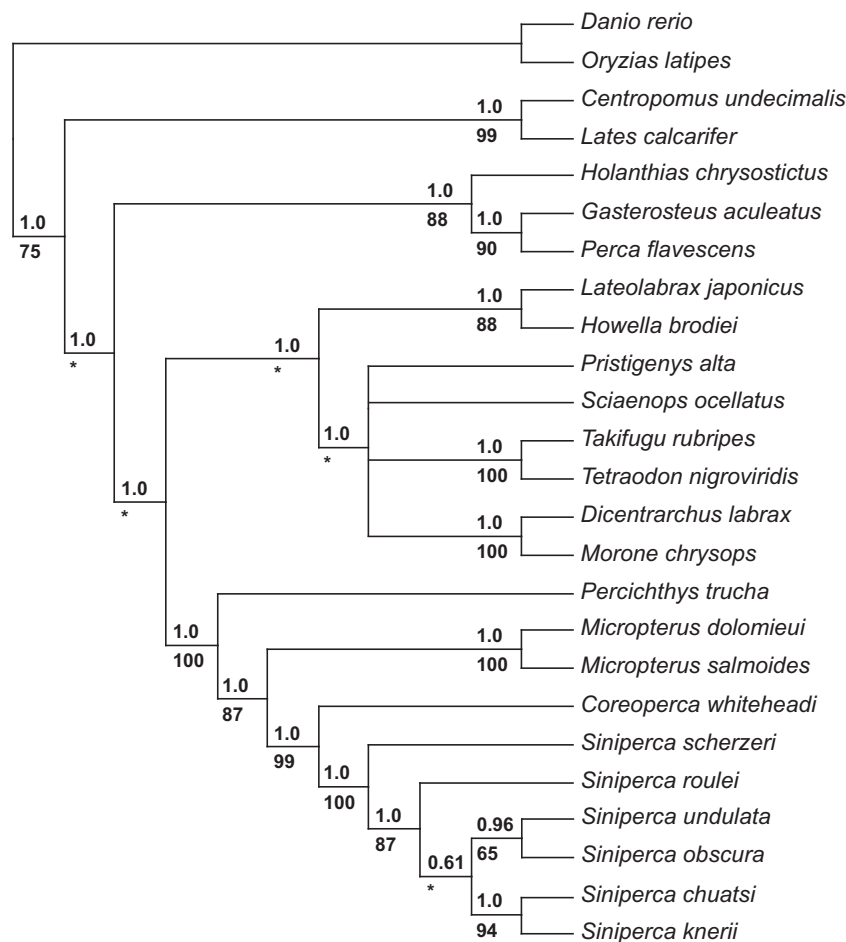


Fig. 2. Consensus phylogeny of the siniperids and closely-related taxa inferred from partitioned Bayesian and ML analysis of 11 nuclear loci. The 11 genes were partitioned into 18 data blocks as shown in Table 5. The numbers above the branches are Bayesian posterior probabilities, whereas the numbers beneath the branches are bootstrap support values from maximum likelihood analysis (ML). The asterisks indicate bootstrap values less than 50%.

Table 6
Shimodaira–Hasegawa tests (Shimodaira and Hasegawa, 1999) on the alternative hypotheses proposed for the phylogenetic placement of the siniperids.

Proposed placement / sister taxon	–ln L	Δ –ln L	p-Value
Serranidae	55,390	259	<0.001
Percichthyidae	55,142	11	1
Moronidae	55,404	273	<0.001
Centropomidae	55,365	234	<0.001
Sister to <i>Lateolabrax japonicus</i>	55,368	237	<0.001
Sister to <i>Perca flavescens</i>	55,495	364	<0.001
Sister to <i>Pristigenys alta</i>	55,396	265	<0.001

nested within *Siniperca*, so this result does not support the recognition of the genus *Coreosiniperca* proposed by previous

authors (Zhou et al., 1988; Zhu, 1985). Within the genus *Siniperca*, the basal node subtends the *S. scherzeri* lineage, followed by branching of *S. roulei*. *Siniperca undulata* and *S. obscura* were grouped together, whereas *S. chuatsi* and *S. kneri* were resolved as the sister taxa; and these four species formed a clade, but with very low support (Fig. 2).

When we rooted the phylogeny with *Danio rerio* and *Oryzias latipes*, the closest relatives to the siniperids were among the Centrarchidae (the two *Micropterus* species), strongly supported by our data (BS = 87; PP = 1.0). The next taxon closer to the centrarchid–siniperid clade was the Percichthyidae (*Percichthys trucha*), again with a high BS value of 100% and PP of 1.0 (Fig. 2). All of the alternative phylogenetic placements of siniperids with respect to the putative sister taxa

Table 7
The phylogenetic signal of individual loci and the patterns supported by them.

Signal or patterns supported by each locus	KIAA-2013	myh6	plagl2	RAG1	sidkey	SLC10-A3	sreb2	UBE3A	UBE-like	zic1	znf503
p-Value of SH-test ^a	0.03	0.21	1	0.13	0.03	0.22	0.007	0.16	1	0.028	0.024
Monophyletic siniperids?	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	Yes
S–C–P clade supported? ^b	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Sister group to the siniperids ^c	C	P	C	C	C	C	PC	P	C	PC	C

^a SH-test calculated the p-value of observing the data with the topology constrained using the tree supported by all genes (Fig. 2). Missing taxa in the particular gene were pruned from the constraining tree.

^b Is the clade consisting the siniperids, centrarchids and percichthyids supported?

^c Who is the sister group to the siniperids, the centrarchids (C) or percichthyids (P)? PC indicates a polytomy among the siniperids, percichthyids and centrarchids.

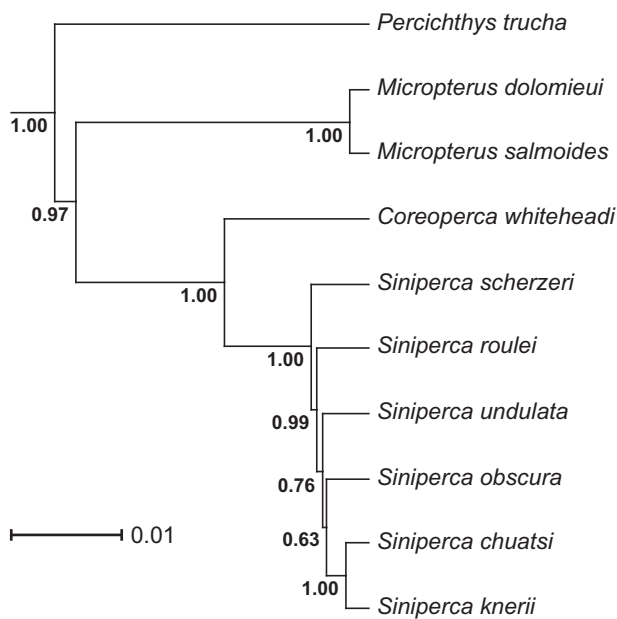


Fig. 3. The species-tree of the sinipercids and their sister taxa constructed from 11 nuclear loci using the BEST method (Liu and Pearl, 2007). The tree was rooted with *Centropomus undecimalis*, which is omitted in this figure in order to show the details within *Siniperca*. The values beneath the branches are posterior probabilities.

proposed in the literature were significantly rejected by the SH test (Table 6).

3.3. Signal from individual loci and the resulting species-tree

To assess the degree of potentially consistent and conflicting phylogenetic signal among loci, we performed SH tests on each locus using the topology obtained with the concatenated data set (Fig. 2) as the constraining topology. Six loci could not reject the phylogeny based on concatenated analysis, while the remaining five loci showed significant discrepancy with it (Table 7). We also reconstructed phylogeny based on individual genes (figures not shown). The monophyly of the sinipercids was supported by seven of the 11 loci, and the clade grouping the sinipercids, centrarchids, and percichthyids (S–C–P) was supported by eight of the 11 loci. The sister group to the sinipercids was found to be either the centrarchids or percichthyids in all individual gene trees (Table 7). The non-monophyly of the sinipercids seen in the individual gene trees was caused mostly by the low resolution or uncertainty in the relationships among taxa in the clade consisting the sinipercids, centrarchids, and percichthyids (S–C–P).

Because of the uncertainty and short internal branches found in the S–C–P clade and among the *Siniperca* species in the individual gene trees, we suspect that the cause of gene-tree discordance might be incomplete lineage sorting. We applied the species-tree method BEST to infer the species phylogeny (Liu and Pearl, 2007). The resulting species phylogeny also produced high support for the S–C–P clade (PP = 1.0). The sister taxon of the sinipercids using the species-tree method again was found among the centrarchid taxa used in the study (Fig. 3). Within the genus *Siniperca*, the same branching pattern among *S. scherzeri* and *S. roulei* was found as in the results of the concatenated analyses (Fig. 2). However, the sister-group relationship among *S. undulata* and *S. obscura* observed in the concatenated analyses was not obtained, and the species-tree revealed a different branching order of *S. undulata* and then *S. obscura*. *Siniperca chuatsi* and *S. kneri* still were highly supported as the sister taxa (Fig. 3).

4. Discussion

4.1. Monophyly of the sinipercids

Coreoperca and *Siniperca* are closely-related taxa that often have been listed under the same family or subfamily (Jordan, 1923; Nichols, 1943; Roberts, 1993; Zhou et al., 1988; Zhu, 1985). The monophyly of sinipercids was confirmed by morphological studies, and several synapomorphy characters were proposed, e.g., the star-shaped transverse section of the first anal pterygiophore (Liu, 1997; Liu and Chen, 1994) and some molecular studies (Zhao et al., 2006a,b). However, Chen et al. (2007) argued that the sinipercids are non-monophyletic based on cytochrome *b* sequence data. They reported that nine species representing Moronidae, Serranidae (*Nippon* in their study), Centropomidae, Acropomatidae, Emmelichthyidae, Siganidae, and Centrarchidae were nested between *Coreoperca* and *Siniperca*, but they also were cautious in interpreting the results because different analyses resulted in different tree and most of the topological tests could not reject the monophyly of the sinipercids (Chen et al., 2007). Our results provide strong support for the monophyly of the sinipercids, with a BS value of 99% and PP of 1.0. The species-tree method also resulted in high confidence for this clade (PP = 1.0). When we looked at the different loci individually, seven of the 11 loci supported the monophyly of the sinipercids while the remaining loci had a less resolved S–C–P clade (Table 7). We reanalyzed the data of Chen et al. (2007) but with minimum evolution or maximum parsimony methods, and we also obtained a S–C–P clade (results not shown), which is similar to some of our individual gene trees. Therefore, the results of Chen et al. (2007) as evidence against the monophyly of the sinipercids should be re-interpreted as a consequence of insufficient signal in a single marker (Cyt *b*) to resolve this question. This same situation also was observed in some of the individual genes used in this study when the markers when analyzed separately. Inference based on multiple loci (either concatenated or under the multispecies coalescent) provides a more robust result.

4.2. Placement of the sinipercids and recognition of the family, Sinipercidae

Previous classifications placed the sinipercids within the families Serranidae, Percichthyidae, or as sister taxon to the centropomids, centrarchids, and *Lateolabrax* (Table 1). Two studies in particular examined in detail morphological characters of the sinipercids in relation to other “lower percoids”. Waldman (1986) suggested that the sinipercids comprise an unrecognized subfamily within centropomids, probably a sister taxon of *Lates*. His conclusions were based on shared traits between the centropomids and *Siniperca*, such as characters associated with the antero-posterior expansion of the second neural, unique ornamentation of the preopercle, enclosure by bone of the preopercular portion of the preoperculomandibular canal, and the derived form of the symplectic (Waldman, 1986). In the other morphological study, Liu (1997) performed a cladistic analysis using eight characters, and found that the sinipercids formed a polytomy with the centropomids, centrarchids, and *Lateolabrax*.

Our results highly supported a sister-group relationship between the sinipercids and centrarchids (BS = 87%; PP = 1.0), and this clade then formed a sister group to the percichthyids (BS = 100%; PP = 1.0). The analysis using a species-tree approach also resulted in the same relationship with high support (Fig. 3), indicating that this pattern was not affected by gene-tree-species-tree discordances. A close relationship between the sinipercids and centrarchids has not been proposed formally in the past (but see Liu, 1997). However, the affinity between the percichthy-

ids (containing siniperchids) and centrarchids was suggested before (McCully, 1962; Smith and Craig, 2007). Perhaps an exhaustive reexamination of the morphological characters in the siniperchids, centrarchids and percichthyids would produce evidence to reconcile the results of the present study and the previous morphological studies. Nevertheless, our data significantly rejected the close relationship between the siniperchids and centropomids, or the close relationship between the siniperchids and *Lateolabrax*. Other previous proposed hypotheses also were significantly rejected (Table 6).

It has also been proposed that *Coreoperca* and *Siniperca* should be recognized as a distinct family, Siniperchidae (Nelson, 2006; Roberts, 1993). Our results showed that the distance between the siniperchids and percichthyids is equal if not larger than the distance between the centrarchids and percichthyids (Fig. 3), thus, we agree with Nelson (2006) and Roberts (1993) in erecting the family Siniperchidae for this group of fishes.

4.3. Interrelationships among the *Siniperca* species

Liu and Chen (1994) examined 34 morphological characters for 12 siniperchid species (three *Coreoperca* and nine *Siniperca*). They used *Lateolabrax* as an outgroup, and coded the characters of *Lateolabrax* as plesiomorphic to polarize character states in the siniperchids. Based on a cladistic study of their character matrix, they published the first cladogram of the siniperchids (Fig. 1A). However, their results were not supported by recent molecular studies using mitochondrial loci (Fig. 1B and C), except for the sister-group relationship between *S. chuatsi* and *S. kneri*. Furthermore, the two molecular phylogenetic studies were not congruent with each other either (Chen et al., 2007; Zhao et al., 2008). The inference from the morphological study might have been different if the character states were polarized using a different outgroup. The morphological characters of *Siniperca* could be reevaluated using the centrarchids as outgroup instead of using *Lateolabrax*. The inconsistency between the results of the previous studies using a single mitochondrial locus and the results of the present study probably was due to the gene-tree-species-tree conflicts or insufficient signal in single locus.

Our results using the species-tree method (Fig. 3) showed a few differences from the results of the concatenated analyses (Fig. 2) in the relationships among the *Siniperca* species. In the species-tree, *S. undulata*, *S. obscura*, and *S. chuatsi* and *S. kneri* formed a sequential branching pattern (Fig. 3), whereas the concatenated methods produced a more symmetrical relationships (Fig. 2), which can be predicted by the coalescence theory, i.e., the symmetrical pattern is often preferred by the concatenated methods even the true species-tree is non-symmetrical (Kubatko and Degnan, 2007).

In a very recent study on the species phylogeny of siniperchids using three mitochondrial and two nuclear loci, Chen et al. (2010) reported that using different genes or different methods led to inconsistent results. The conflicting signal was largely due to differences between the mitochondrial genes and the nuclear loci. Nevertheless, their combined analysis resulted in the same species phylogeny obtained here (Fig. 3), but with one exception. In the phylogeny reported by Chen et al. (2010), *S. undulata* was shown as the basal group in related to the other *Siniperca* species, while *S. scherzeri* was basal in our results. In our separate analyses of 11 individual genes, *S. undulata* was found as the basal *Siniperca* species only once, while *S. scherzeri* was found five times as the basal group. More importantly, our species-tree analysis also confirmed the basal position of *S. scherzeri*, which should eliminated the situation that signal from a single gene dominated the phylogenetic inference.

5. Conclusions

The monophyly of siniperchids was strongly supported by our data. The sister taxon of the siniperchids was found among the taxa included in the present study to be the Centrarchidae, with the Percichthyidae forming a sister group to the siniperchids-centrarchids clade. Analyzing multiple loci using a species-tree method produced a different species phylogeny for *Siniperca* compared to results obtained using the concatenated method. In the species-tree, *S. scherzeri* was found to be the basal lineage, then *S. roulei*, *S. undulata*, and *S. obscura* branched off in a subsequent order, leaving *S. chuatsi* and *S. kneri* as the sister group at the top. Our results support the recognition of a single family, Siniperchidae, to include fishes in the genera *Coreoperca* and *Siniperca*.

Acknowledgments

Most of this work was supported by NSF Assembling the Tree of Life Grant (DEB-0732838 to C. Li and G. Ortí). J. Zhao was supported by Shanghai Leading Academic Discipline Project (No.Y1101), Grants from Key Academic Discipline of Aquaculture of Shanghai Ocean University (No. 04SC11) and Shanghai Education Committee (No. 05KZ02). Some of the fish samples were provided by the Natural History Museum & Biodiversity Research Center, University of Kansas.

References

- Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Syst. Biol.* 54, 373–390.
- Brown, J.M., Lemmon, A.R., 2007. The importance of data partitioning and the utility of bayes factors in Bayesian phylogenetics. *Syst. Biol.* 56, 643–655.
- Caterino, M.S., Reed, R.D., Kuo, M.M., Sperling, F.A., 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Syst. Biol.* 50, 106–127.
- Chang, C.-H.M., 1988. Systematics of the Centrarchidae (Perciformes: Percioidei) with Notes on the Haemal–Anal–Axial Character Complex. Department of Biology, The City University of New York, New York.
- Chen, D., Guo, X., Nie, P., 2007. Non-monophyly of fish in the Siniperchidae (Perciformes) as inferred from cytochrome *b* gene. *Hydrobiologia* 77, 89.
- Chen, D., Guo, X., Nie, P., 2010. Phylogenetic studies of siniperchid fish (Perciformes: Siniperchidae) based on multiple genes, with first application of an immune-related gene, the virus-induced protein (viperin) gene. *Mol. Phylogenet. Evol.* 55, 1167–1176.
- Fang, P.W., Chong, L.T., 1932. Study on the fishes referring to *Siniperca* of China. *Sinensia* 2 (12), 137–200.
- Gosline, W.A., 1966. The limits of the fish family Serranidae, with notes on other lower percoids. *Proc. Calif. Acad. Sci.* 33, 91–111.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp.* Ser. 41, 95–98.
- Jobb, G., 2008. TREEFINDER. Distributed by the Author. Munich, Germany. Available from: <www.treefinder.de> (accessed 01.05.10).
- Johnson, G.D., 1984. Percioidei: development and relationships. In: Moser, H.G., Richards, W.J., Cohen, D.M., Fahay, M.P., Kendall, A.W., Jr., Richardson, S.L. (Eds.), *Ontogeny and Systematics of Fishes* (based on an international symposium dedicated to the memory of Elbert Halvor Ahlstrom). Special Publication Number 1, American Society of Ichthyologists and Herpetologists. Allen Press, Lawrence, Kansas, pp. 464–498.
- Jordan, D.S., 1923. A classification of fishes including families and genera as far as known. *Stanford Univ. Publ., Univ. Ser., Biol. Sci.* 3, 192.
- Katoh, K., Asimenos, G., Toh, H., 2009. Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537, 39–64.
- Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56, 17–24.
- Li, S., 1991. Geographical distribution of the Siniperchidae fishes. *Chin. J. Zool.* 26.
- Li, C., Ortí, G., 2007. Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 44, 386–398.
- Li, C., Ortí, G., Zhang, G., Lu, G., 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol. Biol.* 7, 44.
- Li, C., Lu, G., Ortí, G., 2008. Optimal data partitioning and a test case for ray-finned fishes (Actinopterygii) based on ten nuclear loci. *Syst. Biol.* 57, 519–539.
- Liu, H., 1997. Study on systematic position of siniperchid fishes with discussion on relationships of some lower perciformes. *Trans. Chin. Ichthyol. Soc.* 6, 1–7.

- Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24, 2542–2543.
- Liu, H., Chen, Y., 1994. Phylogeny of the sinipercine fishes with some taxonomic notes. *Zool. Res.* 15, 1–12.
- Liu, L., Pearl, D.K., 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst. Biol.* 56, 504–514.
- Maddison, W.P., 1997. Gene trees in species trees. *Syst. Biol.* 46, 523–536.
- McCully, H.H., 1962. The relationship of the Percidae and the Centrarchidae to the Serranidae as shown by the anatomy of their scales. *Am. Zool.* 2, 430.
- Nelson, J.S., 1976. *Fishes of the world*. Wiley-Interscience, New York.
- Nelson, J.S., 1984. *Fishes of the world*. John Wiley and Sons, New York.
- Nelson, J.S., 1994. *Fishes of the world*. John Wiley and Sons, Inc., New York.
- Nelson, J.S., 2006. *Fishes of the World*. John Wiley and Sons, Inc., New York.
- Nichols, J.T., 1943. *The Fresh-Water Fishes of China*. American Museum of Natural History, New York.
- Nylander, J.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53, 47–67.
- Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24, 581–583.
- Page, R.D., 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12, 357–358.
- Pupko, T., Huchon, D., Cao, Y., Okada, N., Hasegawa, M., 2002. Combining multiple data sets in a likelihood analysis: which models are the best? *Mol. Biol. Evol.* 19, 2294–2307.
- Roberts, C.D., 1993. Comparative morphology of spined scales and their phylogenetic significance in the Teleostei. *Bull. Mar. Sci.* 52, 60–113.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Shirai, S.M., Yabumoto, Y., Kim, I., Zhang, C.G., 2003. Phylogeny of sinipercid fishes and their relatives inferred from mtDNA cytochrome *b* gene: a preliminary study. *Bull. Kitakyushu Mus. Nat. Hist. Hum. Hist. Ser. A* 1, 45–49.
- Smith, W.L., Craig, M.T., 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. *Copeia* 2007, 35–55.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596–1599.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Waldman, J.R., 1986. *Systematics of Morone (Pisces: Moronidae)*, with Notes on the Lower Percoids. Department of Biology, The City University of New York, New York.
- Zhao, J., Li, S., Cai, W., Wang, W., 2005. The preliminary phylogenetic relationships of sinipercine fishes and some lower percoids inferred from 16S ribosomal DNA sequences. *J. Shanghai Fish. Univ.* 14, 364–369.
- Zhao, J., Li, S., Cai, W., Wang, W., 2006a. Phylogenetic relationship of sinipercine fishes in East Asia based on cytochrome *b* sequences analysis. *Acta Zool. Sin.* 52, 676–680.
- Zhao, J., Wang, W., Li, S., Cai, W., 2006b. Structure of the mitochondrial DNA control region of the sinipercine fishes and their phylogenetic relationship. *Acta Genet. Sin.* 33, 793–799.
- Zhao, J., Li, C., Zhao, L., Wang, W., Cao, Y., 2008. Mitochondrial diversity and phylogeography of the Chinese perch, *Siniperca chuatsi* (Perciformes: Sinipercidae). *Mol. Phylogenet. Evol.* 49, 399–404.
- Zhou, C., Yang, Q., Cai, D., 1988. On the classification and distribution of the Sinipercinae fishes (Family Serranidae). *Zool. Res.* 9, 113–125.
- Zhu, Y., 1985. *The Fishes of Fujian Province (Part II)*. Fujian Science and Technology Press, Fuzhou, China.