Molecular Phylogenetics and Evolution 55 (2010) 1167–1176

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenetic studies of sinipercid fish (Perciformes: Sinipercidae) based on multiple genes, with first application of an immune-related gene, the virus-induced protein (viperin) gene

Dali Chen [a,b,*,1], Xianguang Guo [c,1], Pin Nie [b]

[a] Department of Parasitology, School of Preclinical and Forensic Medicine, Sichuan University, Chengdu, Sichuan Province 610041, China
[b] State Key Laboratory of Freshwater Ecology and Biotechnology, Laboratory of Fish Diseases, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei Province 430072, China
[c] Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan Province 610041, China

ARTICLE INFO

Article history:
Received 13 November 2009
Revised 25 January 2010
Accepted 30 January 2010
Available online 4 February 2010

Keywords:
Phylogeny
Sinipercidae
Siniperca
Coreoperca
Viperin gene
S7 ribosomal protein gene

ABSTRACT

The sinipercid fish represent a group of 12 species of freshwater percoid fish endemic to East Asia. To date published morphological and molecular phylogenetics hypotheses of sinipercid fish are part congruent, and there are some areas of significant disagreement with respect to species relationships. The present study used separate and combined methods to analyze 7307 bp of data from three mitochondrial genes (cyt *b*, CO1 and 16S rRNA; ≈2312 bp) and three nuclear genes (viperin, the first two introns of *S7* ribosomal protein gene; ≈4995 bp) for the attempts to estimate the relationships among sinipercids and to assess the phylogenetic utility of these markers. Phylogenetic trees were reconstructed using maximum parsimony, maximum likelihood and partitioned Bayesian analyses. Despite the detection of significant heterogeneity of phylogenetic signal between the mitochondrial and nuclear partitions, the combined data analysis represented the best-supported topology of all data. The sinipercid fish form a monophyletic group with two distinct clades, one corresponding to the genus *Siniperca* and the other to *Coreoperca*. *Coreoperca whiteheadi* is the sister taxon to *Coreoperca herzi* plus *Coreoperca kawamebari*. In the *Siniperca*, *Siniperca undulata* is the sister taxon to the other members of *Siniperca*, within the subclade containing the other members of the genus, *Siniperca chuatsi* and *Siniperca kneri* are sister species, next joined by *Siniperca obscura*, *Siniperca roulei*, *Siniperca scherzeri* and finally by *Siniperca fortis*. The potential utilities of six different genes for phylogenetic resolution of closely related sinipercid species were also evaluated, with special interest in that of the novel virus-induced protein (viperin) gene.

1. Introduction

The sinipercids, a group of freshwater perciform fish, are endemic in China, Vietnam, Japan and Korea, with the majority of species recorded in China (Zhou et al., 1988). A total of twelve species have been described in this group, and disputes over their taxonomy and phylogeny have ever been solved yet. These fish were once assigned to only one genus, *Siniperca* (e.g., Zheng, 1989), or to two genera *Siniperca*, *Coreoperca* (e.g., Liu and Chen, 1994; Nelson, 2006), or even three genera, *Siniperca*, *Coreoperca*, *Coreosiniperca* (e.g., Kong and Zhou, 1993). Moreover, the family to which these fish belong is also a dispute (e.g., Johnson, 1984; Waldman, 1986), although Nelson (2006) considered that these fish can be allocated into Centropomidae, Percichthyidae, or, as is more appropriate and done by Roberts (1993), in their own family Sinipercidae.

Despite the taxonomical disputes, it is a bit surprising that relatively few studies have been carried out in relation to the phylogeny of sinipercids. Using morphological characters and allozymes, Kong and Zhou (1992) considered that sinipercids are of monophyletic origin, confirming the identification of two genera, *Siniperca* and *Coreoperca*. On the basis of morphology, other authors (e.g., Liu and Chen, 1994; Roberts, 1993; Yabumoto and Uyeno, 2000) also suggested that sinipercids should be clarified into the two genera. Much recently, Shirai et al. (2003), using mitochondrial cyt *b* gene sequences, analyzed the phylogenetic relationship of five species of sinipercids, and concluded that these fish were monophyletic with the validity of the two genera. Recently, Zhao et al. (2005, 2006a,b) showed the same conclusion, by using sequences of 16S rRNA, cyt *b* and of mitochondrial control region, but they provided some additional evidence on *Coreosiniperca*, which should be merged into *Siniperca*. It is rather obvious that all these authors, while studying the phylogeny considered that

* Corresponding author. Address: Department of Parasitology, School of Preclinical and Forensic Medicine, Sichuan University, Chengdu, Sichuan Province 610041, China. Fax: +86 28 85503204.
E-mail address: cdl1978119@sina.com (D. Chen).
[1] These authors contributed equally to this work.

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.01.039

sinipercid fish should be clarified into an independent family, Sinipercidae. However, also using cyt *b* but incorporated with many other cyt *b* sequences from a wide range of perciform fish, Chen et al. (2007) found the non-monophyly of these sinipercids, with obviously no comments on their family status. In addition, Chen et al. (2007) also demonstrated that this group does not seem to have a very clear phylogenetic history, for different methods of phylogenetic reconstruction result in different tree topologies.

No nuclear DNA sequences, however, have been employed to test the phylogenetic relationship of sinipercid fish. Single- or low-copy nuclear genes which may represent a source of multiple, unlinked and independently evolving loci, may provide ideal dataset for molecular phylogenetic inference (Cronn et al., 2003; Rokas et al., 2003); and introns which may escape natural selection are considered to have the potential to yield phylogeny-informative sites (Pons et al., 2004; Slade et al., 1994). In the present study, the virus-induced protein (viperin) gene from the nuclear genome was firstly used for phylogenetic resolution among sinipercid fish. To compare with other molecular markers, the first two introns of nuclear *S7* ribosomal protein gene and two genes (16S rRNA and CO1) from mitochondrial genome were also determined from these fish samples. The cyt *b* sequences reported in a previous study (Chen et al., 2007) were included in the present analyses. With 6 molecular markers, including three from nuclear DNA: viperin, the first two introns of S7, and three mitochondrial genes: 16S rRNA, CO1 and cyt *b*, we aimed to examine the potential of viperin gene as a novel marker for phylogenetic analysis of sinipercid fish, and to provide a broader understanding of interspecific relationships among these fish.

2. Materials and methods

2.1. Collection of samples

The specimens used in this study, including 10 species of sinipercids and 2 species of non-sinipercids (*Lateolabrax maculatus* and *Niphon spinosus*), were collected from a variety of localities in East Asia (Table 1). Due to the sampling difficulties, *Siniperca loona* and *S. robusta* were not included in this study. Muscle tissues of each species were collected and preserved in 95% ethanol, and most specimens collected were deposited in the State Key Laboratory of Freshwater Ecology and Biotechnology, in the Institute of Hydrobiology, Chinese Academy of Sciences. *L. maculatus* and *N. spinosus* were chosen as outgroups, which belong to Moronidae and Serranidae, respectively, and are considered as close relatives of sinipercids according to Liu (1997) and Shirai et al. (2003).

2.2. DNA extraction, amplification, cloning and sequencing protocols

The total genomic DNA was extracted from ethanol-preserved muscle following the method of Sambrook et al. (1989). The CO1 gene (≈666 bp) was amplified with the primers FishF1 and FishR1 as described by Ward et al. (2005). One cyt *b* gene sequence in *Coreoperca herzi* was amplified with the primers L14724 and H15915 as reported by Xiao et al. (2001), and other cyt *b* genes were reported in a previous paper (Chen et al., 2007). To sequence the first two introns of the *S7*, the primers used by Chow and Hazama (1998) were employed in this study. The 16S rRNA gene (≈580 bp) was obtained with the universal primers 16Sar-L and 16Sbr-H (Palumbi, 1996). The PCR cycling conditions were 94 °C for 3 min followed by 30 cycles of 94 °C for 30 s, 55–63 °C (adjusted according to the quality of template DNA) for 30 s and 72 °C for 1 min, and then a final elongation step at 72 °C for 10 min. The amplified DNA fragments were purified via spin columns and sequenced with an ABI 3730 automated DNA sequencer following the manufacturer's protocol.

Viperin gene sequences were constructed by linking together overlapping segments amplified with three primers designed on the basis of Sun and Nie (2004). The primer sequences and the positions in the complete viperin gene of *Siniperca chuatsi* are listed in Table 2 and Fig. 1, respectively. The PCR cycling conditions for primers 1(G1f and G1r) and 2 (G2f and G2r) were 94 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 63 °C for 1 min and 72 °C for 1 min 30 s, and then a final elongation step at 72 °C for 10 min. The PCR cycling conditions for primers 3 (G3f and G3r) were 94 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 54 °C for 1 min and 72 °C for 1 min, and then a final elongation step at 72 °C for 10 min.

Each viperin segment amplified as a single fragment in all taxa, and an initial screening for the presence of heterozygous sites was done by cloning PCR products into pMD18-T vector (TAKARA). The first segment of viperin (amplified with G1f and G1r, ≈1039 bp) was sequenced with the universal primers M13 for five clones each of 4 genera (*Siniperca*, *Coreoperca*, *Niphon* and *Lateolabrax*). No evidence of heterozygosity was found, and all subsequent sequencing was performed with one clone of a segment in each taxon. Sequences were then determined in both directions for each species and submitted for BLAST searching (Altschul et al., 1997) in GenBank to verify the data, with the accession numbers listed in Table 1.

2.3. Sequence alignment and analyses

Considering the first three bases on the 5′ end and the last 188 bases on the 3′ end of the viperin gene were excluded in the PCR

Table 1
List of species, origin, and database accession numbers.

Family	Species	Origin	Viperin	*S7* intron 1	*S7* intron 2	Cyt *b*	CO1	16S rRNA
Sinipercidae	*Siniperca chuatsi*	Changde, China	AY395718[a]	DQ864745	EF143358	DQ274042	EF143387	AY898947[b]
	S. kneri	Changde, China	EF143401	DQ864749	EF143365	DQ274047	EF143389	AY898948[b]
	S. fortis	Shiliting, China	EF143400	DQ864747	EF143361	DQ274045	EF143388	EF143378
	S. obscura	Cili, China	EF143402	DQ864750	EF143364	DQ274048	EF143390	DQ345335[b]
	S. scherzeri 1	Changde, China	EF143403	DQ864751	EF143366	DQ274043	EF143391	AY898949[b]
	S. scherzeri 2	Changju, South Korea	EF143404	DQ864752	EF143367	DQ274044	EF143392	EF143380
	S. undulata	Lianzhou, China	EF143405	DQ864753	EF143363	DQ274046	EF143393	DQ345334[b]
	S. roulei	Changde, China	EF143394	DQ864736	EF143360	DQ274049	EF143381	DQ345333[b]
	Coreoperca herzi	South Korea	EF143395	DQ864737	EF143370	DQ864731	EF143382	EF143374
	C. kawamebari	Niigata, Japan	EF143396	DQ864738	EF143371	DQ274051	EF143383	EF143375
	C. whiteheadi	Changde, China	EF143397	DQ864740	EF143368	DQ274050	EF143384	AY898950[b]
Serranidae	*Niphon spinosus*	Fukuoka, Japan	EF143399	DQ864743	EF143355	DQ274053	EF143386	EF143377
Moronidae	*Lateolabrax maculatus*	Guangzhou, China	EF143398	DQ864742	EF143357	DQ274054	EF143385	EF143376

[a] Source of additional sequences: Sun and Nie (2004).
[b] Source of additional sequences: Zhao et al. (2006a).

Table 2
The primers used to amplify the viperin gene in this study.

Primer	Primer sequence	Approximate product length (bp)	References
G1f	5′-GCG AAC AGA GCA GGT GTG ATT C-3′	1039	This study
G1r	5′-AAG GTG TTG ATG ACC GAG TTG ATC-3′		This study
G2f	5′-TCA ACT CGG TCA TCA ACA CCT TCA-3′	1720	This study
G2r	5′-TCC ACA CAT ATT TCC CTC CTC TC-3′		This study
G3f	5′-ATG CGT TTC CTG GAC TGT CG -3′	377	This study
G3r	5′-CTA TCT ATC TAT CTA TCT ATC TAT-3′		This study

amplification, the corresponding bases of the complete viperin gene in *S. chuatsi* (Sun and Nie, 2004) were removed before alignment. The alignments of viperin gene sequences and three mtDNA segments (16S, cyt *b* and CO1) were conducted respectively with Clustal X 1.83 (Thompson et al., 1997) with default gap penalties. The mtDNA alignments were straightforward, while the alignment of the viperin gene showed some variations in the lengths of the sequences. The actual intron of viperin gene was located by aligning the nucleotide sequences with the published complete viperin gene sequence of *S. chuatsi* (Sun and Nie, 2004). Alignment of the 16S data was subsequently adjusted manually on the basis of secondary structure following the published model of *Galaxias breyipinnis* (Waters et al., 2000). All the first two introns of the *S7* were aligned using the direct optimization method implemented in the computer program POY 3.0.11 (Gladstein and Wheeler, 2002), with parsimony as the optimality criterion. For all searches reported, we presented the best (lowest cost) trees from 100 random addition replicates (commands: *-random 100 –maxtrees 3*). The aligned matrix from this procedure was checked by eye, and the minor adjustments were made manually with SEAVIEW (Galtier et al., 1996). Gaps were considered as missing data rather than fifth characters, to prevent those longer than one or two bases from being taken as representing multiple events (Swofford, 1993). Pairwise distances based on the F84 model (Felsenstein, 1993) were calculated with the DAMBE program (Xia and Xie, 2001). Stationarity of nucleotide composition across taxa were examined using chi-square (χ^2) tests implemented in PAUP* 4.0b10 (Swofford, 2002).

2.4. Phylogenetic analyses

The separate and combined dataset were used to infer phylogenies by maximum parsimony (MP) using PAUP* 4.0b10 (Swofford, 2002), maximum likelihood (ML) using PHYML 2.4.4 (Guindon and Gascuel, 2003) and recently developed partitioned Bayesian analyses (Nylander et al., 2004; Brandley et al., 2005) with MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). The use of partitioned Bayesian analyses can incorporate complex mixed models of nucleotide evolution by allocating independent models to partitions of a heterogeneous dataset. This property has made it suitable for merging heterogeneous data matrices and should reduce systematic error, thus providing more accurate posterior probability estimates (e.g., Brown and Lemmon, 2007; Castoe and Parkinson, 2006; Guo and Wang, 2007). In MP analysis, a branch-and-bound search strategy was employed with all characters treated as equal weights and gaps as missing data. The best-fitting models of sequence evolution for ML and Bayesian analyses were estimated using Modeltest 3.7 (Posada and Crandall, 1998) under the Akaike information criterion (AIC; Akaike, 1974), following recent recommendations by Posada and Buckley (2004). To test for nodal reliabilities, heuristic bootstrap analyses (Felsenstein, 1985; 1000 replicates for MP and ML) were applied, with groups appearing in 50% or more of the trees in bootstrap analysis retained.

In Bayesian analyses, cyt *b* and CO1 was partitioned by codon position, 16S gene by the paired position and unpaired position, and the viperin gene by the individual exon and intron. The models implemented in our Bayesian phylogenetic analyses are listed in Table 3. Bayesian analyses started with randomly generated trees and four Markov chains under default heating values were run for 2×10^7 generations, with sampling at intervals of 200 generations. To ensure that the analyses were not trapped on local optima, the dataset was run three times independently. The length of the burn-in was determined by plotting the ln likelihood of the trees sampled, by examining the potential scale reduction factor for all variables, and plotting the frequency of bifurcations in the sampled trees in the two runs using AWTY (Wilgenbusch et al., 2004). After determining chain convergence, which generally occurred within the first 2–5 million generations of each analysis, we followed a conservative approach by discarding all samples obtained during the first 8 million generations as "burn-in". We then generated 50% majority-rule consensus trees with posterior probability values for each node in the context of the final 12 million generations obtained during each analysis.

Partitioned Bremer support analysis (PBS; Bremer, 1994), as calculated using PAUP* 4.0b10 (Swofford, 2002) and TreeRot V2 (Sorenson, 1999), was used to assess the respective contribution of each gene to the total nodal Bremer support. We also tested for incongruent phylogenetic signal between nuclear and mt genes using the partition homogeneity (PHT) test (essentially the incongruence length difference (ILD) test of Farris et al. (1994) in PAUP* 4.0b10 (Swofford, 2002). Even if incongruence among datasets was

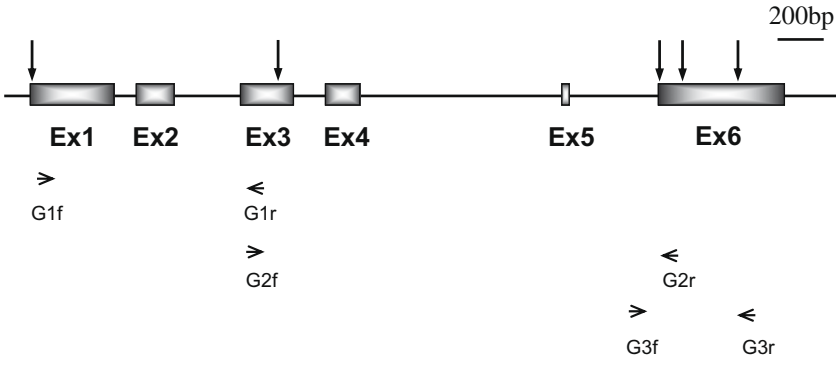

Fig. 1. Diagram of *Siniperca chuatsi* viperin gene structure drawn according to Sun and Nie (2004) and the primer positions in this study. The gene diagram is drawn approximately to scale, except for the arrows designating the primer localization and direction of PCR. Black boxes represent exons.

Table 3
Data partitions, their estimated models of sequence evolution, and total number of characters of each partition used in Bayesian analyses.

Data partition	Model	Number of characters in partition
Cyt b all sites	GTR + I + G	1140
Cyt b 1st codon	K80 + G	380
Cyt b 2nd codon	TVM + I	380
Cyt b 3rd codon	GTR + I + G	380
CO1 all sites	HKY + I + G	666
CO1 1st codon	GTR + I + G	222
CO1 2nd codon	F81 + I	222
CO1 3rd codon	GTR + G	222
16S all sites	GTR + I + G	506
16S stems sites	SYM + I	223
16S loops sites	GTR + I + G	283
S7 intron 1	GTR + G	851
S7 intron 2	GTR + G	642
Exon 1 of viperin gene	K80 + I	358
Intron 1 of viperin gene	GTR	106
Exon 2 of viperin gene	K80	62
Intron 2 of viperin gene	GTR + I	98
Exon 3 of viperin gene	GTR + I	233
Intron 3 of viperin gene	GTR	105
Exon 4 of viperin gene	GTR	156
Intron 4 of viperin gene	GTR + I	941
Exon 5 of viperin gene	K80	33
Intron 5 of viperin gene	GTR + I	736
Exon 6 of viperin gene	GTR + I	374
Viperin all sites	HKY + G	3502

observed, the combined data approach was adopted, which has been favored by many to fully utilize genetic information in deriving phylogenetic estimates (Rokas and Carroll, 2006).

2.5. Alternative phylogenetic hypotheses tests

To compare competing sinipercids topologies, sitewise log-likelihoods were calculated for each topology in PAUP* and used as input for CONSEL 0.1f program (Shimodaira and Hasegawa, 2001). CONSEL was used to calculate the probability values according to the approximately unbiased (AU) test using the multiscale bootstrap technique (Shimodaira, 2002). For comparison, we also performed the Templeton test (Templeton, 1983) and Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa, 1999) using 1000 bootstrap replicates with the resampling of estimated log-likelihood (RELL) optimization as implemented in PAUP*. SH and AU tests are appropriate in comparing both a priori and a posterior hypothesis. However, the SH test is conservative, being less likely to reject alternative topologies under consideration, while the AU test uses a multiscale bootstrap approach to remove this bias (Strimmer and Rambaut, 2002).

3. Results

3.1. Sequence characteristics

Summaries of sequence characteristics for individual and combined gene partitions are given in Table 4 from 7307 characters of sequence data for 13 taxa. The concatenated mtDNA gene fragments consisted of 2312 sites, with 733 variable characters (VC) and 489 parsimony-informative characters (PIC). The concatenated nuclear datasets consisted of a matrix of 4995 base pairs (bp). Of these, 1737 bp were VC and 667 bp were PIC. The aligned nuclear dataset was assembled from three fragments: Viperin gene dataset (3502 bp/939 VC/300 PIC); *S7* intron 1(851 bp/391 VC/224 PIC); *S7* intron 2 (642 bp/407 VC/143 PIC). Viperin gene showed the secondarily lowest proportion of PIC (8.57%) compared with the other gene sequences. Base composition was AT-rich biased in all the six genes. The viperin gene was the secondly most AT-rich (59.02%), and showed the lowest transition/transversion ratio (maximum likelihood Ti/Tv = 1.09), contrasted with obvious tendency of transitions in the mtDNA genes (4.57 in cyt *b*, 6.23 in CO1 and 7.68 in 16S). F84 distances within the ingroup taxa ranged from 0.7 to 43.2% for *S7* intron 1 (19.5% on average), from 0.8 to 74.2% for *S7* intron 2 (26.8% on average), from 0.3–12.5% for viperin (6% on average), and from 0.5–20.1% for combined mt DNA sequences (11.1% on average).

As shown in Table 5, we were unable to obtain sequences from intron 3 to exon 6 of viperin gene in *C. herzi*, from exon 1 to intron 1 in *N. spinosus* and from intron 4 to exon 6 in *L. maculatus*, and therefore the three species have fewer characters than all other species in the matrix. The effects of such incomplete matrices are difficult to predict, but simulations suggest the inclusion of a limited amount of missing data was unlikely to distort the phylogenetic results (Wiens, 2003). The viperin gene fragment obtained ranges in size from 979 to 3181 bp. The detailed length of each exon and intron in viperin gene sequences across all taxa was listed in Table 5. For example, the first intron varies from 79 (*Coreoperca whiteheadi*) to 104 bp (*L. maculatus*) in size, but within the genus *Siniperca*, the size is constant (86 bp).

3.2. Phylogenetic analyses of different genes and gene combinations

3.2.1. Mitochondrial genes

Three mt genes (16S, cyt *b* and CO1) were combined and analyzed simultaneously (≈2312 bp). The topologies obtained by various analytic methods were similar (MP tree length = 1680, CI = 0.6232, and RC = 0.3058) and exhibited much improved resolution and nodal support than either mtDNA gene alone (see

Table 4
Summaries of sequence characteristics for the separate gene partitions and simultaneous analysis datasets.

	Nuclear datasets				Mitochondrial datasets				Combined dataset
	Viperin	*S7 intron 1*	*S7 intron 2*	Nuclear DNA	Cyt *b*	*16S*	*CO1*	mtDNA	
Aligned sites	3502	851	642	4995	1140	506	666	2312	7307
%A	29.03	22.63	32.01	28.26	25.00	29.84	23.85	25.72	27.30
%C	19.91	19.71	15.23	19.30	31.48	24.75	29.49	29.44	23.13
%G	21.07	23.84	21.65	21.63	15.17	23.45	18.64	17.97	20.25
%T	29.99	33.82	31.11	30.81	28.35	21.96	28.02	26.86	29.32
Variable sites	939	391	407	1737	438	81	214	733	2470
Parsimony-informative sites	300	224	143	667	302	35	152	489	1156
Maximum likelihood Ts/tv ratio	1.09	1.18	1.12	1.06	4.57	7.68	6.23	4.82	1.72
Mean F84 distance(%) within ingroup	6 (0.3–12.5)	19.5 (0.7–43.2)	26.8 (0.8–74.2)	10.4 (0.86–26.1)	15.4 (0.88–30.1)	3..3 (0.20–7.43)	11.8 (0–20.4)	11.1(0.5–20.1)	10.7 (1.1–22.9)

Table 5
The length of each exon and intron in viperin gene sequences across all taxa.

Species	Exon 1	Intron 1	Exon 2	Intron 2	Exon 3	Intron 3	Exon 4	Intron 4	Exon 5	Intron 5	Exon 6	Total
S. chuatsi	358*	86	162	259	230	94	150	865	33	362	373*	2972
S. scherzeri 1	358*	86	162	257	209	94	150	891	33	368	374*	2982
S. scherzeri 2	358*	86	162	291	209	94	150	895	33	368	374*	3020
S. kneri	358*	86	162	261	230	94	150	874	33	363	345*	2985
S. fortis	358*	86	162	236	209	94	150	888	33	374	374*	2964
S. obscura	358*	86	162	256	230	94	150	888	33	374	374*	3005
S. undulata	358*	86	162	264	230	94	150	881	33	374	374*	3006
S. roulei	358*	86	162	259	230	94	150	870	33	372	374*	2988
C. herzi	358*	81	162	206	172*	–	–	–	–	–	–	979
C. kawamebari	358*	81	162	186	230	94	150	795	33	382	208*	2679
C. whiteheadi	358*	79	162	194	230	99	150	799	33	704	373*	3181
N. spinosus	–	–	83*	177	230	89	150	810	33	155	287*	2014
L. maculatus	349*	104	162	178	233	101	156	402*	–	–	–	1685

"*" represents that the obtained fragment was not complete.
"-" represents that the fragment was not obtained.
"ex" and "in" is the abbreviation of exon and intron, respectively.

Appendix 1–3). MP and ML analyses of the combined mt data were shown respectively in Appendix 4A and 5A, and Bayesian tree was presented in Fig. 2A. As expected, the monophyly of both *Siniperca* and *Coreoperca* was recovered with strong posterior probability (PP = 1.0 and 1.0, respectively). Similar support values for the corresponding nodes were also obtained by MP (100% and 90%, respectively) and ML (100% and 90%, respectively). However, the monophyly of all recognized sinipercids, i.e., the family Sinipercidae, was not corroborated in all analyses, with *N. spinosus* joining *Siniperca* in Bayesian tree (PP = 0.97) or *Coreoperca* in MP tree (BP = 57). The interlationships of *Siniperca* were well-resolved and strongly supported for almost all nodes. The basal split in the genus is between *Siniperca roulei* and a strongly supported subclade of other *Siniperca* species (PP = 1.0, BP = 85% and 93% for MP and ML, respectively). Within the remaining *Siniperca* species, *S. chuatsi* and *S. kneri* consistently showed a robust sister species and *Siniperca obscura* was identified as the sister taxon to them. Similarly, *Siniperca fortis* and *Siniperca undulata* consistently showed a sister species and *Siniperca scherzeri* was identified as the sister taxon to them, despite the MP bootstrap value lower than 50%. In the genus *Coreoperca*, *C. herzi* was moderately placed as the sister taxon to *Coreoperca kawamebari* and *C. whiteheadi* under the

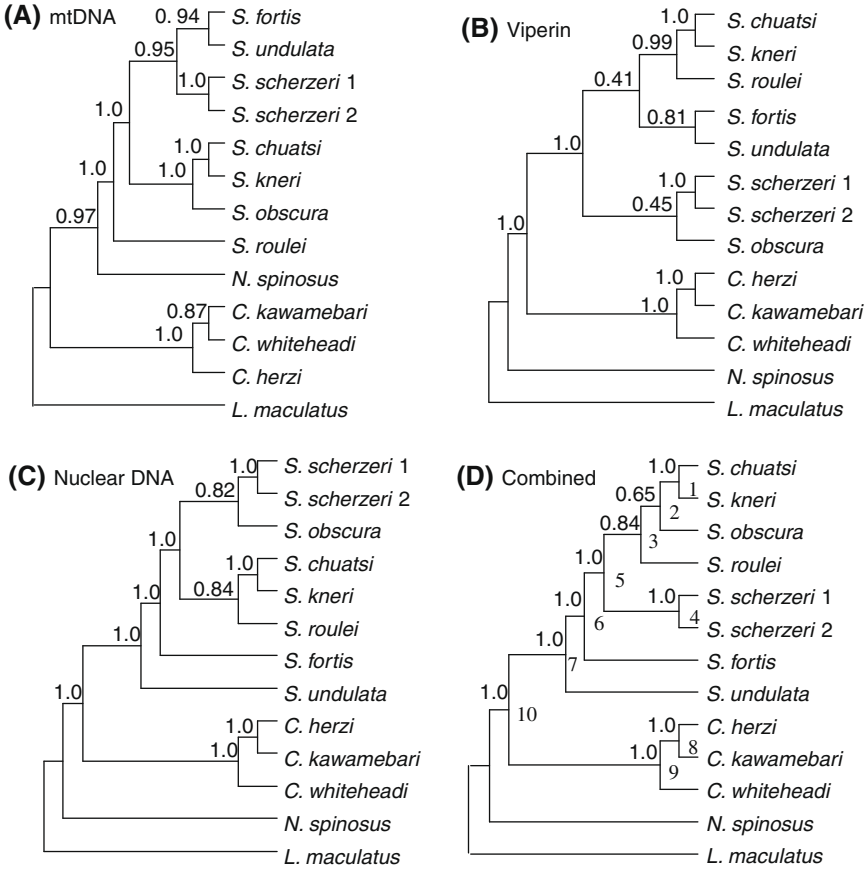

Fig. 2. A majority-rule consensus of trees sampled from the posterior distribution (at stationarity) of the Bayesian analysis, with posterior probabilities above the branch. Numbers behind the nodes are arbitrarily defined numbers represent the clade that follow and which can be used in Table 6.

Bayesian method (PP = 0.87), as seen in Fig. 2A, or collapsed into a polytomy under MP and ML methods.

3.2.2. Nuclear genes

MP consensus trees based on analyses of individual viperin dataset and combined nuclear dataset (viperin + *S7* intron 1 and intron 2; ≈4995 bp) were shown in Appendix 4B and C, respectively. The topologies of two genelogical trees contrasted in several aspects, especially for the placement of *S. undulata* and *S. fortis*. Moreover, the 50% majority consensus trees from the ML and Bayesian analyses (Appendix 5B and C, Fig. 2B and C) were not completely consistent with the corresponding MP tree. Notably, the topological differences from different methodologies were weakly supported (BP < 50%, PP < 0.9). The combined nuclear data tree tended to have overall greater resolution and higher nodal supports. Combined nuclear data analyses consistently identified that the basal split in the genus *Siniperca* is between *S. undulata* and a strongly supported subclade of the remaining *Siniperca* species (PP = 1.0, BP = 85 and 93% for MP and ML, respectively). Within the latter subclade, *S. fortis* showed to be the sister taxon to other *Siniperca* species except *S. undulata* (MP, 68%; ML, 89%; Bayesian, 1.0). As with the mt DNA analyses, our combined nuclear data analyses clearly confirmed closet affinity between *S. chuatsi* and *S. kneri*. Contrast to mtDNA, our nuclear data analyses strongly supported that *C. herzi* and *C. kawamebari* formed as sister species, being sister to *C. whiteheadi*. As shown in Fig. 2 B and C, Appendix 4B and C, and Appendix 5B and C, the reciprocal monophyly of the genus *Siniperca* and *Coreoperca* as well as the family Sinipercidae was all recovered in viperin and combined nuclear analyses.

3.2.3. Combined data (mt DNA + nuclear DNA)

Parsimony analysis using equal weights only resulted in a single MP tree (tree length = 4046, CI = 0.7731, and RC = 0.6183), which is shown in Appendix 4D. ML and Bayesian analyses, identical topology with each other, were essentially consistent with the MP tree except for the placement of *S. undulata* and *S. fortis*. Clearly, comparisons between the MP and Bayesian trees suggested that the Bayesian tree provides higher resolution than the MP. As well with combined nuclear data, our combined data (mtDNA + nuclear DNA) analysis clearly confirmed a well-supported monophyletic sinipercid fish with two distinct genera in accordance with Liu and Chen (1994), one represented by *Siniperca* and the other by *Coreoperca*. ML analysis of the combined data is presented in Appendix 5D. This multiple gene tree is completely resolved and strongly supported for all nodes except for the positions of *S. roulei* (69%) and *S. obscura* (53%). Similar support values for the two nodes were also obtained by Bayesian analyses (Fig. 2D; 0.84 for *S. roulei*, 0.65 for *S. obscura*). In the *Siniperca*, *S. undulata* occupied the most basal position, followed by *S. fortis*, *S. scherzeri*, *S. roulei*, *S. obscura*, and last tow most recently diverged sister species *S. chuatsi* and *S. kneri*. As for the interrelationships in *Coreoperca*, they are identical with those of nuclear datasets.

The partition homogeneity tests showed that no heterogeneity in phylogenetic signal was detected between any of the mtDNA genes, and that partitions were significantly heterogeneous between viperin and *S7* intron 1, 2 ($P = 0.013$), and between combined mtDNA and combined nuclear DNA ($P = 0.001$), etc. Partitioned Bremer values (Table 6) reveals that the majority of phylogenetic information from the combined gene-based topologies was contributed by nuclear character, with *S7* intron 1 holding the highest percentage (59.35%), followed by viperin (27.43%), and *S7* intron 2 (10.22%), in sharp contrast to the mitochondrial characters (2.99% in total). The relatively weak influence of mitochondrial partition upon analysis of complete set of genes was somewhat surprising, given their relatively high degree of genetic divergence and rich informative characters. The breakdown of the PBS values indicated some conflict between the mt DNA datasets and the nuclear DNA datasets, because 6 of 10 nodes resolved in our combined dataset consensus tree had conflicting PBS values. This pattern may have resulted from either differences in internal homoplasy within each partition or potential conflicting signals, as indicated by PHT test, between mtDNA and nuclear DNA.

The competing hypotheses inferred using different partition datasets were also statistically tested among each other by AU, SH and Templeton tests (Table 7). Under the combined dataset, the three tests supported that *S7* intron 2 Bayesian tree and mtDNA Bayesian topologies including those from combined mtDNA, 16S, CO1 and cyt *b* are all safely rejected ($P < 0.05$). However, the hypotheses from viperin and *S7* intron 1 can only be significantly rejected by AU test ($P = 0.029$, <0.001, respectively) and Templeton test ($P = 0.0061$, 0.0005, respectively), but cannot by SH test ($P = 0.611$, 0.076, respectively). This echoes Strimmer and Rambaut (2002) in that the SH test is more conservative (i.e. less likely to reject alternative topologies under consideration) than AU test. The combined dataset Bayesian hypothesis and monophyly constraint hypothesis were not significantly different, which were covered similar log-likelihood score and tree length. Thus, in our view, the hypotheses based on ML and Bayesian analyses of combined datasets (Appendix 5D and Fig. 2D) represent the best current estimate of sinipercids phylogeny in the present phylogenetic reconstruction.

4. Discussion

4.1. Monophyly and interspecific relationships of sinipercid fish

Traditionally, sinipercid fish are composed of twelve recognized extant species. With the unusual morphology and relatively few perciform fish in freshwater, these fish have been the subject of

Table 6
Results of partitioned Bremer support (PBS) analyses for each node on the total evidence Bayesian tree (combined nuclear and mt dataset).

Nodes[a]	Mitochondrail datasets				Nuclear datasets				Combined nuclear and mt gene trees
	16S	*CO1*	*Cyt b*	Combined mt	Viperin	*S7* intron 1	*S7* intron 2	Combined nuclear	
1	−1	5	13	17	4	21	0	25	42
2	−2	−2	−4	−8	−4	14	−1	9	1
3	−2	−2	−4	−8	−4	14	−1	9	1
4	0	3	6	9	11	16	−1	26	35
5	−2	−2	−4	−8	−4	14	−1	9	1
6	−2	−2	−4	−8	−4	14	−1	9	1
7	2	13	16	31	45	45	12	102	133
8	0	3	−16	−13	8	49	2	59	46
9	1	7	−4	4	51	47	25	123	127
10	−3	0	−1	−4	7	4	7	18	14
Total PBS	−9	23 (5.74%)	−2	12 (2.99%)	110 (27.43%)	238 (59.35%)	41 (10.22%)	389 (97.01%)	401

[a] Nodes are numbered as the combined data Bayesian tree in Fig. 2D.

Table 7
Statistical comparison of alternative topologies including different Bayesian trees based on the separate gene partitions and simultaneous analysis datasets using approximately unbiased (AU) test of Shimodaira (2002) and Shimodaira and Hasegawa (1999) (SH) and Templeton (1983) test under the combined dataset.

Topology	Approximately unbiased test and Shimodaira–Hasegawa test				Templeton test			
	−Ln L	*−Ln L diff.*	AU test: *P*	SH test: *P*	Tree length	*N*	*Z*	*P*
Combined dataset (mt + nuclear DNA): Fig. 2D	28398.62857	1.32904	0.593	0.950	4046	Best		
Viperin: Fig. 2B	28420.86892	23.56939	0.029*	0.611	4071	80	−2.7441	0.0061*
S7 intron 1: Appendix 3D	28470.28320	72.98367	0.000*	0.076	4096	193	−3.4677	0.0005*
S7 intron 2: Appendix 3E	28537.27805	139.97852	0.000*	0.000*	4097	135	−4.2953	<0.0001*
Combined nuclear DNA dataset: Fig. 2C	28413.10487	15.80534	0.126	0.725	4055	45	−1.3416	0.1797
Combined mtDNA dataset: Fig. 2A	28481.81722	84.51769	0.000*	0.016*	4106	303	-3.3975	0.0007*
Cyt *b*: Appendix 3A	28852.50731	455.20778	0.000*	0.000*	4284	469	−10.3979	<0.0001*
CO1: Appendix 3B	28515.19434	117.89480	0.000*	0.002*	4238	363	−9.3672	<0.0001*
16S: Appendix 3C	28824.37365	427.07412	0.000*	0.000*	4289	526	−10.2228	<0.0001*
Liu and Chen (1994) [a]	28519.57607	122.27653	0.000*	0.003*	4127	188	−5.4601	<0.0001*
Sinipercids monophyly[a]	28397.29953	Best	0.594	–	4048	76	−0.2294	0.8185

Likelihood settings from best-fit model for combined mt and Nuclear DNA (GTR + I + G) selected by AIC in Modeltest
* Associated probabilities are given and significantly ($P < 0.05$) worse topologies are indicated by asterisk.
[a] Maximum likelihood tree recovered from constraint search under the combined dataset.

several taxonomic and phylogenetic studies for decades. However, as summarized by Chen et al. (2007), different views still exist over the taxonomic status and phylogeny of sinipercids. In support of Zhou et al. (1988), Kong and Zhou (1993) suggested that the sinipercids should be divided into three genera: *Coreoperca*, *Siniperca* and *Coreosiniperca* on the basis of osteological work. Kong and Zhou (1992) anticipated the monophyly of sinipercid fish based on similar relative-mobility of LDH isozyme in 9 sinipercid species. On the basis of cladistic analysis of their osteological characters, Liu and Chen (1994) confirmed that the sinipercid fish is a monophyletic group consisting of two genera, namely, *Siniperca* and *Coreoperca*, and the genus *Coreosiniperca* was unwarranted and *Coreosiniperca roulei* is a member of the genus *Siniperca*. The interspecific relationships within this group have been also equivocal probably due to rapid cladogeneis resulting in short internodes that have not been resolved because of insufficient sequence data used in previous studies (e.g., Chen et al., 2007; Zhao et al., 2006b).

In the present study, although individual genes examined failed to provide a well-supported phylogeny, the ML and Bayesian analyses of concatenated mt and nuclear data, consistently produced a robust, well-resolved tree. Our phylogenies strongly uphold the monophyly of sinipercid fish (Sinipercidae), *Siniperca* and *Coreoperca*, which agrees well with several key features of recent perspectives of sinipercid systematics. *Coreosiniperca roulei* (= *Siniperca roulei*) does not form an independent group, instead merges into the genus *Siniperca*, which is in accord with the results of Liu and Chen (1994) and Zhao et al. (2006a,b). The internal affinities within *Coreoperca* are shown to be novel, and our results differ from all previous hypotheses (Fig. 3). Our analyses clearly indicate that *C. whiteheadi* is the sister taxon to *C. herzi* plus *C. kawamebari*. Similarly, the interrelationships in the genus *Siniperca* are different from previous hypotheses (Fig. 3). Our results demonstrate that *S. undulata* is the sister taxon to the other members of *Siniperca*, within the clade containing the other members of the genus, *S. chuatsi* and *S. kneri* are sister species, next joined by *S. obscura*, *S. roulei*, *S. scherzeri* and finally by *S. fortis*.

Several interesting phenomena with respect to the phylogenetic relationships of sinipercid fish are observed in the present study. Firstly, mitochondrial gene whenever single or simultaneous analysis cannot recover the monophyly of sinipercids, while the nuclear viperin gene, the *S7* intron 2 and the total data supported the monophyly of this family with higher bootstrap value or posterior possibility. Secondly, the monophyly of two genera *Siniperca* and *Coreoperca* were supported by viperin gene, simultaneous mtDNA or nuclear DNA and combined data in three analytical methods. Thirdly, most single gene and the total data supported the mono-

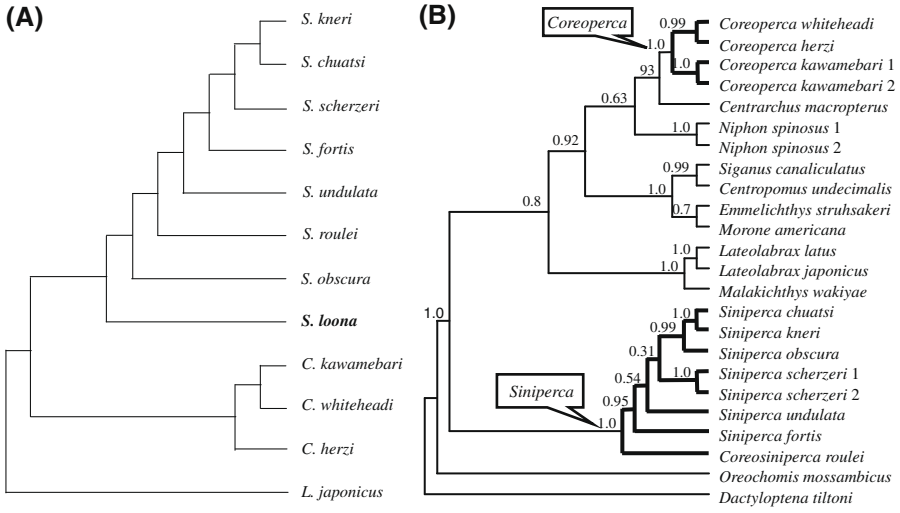

Fig. 3. Competing hypotheses of phylogenetic relationships within sinipercid fish deduced from (A) morphology (Liu and Chen, 1994). (B) Complete cyt *b* gene sequence (Chen et al., 2007).

phyly of genus *Siniperca*, while single cyt *b* and 16S cannot corroborate the genus *Coreoperca* as a monophyletic group.

Taking all evidence together, we can find that nearly no two analyses have come to a completely identical result regarding affinities among theses seven species in *Siniperca* (with the exception of ML and Bayesian trees from combined mt and nuclear DNA dataset in the present study), with conclusions varying depending on the analytic methods and character type used. However, given the large size of our dataset and the robust support of many nodes, we consider our total data gene tree as the preferred interpretation of sinipercid fish relationships. AU, SH and Templeton tests suggest that the combined data topology we recovered (Fig. 2D and Appendix 5D) were most possible estimate for present dataset, although the results from combined nuclear DNA cannot be statistically rejected by all the tests at 5% significant level.

4.2. Molecular evolution patterns and phylogenetic utility of different mt and nuclear genes in sinipercid fish

Our results indicate that nuclear genes/introns have much advantage over mitochondrial genes in equal weights parsimony analysis. Nuclear genes/introns had universally higher values of CI as compared to mitochondrial genes (Table 4), and generally provided more in the way of partitioned Bremer support than the mitochondrial genes (Table 6). These results suggest the view among sinipercid fish that mitochondrial genes show higher levels of homoplasy and are of less utility than nuclear genes.

Given the detected conflict between the mitochondrial genes and nuclear DNA, it is important to identify the source of this disagreement. By performing Bayesian analysis using a GTR + site-specific rates (SSR) model with rate categories corresponding to gene, it was possible to quantitatively compare the substitution patterns presented by different mitochondrial genes and nuclear genes/introns. Simon et al.(1994) concluded in general terms that nuclear genes evolve more slowly than mitochondrial genes, making nuclear genes better markers for deep divergence. This may be the general tendency for exons, but it does not necessarily apply when comparing nuclear introns with mitochondrial genes. When we consider our data alone this pattern is not seen because the relative rate presented by *S7* inton1 (0.905) and intron 2 (1.432) was greater than 16S (0.327). Indeed, this pattern could already be expected because the mean F84 distance values obtained for S7 intron 1 (0.195) and intron 2 (0.268) was much greater than that for 16S.

It was previously seen that the nuclear DNA genes/introns showed greater base compositional bias than the mitochondrial genes, but other patterns of nucleotide substitution, such as the Q matrix of transformation, may also be important. In general, the instantaneous rate matrices for mitochondrial genes was more asymmetrical relative to those presented by nuclear DNA (viperin, *S7* intron 1 and intron 2) and were also more skewed towards one type of change over another (Fig. 4B), although in all cases there was a higher overall rate of transitions. The obvious consequence of these highly skewed transformation rate matrices is greater levels of homoplasy, which are not easily corrected for.

Another parameter that often differs between nuclear intron and mitochondrial gene is the shape of the gamma distribution, as given by α value, describing the among-site rate variation. As can be seen from Fig. 4C, there is much more heterogeneity in among-site rate variation in mitochondrial genes than in nuclear *S7* intron 1 and intron 2, with the most heterogeneous in cyt *b* gene, which had the lowest α value. However, all the nuclear sequences presented lower proportion of invariable sites (Pi values), which is contrast to that of Lin and Danforth (2004), who demonstrated that the proportion of invariable sites is positively correlated with α value. These authors also showed a positive

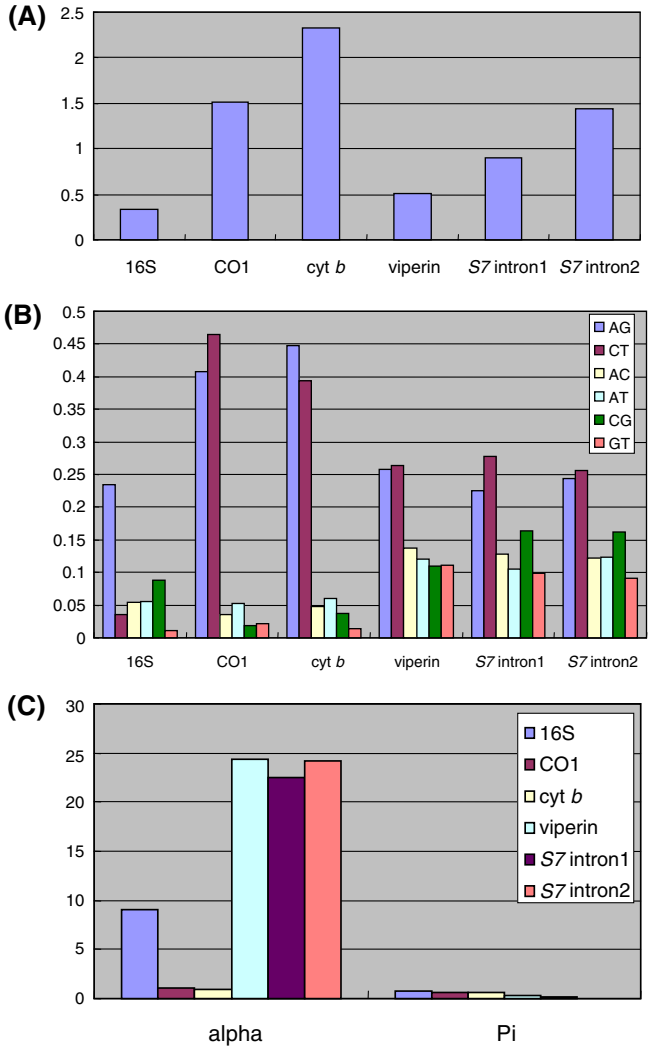

Fig. 4. The mt and nuclear datasets nucleotide substitution patterns based on the GTR + SSR model, with sites partitioned by gene. (A) Relative rates of nucleotide substitution; (B) transfomation rate mitrices; (C) gamma shape, proportion of invariable sites (Pi).

correlation between α value and CI, the consistency index, which suggests that data partitions with more heterogeneous substitution rates show a higher level of homoplasy. This correlation was detected by us, with the nuclear datasets having higher CI values (viperin, 0.91; *S7* intron 1, 0.86; *S7* intron 2, 0.9) than the mt genes (cyt *b*, 0.62; 16S, 0.71; CO1, 0.62).

The three mt genes included in this study failed to recover the monophyly of sinipercid fish, thus making the combination of all mt genes unwarranted for providing persuasive evidence of the resolution of all parts of the trees. *S7* intron 1 and intron 2 datasets in present study proved especially informative in phylogenetic reconstruction of the sinipercid fish. Both introns evolved at a rapid rate and no sign of significant saturation were observed. The *S7* intron 1 was found to contribute the most phylogenetic signal while *S7* intron 2, only less than *S7* intron 1 and viperin to the final tree (Fig. 2D; see Table 6). However, like the mitochondrial genes in this study and cyt *b* in Chen et al. (2007), S7 intron 1 failed to confirm the monophyly of sinipercid fish either.

It is notable that our study provides valuable information on the utility of nuclear viperin gene in the phylogenetic reconstruction for the first time. The viperin gene was explored in the present study to reconstruct phylogeny of sinipercid fish, encouraged by

the fact that this gene has been shown to contain valuable phylogenetic signals corroborating the monophyly of *Siniperca*, *Coreoperca*, and a single origin of sinipercid fish. However, the viperin gene alone cannot resolve the interspecific relationships of the genus *Siniperca*. It is expected that this novel nuclear marker would be better suited for resolving supergeneric (e.g., interfamilial) relationships among perciform and other fish.

In conclusion, our results demonstrate that using just one of these markers would not have provided a well-resolved phylogeny, even with the most sophisticated analytical strategies we could find in the literature. This lack of resolution is largely due to insufficient phylogenetic information in individual loci. Analysis of these heterogeneous data using partitioned models of sequence evolution removes the problems associated with combining data and appears to be the most logical way to analyze heterogeneous data. By further combining mt DNA and nuclear datasets, we have been able to provide a well-supported, completely sampled phylogenetic hypothesis for sinipercid fish.

Acknowledgments

We are grateful to Dr T. King, L. Yang, Y. Zhang, Z. Li, A. Li, D. Wang, L. Zhou and Mrs. M. Pan, for assistance in collecting specimens. We also thank two anonymous reviewers for their insightful comments on this manuscript. This study was supported by a joint fund from the National Natural Science Foundation of China and the Government of Guangdong Province (U0631010). Dali Chen was supported by the National Natural Science Foundation of China (30800094) and Education Foundation for Young Teachers of Sichuan University (07056). Xianguang Guo was supported by the National Natural Science Foundation of China (30700062), and the Knowledge Innovation Program of the Chinese Academy of Sciences (CIB-2007-LYQY-Q01, KSCX2-YW-Z-005).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.01.039.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control 19, 716–723.
Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
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.
Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.
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.
Castoe, T.A., Parkinson, C.L., 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol. Phylogenet. Evol. 39, 91–110.
Chen, D., Guo, X., Nie, P., 2007. Non-monophyly of fish in the Sinipercidae (Perciformes) as inferred from cytochrome *b* gene. Hydrobiologia 583, 77–89.
Chow, S., Hazama, K., 1998. Universal PCR primers for *S7* ribosomal protein gene introns in fish. Mol. Ecol. 7, 1255–1256.
Cronn, R.C., Small, R.L., Haselkorn, T., Wendel, J.F., 2003. Cryptic repeated genomic recombination during speciation in Gossypium gossypioides. Evolution 57, 2475–2489.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
Felsenstein, J.P., 1985. Confidence limits on phylogenetics: an approach using the bootstrap. Evolution 39, 783–791.
Felsenstein, J.P., 1993. PHYLIP 3.5 (phylogeny inference package). Department of Genetics, University of Washington, Seattle.
Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO-WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.
Gladstein, D., Wheeler, W.C., 2002. POY. Program and Documentation, Version 3.0. American Museum of Natural History, New York.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Guo, X., Wang, Y., 2007. Partitioned Bayesian analyses, dispersal-vicariance analysis, and the biogeography of Chinese toad-headed lizards (Agamidae: *Phrynocephalus*): a re-evaluation. Mol. Phylogenet. Evol. 45, 643–662.
Johnson, G.D., 1984. Percoidei: development and relationships. In: Moser, H.G., Richards, W.J., Cohen, D.M., Fahay, M.P., Kendall, A.W., Richardson, S.L. (Eds.), Ontogeny and Systematics of Fishes. Am. Soc. Ichthyol. Herpetol. Spec. Publ. vol. 1, pp. 464–498.
Kong, X., Zhou, C., 1992. Comparative studies on LDH isozyme in sinipercinae fish of China. J. Ocean Univ. Qingdao 22, 103–110 (In Chinese with English abstract).
Kong, X., Zhou, C., 1993. Comparative studies on the skeletal characteristics of seven Sinipercinae fish of China. J. Ocean Univ. Qingdao 23, 116–124 (In Chinese with English abstract).
Lin, C.-P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analysis of combined datasets. Mol. Phylogenet. Evol. 30, 686–702.
Liu, H., 1997. Study on systematic position of sinipercine fish with discussion on relationships of some lower perciforms. Trans. Chin. Ichthyol. Soc. 6, 1–7 (In Chinese with English abstract).
Liu, H., Chen, Y., 1994. Phylogeny of the sinipercine fish with some taxonomic notes. Zool. Res. 15 (Suppl.), 1–12 (In Chinese with English abstract).
Nelson, J.S., 2006. Fishes of the World, fourth ed. John Wiley and Sons, New Jersey.
Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogeneitc analysis of combined data. Syst. Biol. 53, 47–67.
Palumbi, S.R., 1996. Nucleic acids II: the polymerase chain reaction. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics, second ed. Sinauer Associates, Sunderland, Massachusetts, pp. 205–247.
Pons, J., Barraclough, T.G., Theodorides, K., Cardoso, A., Vogler, A.P., 2004. Using exon and intron sequences of the gene MP20 to resolve basal relationships in *Cicindela* (Coleoptera: Cicindelidae). Syst. Biol. 53, 554–570.
Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics advantages of akaike information criterion and Bayesian approach. Syst. Biol. 53, 793–808.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Roberts, C.D., 1993. Comparative morphology of spined scales and their phylogenetic significance in the Teleostei. Bull. Mar. Sci. 52, 60–113.
Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804.
Rokas, A., Carroll, S.B., 2006. Bushes in the tree of life. PLoS Biol. 4, 1–6.
Ronquist, F.R., Huelsenbeck, J.P., 2003. MRBAYES: bayesian inference of phylogeny. Bioinformatics 19, 1572–1574.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, New York.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.
Shirai, S.M., Yabumoto, Y., Kim, I., Zhang, C.G., 2003. Phylogeny of sinipercid fish and their relatives inferred from mtDNA cytochrome *b* gene: a preliminary study. Bull. Kitakyushu Mus. Nat. Hist. Hum. Hist. Ser. A 1, 45–49.
Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.
Slade, R.W., Moritz, C., Heideman, A., 1994. Multiple nuclear gene phylogenies: application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. Mol. Biol. Evol. 11, 341–356.
Sorenson, M.D., 1999. TreeRot, Version 2. Boston University, Boston, Massachusetts.
Strimmer, K., Rambaut, A., 2002. Inferring confidence sets of possibly misspecified gene trees. Proc. R. Soc. Lond. B 269, 137–142.
Sun, B., Nie, P., 2004. Molecular cloning of the viperin gene and its promoter region from the mandarin fish *Siniperca chuatsi*. Vet. Immunol. Immunopathol. 101, 161–170.
Swofford, D.L., 1993. PAUP: Phylogenetic Analysis Using Parsimony. Ver. 3.1.1. Illinois Natural History Survey, Champaign.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.
Templeton, A.R., 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37, 221–244.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface. Flexible strategies for multiple sequences alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
Waldman, J.R., 1986. Systematics of *Morone* (Pisces: Moronidae), with notes on the lower percoids. Ph. D. thesis, Department of Biology, The City University of New York, New York.
Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, D.N., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 360, 1847–1857.
Waters, J.M., Lopez, J.A., Wallis, G.P., 2000. Molecular phylogenetics and biogeography of galaxilid fish (Osteichthyes: Galaxiidae): dispersal,

vicariance, and the position of *Lepidogalaxias salamandroides*. Syst. Biol. 49, 777–795.
Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.
Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2004. AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.
Xiao, W., Zhang, Y., Liu, H., 2001. Molecular systematics of Xenocyprinae (Teleostei: Cyprinidae): taxonomy, biogeography, and coevolution of a special group restricted in East Asia. Mol. Phylogenet. Evol. 18, 163–173.
Xia, X., Xie, Z., 2001. DAMBE: data analysis in molecular biology and evolution. J. Hered. 92, 371–373.
Yabumoto, Y., Uyeno, T., 2000. *Inabaperca taniurai*, a new genus and species of Miocene percoid fish from Tottori Prefecture, Japan. Bull. Natn. Sci. Mus. Tokyo: Ser. C 26, 93–106.
Zhao, J.-L., Li, S.-F., Cai, W.-Q., Wang, W.-W., 2005. The preliminary phylogenetic relationships of sinipercine fish and some lower percoids inferred from 16S ribosomal DNA sequences. J. Shanghai Fisheries Univ. 14, 364–369 (In Chinese with English abstract).
Zhao, J.-L., Li, S.-F., Cai, W.-Q., Wang, W.-W., 2006a. Phylogenetic relationship of sinipercine fish in East Asia based on cytochrome *b* sequence analysis. Acta Zool. Sinica 52, 676–680 (In Chinese with English abstract).
Zhao, J.-L., Wang, W.-W., Li, S.-F., Cai, W.-Q., 2006b. Structure of the mitochondrial DNA control region of the sinipercine fish and their phylogenetic relationship. Acta Gene. Sinica 33, 793–799 (In Chinese with English abstract).
Zheng, C., 1989. Fish of the Zhujiang River. Science Press (In Chinese), Beijing.
Zhou, C., Yang, Q., Cai, D., 1988. On the classification and distribution of the sinipercinae fish (family Serranidae). Zool. Res. 9, 113–125 (In Chinese with English abstract).