Mol Biol Rep (2011) 38:4673–4685
DOI 10.1007/s11033-010-0602-4

The complete mitochondrial genome of bighead croaker, *Collichthys niveatus* (Perciformes, Sciaenidae): structure of control region and phylogenetic considerations

Tian-jun Xu · Yuan-zhi Cheng · Yue-na Sun · Ge Shi · Ri-xin Wang

Received: 7 August 2010 / Accepted: 22 November 2010 / Published online: 4 December 2010

Abstract Sciaenidae is a diverse, commercially important family. To understand the phylogenetic position of *Collichthys niveatus* in this family, we present its complete mitochondrial genome sequence. The genome is 16469 bp in length and contains 37 mitochondrial genes (13 protein-coding genes, 2 ribosomal RNA genes and 22 transfer RNA genes) and a control region (CR) as in other bony fishes. Further sequencing for the complete control region was performed on *Collichthys lucida*. Although the conserved sequence domains such as extend termination associated sequence (ETAS) and conserved sequence block domains (CSB-1, CSB-2 and CSB-3) are recognized in the control region of the two congeneric species, the typical central conserved blocks (CSB-F, CSB-E and CSB-D) could not be detected, while they are found in *Miichthys miiuy* and *Cynoscion acoupa* of Sciaenidae and other Percoidei fishes. Phylogenetic analyses do not support the monophyly of Pseudosciaeniae, which is against with the morphological results. *C. niveatus* is most closely related to *Larimichthys polyactis*, and *Collichthys* and *Larimichthys* may be merged into one genus, based on the current datasets.

Keywords *Collichthys niveatus* · Sciaenidae · Control region · Phylogeny

Electronic supplementary material The online version of this article (doi:10.1007/s11033-010-0602-4) contains supplementary material, which is available to authorized users.

T. Xu · Y. Cheng · Y. Sun · G. Shi · R. Wang (✉)
Key Laboratory for Marine Living Resources and Molecular Engineering, College of Marine Science, Zhejiang Ocean University, Zhoushan 316000, People's Republic of China
e-mail: wangrixin1123@126.com

Introduction

The vertebrate mitochondrial DNA (mtDNA) is a circular molecule with a length of 16–19 kb that contains 37 genes encoding 22 transfer RNAs (tRNAs), 13 proteins, 2 ribosomal RNAs (rRNAs) and a putative control region (CR) [25, 54]. Mitochondrial DNA exhibits several properties that make it a useful tool in the study of phylogenetics, molecular evolution and even conservation genetics, due to its relatively simple genetic structure, maternal mode of inheritance (in most situations), and high rate of evolution [11, 15, 18, 36].

The bighead croaker, *Collichthys niveatus* is distributed in the Western Pacific, Yellow Sea and East China Sea, and inhabits in the costal waters of subtropical sea with the bottom composed of sand or muddy sand [8]. It is a very popular fish for Chinese consumers, and is one of the most important commercial fishes in China. Thus, intensive studies have been carried out in the areas of feeding habits, category composition and abundance of ichthyoplankton, and early life history [49, 52]. However, almost no molecular studies have so far been conducted on this species, and no one has yet analyzed its mitochondrial genome.

The *C. niveatus*, belongs to the family Sciaenidae, which comprises 70 genera and about 270 species [27], fish from this family being popularly known as croakers or drums due to the sound they produce using muscles associated with the gas bladder. Numerous studies of morphological and molecular phylogenetic relationships among Sciaenidae species have been conducted. Based on the characters of gas bladder, sagitta and mental pores, Zhu et al. [59] divided the family into seven subfamilies: Johniinae, Megalonibinae, Bahabinae, Sciaeninae, Otolithinae, Argyrosominae, Pseudosciaeniae. *Johnius belengerii* was placed as a basal

species of Sciaenidae. Pseudosciaeniae was composed of *Miichthys*, *Larimichthys* (*Pseudosciaenidea*) and *Collichthys*, and a close relative to Argyrosominae. *Collichthys* and *Larimichthys* were also supported by many other morphological studies [20, 42, 43].

Compared with morphological studies, the molecular phylogenies have given different results. Tong et al. [46] investigated the phylogenetic relationships of 10 Sciaenidae fishes by using partial 16S rRNA gene sequence. Molecular phylogenetic analysis revealed a non-monophyletic *Larimichthys*, where *L. polyactis* was grouped with *Collichthys* and then grouped with *L. crocea*. Chen [4] also utilized 16S rRNA characters to resolve a non-monophyletic *Collichthys*. However, in a phylogeny proposed by Meng et al. [23] based on the 16S rRNA gene sequences, monophyletic *Larimichthys* and *Collichthys* were supported well.

Because of instability of the phylogenetic position of *C. niveatus* based on these limited data, the aim of this study was to obtain new sequence data of *C. niveatus* for uses in phylogenetic studies of Sciaenidae and population genetic structures. We present the first complete mitochondrial genome of the *C. niveatus*. The mitogenome sequence was described and compared with its coordinal species. Finally, we performed NJ, ME, and ML analyses in order to gain insights into the position of *C. niveatus* within the Sciaenidae by adoption of the three datasets (COI, 16S and Cytb).

Materials and methods

Fish sampling and DNA extraction

The specimens of *C. niveatus* were captured at the Zhoushan fishing ground (Zhejiang, China). A dorsal fin of an individual was partially excised. Genomic DNA was extracted using the conventional SDS\proteinae K method followed by organic extraction and ethanol precipitation [37].

PCR amplification

13 sets of primers were used to amplify contiguous, overlapping segments of the complete mitochondrial genome of *C. niveatus* (Table 1). The 50 μl PCR mixture contained 0.2 μM of each primers, 5.0 μl of 10 × *Taq* Plus polymerase buffer, 0.2 mM dNTPs, 2 unit of *Taq* Plus DNA polymerase with proof-reading characteristic (TIANGEN), and 1 μl of DNA template. PCR was performed on a PTC-200. The conditions of the PCR are as follows: predenaturalization at 94°C for 4 min; 35 cycles of denaturation at 94°C for 50 s, annealing at 60°C for 60 s, extension at 72°C for 2–3 min; and final extension at 72°C for 10 min. The PCR products were electrophoresed on a 1% agarose gel to check integrity and visualized by the Molecular Imager Gel Doc XR system (BioRad, USA). The PCR products were purified using a QIAEX II Gel Extraction Kit (Qiagen).

Cloning, sequencing and sequence analysis

The purified fragments were ligated into PMD18-T vectors (Takara) and transformed to TOP10 cells (TIANGEN) according to the standard protocol. Positive clones were screened via PCR with M13+/− primers. Amplicons were sequenced using the ABI 3730 automated sequencer with M13+/− primers. The obtained sequence fragments were edited in Sequencher™ (Gene Code, Ann Arbor, MI, USA) for a contig assembly to make the complete mitochondrial genome.

Annotation of protein-coding and ribosomal RNA (rRNA) genes and determination of their gene boundaries were carried out with reference sequences of Sciaenidae available in GenBank. Most tRNA genes and their

Table 1 PCR primers in the analysis of bighead crocaker *C. niveatus*

Primer	Sequence (5′–3′)
Coni-1F	ATAAAGACCCGTATGAATGGC
Coni-1R	ATGAGCGGTAAGATAGCAAGG
Coni-2F	CCCTCCAACTCCTTAGAAAAG
Coni-2R	GGTGACCGAAGAATCAGAATA
Coni-3F	GACATTGGCACCCTCTATCTAA
Coni-3R	AGGCAAGGTCTTCGTAATCAGT
Coni-4F	TCATCATCGGCTCTACATTCCTG
Coni-4R	GGCGAGACTGGCAATAAATCATC
Coni-5F	AACTGGCAAATCAGCACAA
Coni-5R	CGAATGAGTCAGCCGTAAT
Coni-6F	TGATACTTCCTATTCGCCTAC
Coni-6R	TCAGATGACAAAGCCACAG
Coni-7F	AACCAACCATAGCCCACGACA
Coni-7R	ATTCATTTCCCAGGCAACCAG
Coni-8F	CCTATTTACCGCTACCTGTGC
Coni-8R	ACGGATGAGAAGGCTATGGA
Coni-9F	GTATGAATGGCAAAACGAGG
Coni-9R	TTTCTAAGGAGTTGGAGGGA
Coni-10F	ATAAGCGGGAGCAACCACCAT
Coni-10R	ATAGGCGGGAGAAAGACAAGG
Coni-11F	TGGAGGCATACCAGTAGAACAC
Coni-11R	CACTCTTTACGCCGTTGACTAT
Coni-12F	TCAGGCACAACCAACCATAGC
Coni-12R	GGTCCGTTCCGAGTTACACTT
Coni-13F	TTCCTTCCACTAACACTTGCC
Coni-13R	GACTGCGTCTATTTTGATGCC

secondary cloverleaf structure were identified in tRNAscan-SE1.21 [22]. The remaining tRNA genes, which could not be found by tRNAscan-SE were identified by sequence homology, secondary structure and specific anti-codon. Nucleotide base frequencies and codon usage of protein-coding genes were determined using MEGA 4 [41]. The complete mitochondrial genome (mitogenome) sequence of *C. niveatus* was deposited in the public database GenBank under accession number HM219223.

Phylogenetic analysis

To determine the phylogenetic position of *C. niveatus*, three data sets (COI, 16S and Cytb) were collected from the sequences available in GenBank/EMBL/DDBJ (Table 2). For each data set, the nucleotide sequences were aligned using the program Cluster X with the default settings [44], and then the sequences were checked and adjusted manually. Phylogenetic trees were estimated for each data set using NJ and ME methods as implemented in MEGA 4. The model GTR + I+G was selected for ML analyses by ModelTest 3.7 [34]. ML analyses were conducted using PhyML 3.0 [9]. Robustness of the inferred trees was evaluated using bootstrap analysis on 1000 replications [7].

Results

Genomic organization

The total length of the *C. niveatus* mitogenome is 16469 bp. The mitogenome content (13 protein-coding genes, 2 rRNAs, 22 tRNAs), gene order and gene coding strand of the bighead croaker mitogenome conform to the vertebrate consensus (see Fig. 1 in Supplementary Material; Table 3) [48, 57]. The overall base composition is: T, 25.0%; C, 31.3%; A, 27.6; G, 16.2% (Table 4). The A + T content is slighter than the G + C content, which is similar to other fishes [1, 47]. An anti-G bias (8.5%) was observed in the third codon position of the protein-coding genes as reported in other vertebrate mitogenomes [17, 50].

Protein-coding genes

The size of the protein-coding genes in the *C. niveatus* mitogenome is similar to the orthologs in the other fishes. Among the 13 protein-coding genes, two reading frame overlaps occurred on the same strand (Table 3): ATPase8 and ATPase6 overlapped by ten nucleotides, and ND4L and ND4 overlapped by seven nucleotides. This is a common vertebrate feature and has been found in other bony fishes [24]. At the second codon position, pyrimidines (T + C=67.8%) are overrepresented in comparison with purines, owing to hydrophobic character of the proteins [26]. All protein-coding genes start with ATG codon. Variance in the stop codons seems to be a common tendency in the fish mitochondrial genome [14, 53]. Open reading frames of the bighead croaker end with TAA (ATPase8, ATPase6, ND4L, ND5 and ND6), TAG (ND1) and AGA (COI) or incomplete stop codons, either TA (ND2, COIII) or T (COII, ND3, ND4, and Cytb). This condition is common among vertebrate mitochondrial genome, and it appears that TAA stop codons are created via posttranscriptional polyadenylation [30].

Codons in 13 protein-coding genes identified in *C. niveatus* are shown in Table 5. For amino acids with fourfold degenerate third position, codons ending in C are mostly seen, followed by codons ending in A and T for alanie, proline, serine and threonine. However, for arginine, glycine, leucine and valine, A is more frequent than C. Among twofold degenerate codons, C appears to be used more than T in pyrimidine codon family, whereas purine codon families end mostly with A. Except for arginine, G is the least common third position nucleotide in all codon families. All these features are very similar to those observed in vertebrates [19, 51].

Non-coding regions

The origin of the light strand replication (O_L) of *C. niveatus* is located in a cluster of five tRNA genes (WANCY) as in other vertebrates [29, 55]. This region is 57 bp long and has potential to fold into a stem-loop secondary structure. The folding of the O_L requires eight nucleotides of tRNA-Asn and 12 nucleotides of tRNA-Cys including the conserved sequence (5′-GCCGG-3′) at the base of the stem-loop structure.

The mitogenome of *C. niveatus* contains a non-coding region up to 799 bp (Table 3). The CR has a highter A + T content (63.1%) than the average value of the whole genome (52.6%) of *C. niveatus*, a feature that has been reported in all Percoidei fishes. Structurally, the CR is divided into three domains, including the termination associated sequence domain, the central conserved sequence block domain and the conserved sequence block domain [38, 39]. By comparing with the recognition sites in some reported Percoidei species, the conserved blocks ETAS and CSB-1, -2, and -3 can be easily identified in the control regions of *Collichthys* species, no CSB-F, -E and -D were detected (Fig. 1). The sequences corresponding to CSB-F, CSB-E and CSB-D could not be aligned with the respective sequences in the control region of *Miichthys miiuy* (HM447240) (see Fig. 2 in Supplementary Material) and *Cynoscion acoupa* [35] (see Fig. 3 in Supplementary Material). The lack of the typical sequences of the central

Table 2 List of species used in this study, with GenBank/EMBL/DDBJ accession numbers

Family	Species	Accession number
Sciaenidae	*Aplodinotus grunniens*	AY225662
	Aplodinotus grunniens	AY520093
	Argyrosomus regius	EF455987
	Atractoscion aequidens	DQ107826
	Atractoscion aequidens	DQ197926
	Atractoscion nobilis	GQ220018
	Atractoscion nobilis	GU440241
	Bairdiella armata	GQ220024
	Bairdiella ronchus	GQ220025
	Cheilotrema saturnum	GU440274
	Chrysochir aureus	EF607347
	Cilus gilberti	GU585937
	Collichthys lucida	HM447230
	Collichthys niveatus	HM219223
	Cynoscion acoupa	EU562597
	Cynoscion arenarius	AY857948
	Cynoscion arenarius	EU180144
	Cynoscion guatucupa	EU074399
	Cynoscion nebulosus	AY857949
	Cynoscion nothus	AY857950
	Cynoscion parvipinnis	DQ090069
	Cynoscion parvipinnis	GU440301
	Cynoscion regalis	AY857947
	Isopisthus remifer	GQ220017
	Johnius belangerii	FJ347918
	Johnius borneensis	FJ347920
	Johnius dussumieri	EF528208
	Johnius elongates	EF528213
	Larimichthys crocea	EU366949
	Larimichthys polyactis	GU586227
	Macrodon ancylodon	GQ220014
	Macrodon mordax	GQ220015
	Menticirrhus americanus	EU074465
	Menticirrhus undulatus	GU440404
	Micropogonias furnieri	EU074482
	Miichthys miiuy	HM447240
	Nebris microps	GQ220022
	Nebris occidentalis	GQ220021
	Nibea maculata	EU014250
	Otolithes cuvieri	FJ347925
	Otolithes cuvieri-1	EF051049
	Otolithes cuvieri-2	EF528228
	Otolithes cuvieri-3	EF528199
	Otolithes ruber	EF528218
	Otolithes ruber	EF607451
	Otolithoides biauritus	EF528222
	Otolithoides biauritus	EF534127
	Pennahia anea	EF607488
	Plagioscion squamosissimus	GQ220020
	Pogonias cromis	EU752167
	Protonibea diacanthus	EF528203

Table 2 continued

Family	Species	Accession number
	Protonibea diacanthus	EF528233
	Pseudotolithus elongatus	EF456017
	Pseudotolithus senegalensis	DQ197986
	Pseudotolithus senegallus	DQ197987
	Pseudotolithus typus	DQ197988
	Roncador stearnsii	EU752172
	Sciaenops ocellatus	EU752181
	Sciaenops ocellatus-1	AY857951
	Sciaenops ocellatus-2	FJ175399
	Stellifer illecebrosus	GQ220023
	Umbrina canariensis	EF392638
Heterodontidae	*Heterodontus francisci*	AJ310141
Petromyzonidae	*Lampetra fluviatilis*	Y18683
Percidae	Etheostoma radiosum	AY341348
Chaetodontidae	Chaetodon auripes	AP006004
Lutjanidae	Lutjanus rivulatus	AP006000
Pomacanthidae	Chaetodontoplus septentrionalis	AP006007
	Centropyge loriculus	AP006006

conserved sequence block was also confirmed in *L. crocea* and *L. polyactis* [6]. In the termination associated sequence domain, an ETAS was identified, the sequence of ETAS is TATATATATGTATTATCAAC ATACAATTATATTA ACCAT, whose motif sequence is TATAT with one palindormic sequence ATGTA. Moreover, another motif (TACAT) was detected at the downstream of ETAS. Although the central conserved sequence block is absence, however, *C. niveatus* dose contain a GTGGGG box which is a typical feature of CSB-E in teleosts. In the conserved sequence block, CSB-1, CSB-2 and CSB-3 of *C. niveatus* were found at the 3'-end of the control region, whose sequences are ATTTTAAGTATTCAAGTGCATAA, TAG ACCCCCCCCTACCCCCCCC and TAAAACCCCAT AAAACA, respectively (Fig. 1). CSB-1 is at the start of the conserved sequence block, and is relatively less conserved than CSB-2 and CSB-3.

Transfer and ribosomal RNA genes

The mitochondrial genome of *C. niveatus* encodes 22 tRNA genes, ranging from 67 to 75 bp, which can be fold into the typical clover-leaf secondary structures with several mismatch pairs (see Fig. 4 in Supplementary Material). Of these tRNAs, we identified two forms of tRNA-Leu (UUR and CUN) and tRNA-Ser (UCN and AGY) (see Fig. 1 in Supplementary Material; Table 3). The three tRNA clusters (IQM, WANCY, and HSL) are well conserved in *C. niveatus* as those of typical vertebrate mitogenomes.

Although putative gene boundaries for the two rRNA genes in the mitogenome have been found, these cannot be accurately determined until transcript mapping is carried out. As in other vertebrate mitogenomes, these genes are located between tRNA-Phe and tRNA-Val and between tRNA-Val and tRNA-Leu(UUR). Preliminary assessment of their secondary structure indicated that the sequence could be reasonably superimposed on the proposed secondary structure of carp 12S RNA and loch 16S rRNA, respectively. The lengths of 12S rRNA gene and 16S rRNA gene are 949 and 1698 bp, respectively. The base composition of the two rRNAs gene sequence is T: 26.2%, C: 22.6%, A: 28.1%, G: 23.1%. The overall A + T contents of ribosomal RNAs being 54.3%, which is slightly A + T rich than other bony fishes [21].

Phylogenetic analysis of Sciaenidae and the position of *C. niveatu*

Phylogenetic analysis of 16S rRNA using NJ, ME, and ML methods revealed three distinct clades (Fig. 2). Clade I included *Johinus* (*Johinus dussumieri* and *Johinus elongatus*) distant from the other Sciaenidae fishes. Clade II included *Otolithes* (*Otolithes ruber* and *Otolithes cuvieri*), which was more closely related to *Collichthys*, *Larimichthys* and *Miichthys* were placed in clade III. Although *Johinus belangrii* and *Johinus borneensis* were not formed into a clade in the analyses based on COI gene, however, within Sciaenidae, *Johinus* was strongly supported as the most basal genus in this study and *Otolithes* was also more closely related to *Collichthys*, *Larimichthys* and *Miichthys* than *Johinus* with high bootstrap value (Fig. 3).

According to the morphological characters, *Collichthys*, *Larimichthys* and *Miichthys* were grouped as an independent subfamily Pseudosciaeniae; however, the

Table 3 Organization of the mitochondrial genome of *C. niveatus*

Gene	Position		Size (bp)		Codon		Intergenic nucleotide[b]	Strand
	From	To	Nucleotide	Amino acid	Initiation	Stop[a]		
tRNA-Phe	1	69	69					H
12S rRNA	70	1018	949					H
tRNA-Val	1019	1089	71				1	H
16S rRNA	1091	2788	1698				1	H
tRNA-Leu(UUR)	2790	2863	74					H
ND1	2864	3838	975	324	ATG	TAG	4	H
Trna-Ile	3843	3912	70				−1	H
tRNA-Gln	3912	3983	72				−1	L
tRNA-Met	3983	4052	70					H
ND2	4053	5098	1046	348	ATG	TA-	0	H
tRNA-Trp	5099	5169	71				1	H
tRNA-Ala	5171	5239	69				1	L
tRNA-Asn	5242	5314	73				37	L
tRNA-Cys	5352	5418	67				−1	L
tRNA-Tyr	5418	5487	70				1	L
COI	5489	7045	1557	518	ATG	AGA	−5	H
tRNA-Ser(UCN)	7041	7112	72				3	L
tRNA-Asp	7116	7184	69				8	H
COII	7193	7883	691	230	ATG	T–		H
tRNA-Lys	7884	7958	75				1	H
ATPase8	7960	8127	168	55	ATG	TAA	−10	H
ATPase6	8118	8801	684	227	ATG	TAA	−1	H
COIII	8801	9585	785	261	ATG	TA-		H
tRNA-Gly	9586	9656	71					H
ND3	9657	10005	349	116	ATG	T–		H
tRNA-Arg	10006	10074	69					H
ND4L	10075	10371	297	98	ATG	TAA	−7	H
ND4	10365	11745	1381	460	ATG	T–		H
tRNA-His	11746	11814	69					H
tRNA-Ser(AGY)	11815	11881	67				5	H
tRNA-Leu(CUN)	11887	11958	72					H
ND5	11959	13797	1839	612	ATG	TAA	−4	H
ND6	13794	14312	519	172	ATG	TAA		L
tRNA-Glu	14313	14381	69				4	L
Cytb	14386	15526	1141	380	ATG	T–		H
tRNA-Thr	15527	15598	72				2	H
tRNA-Pro	15601	15670	70					L
Control Region	15671	16469	799					H

[a] TA- and T-represent incomplete stop codons

[b] Numbers correspond to the nucleotides separating adjacent genes. Negative numbers indicate overlapping nucleotides

monophyly of Pseudosciaeniae was not supported by this study. *M. miiuy* was placed in a clade grouping *Nibea maculata*, *Chrysochir aureus*, *Protonibea diacanthus* and *Pennahia anea* while *Otolithoides biauritus* was grouped with *Collichthys* and *Larimichthys* (Fig. 3). Analyses based on 16S rRNA gene also showed a non-monophyletic Pseudosciaeniae, where *M. miiuy* formed an independent clade (Fig. 2). In contrast to these findings from COI and 16S rRNA analyses, the monophyly of Pseudosciaeniae was supported in the analyses based on partial Cytb gene, however, the bootstrap value is very poor (Fig. 4).

Table 4 Base composition of *C. niveatus* mitochondrial genome

Gene/region	Base composition (%)				
	T	C	A	G	A + T
Protein coding					
1st	20.0	29.0	25.9	25.3	45.9
2nd	40.0	27.8	18.2	13.8	58.2
3rd	21.0	40.7	30.3	8.5	51.3
Total	26.9	32.5	24.8	15.9	51.7
tRNA	26.2	22.6	28.1	23.1	54.3
sRNA	20.5	26.7	32.6	20.2	53.1
Control region	30.4	22.9	32.7	14.0	63.1
Overall	25.0	31.3	27.6	16.2	52.6

Relationships of taxa of *Collichthys* and *Larimichthys* derived from NJ, ME, and ML analyses of the COI, 16S rRNA and Cytb sequences were identical. *C. niveatus* is found to be most closely related to *L. polyactis*, the two species were grouped with *L. crocea* and then grouped with *C. lucida*. This result was against with morphological affiliations, *Larimichthys* and *Collichthys* were not supported.

Discussion

General features of mitogenomes of *C. niveatus* and its coordinal species

The complete mitochondrial genome of *C. niveatus* is the third to be reported for a member of the family Sciaenidae. The mitochondrial genome is 16469 bp in length, and consisted of 37 genes (13 protein-coding genes, 2 rRNAs and 22 tRNAs), which is nearly identical to *L. crocea* and *L. polyactis*, and longer than *Collichthys lucida* but shorter than *M. miiuy* (Table 6). This length variation of mitogenomes in these species is largely due to the number and size of non-coding spacer and length of main non-coding regions. The mitochondrial genome of *C. niveatus* has an overall 52.6% A + T content, identical to the value in *L. polyactis* (GU586227), but lower than *L. crocea* and *C. lucida* [6]; NC_014350). As in other vertebrates, most of genes of *C. niveatus* are encoded on the H-strand, with only the ND6, ND8 and eight tRNA genes encoded on the L-strand. In addition, all genes are nearly identical to those of other Sciaenidae species in length (Table 3). The length of protein-coding region in the mitochondrial genomes of *C. niveatus* and its coordinal species is also nearly identical with only several base differences (Table 6), and the presence of incomplete stop codon resulted in these variations. All the initiation codons have been identified as ATG in these species and some of the stop codons are incomplete, with TA or T. Such incomplete stop codons are common among fish mitogenomes. However, the stop codons of these species are TAG and AGA in ND1 and COI respectively, while the corresponding ones in other fishes are often TAA [33, 40]. The mitochondrial genome of *C. niveatus* contains 22 tRNA genes, which are interspersed between the rRNA and protein-coding genes (Table 3). Length of tRNA genes in these species all varied from 67 bp ($tRNA^{Cys}$ and $tRNA^{Ser}$(AGY)) to 75 bp ($tRNA^{Lys}$). All tRNA genes were predicted to have the typical cloverleaf structures except the $tRNA^{Ser}$(AGY) showing the deviated secondary structure [3, 12, 13]. They harbor identical anticodons used in other vertebrates, and also conserved aminoacyl, DHU (dihydrouridine), anticodon and TΨC (thymidine-pseudouridine-cytidine) stems

Table 5 Condon usage in *C. niveatus* mitochondrial protein-coding genes

Amino acid	Codon	Number	Amino acid	Codon	Number
Phe	TTT	104	Tyr	TAT	47
	TTC	134		TAC	61
Leu	TTA	56	Stop	TAA	5
	TTG	19		TAG	1
	CTT	125	His	CAT	28
	CTC	224		CAC	79
	CTA	182	Gln	CAA	89
	CTG	69		CAG	9
Ile	ATT	124	Asn	AAT	32
	ATC	145		AAC	87
Met	ATA	96	Lys	AAA	62
	ATG	61		AAG	9
Val	GTT	58	Asp	GAT	22
	GTC	58		GAC	55
	GTA	52	Glu	GAA	75
	GTG	29		GAG	27
Ser	TCT	32	Cys	TGT	8
	TCC	66		TGC	22
	TCA	67	Trp	TGA	101
	TCG	8		TGG	20
Pro	CCT	46	Arg	CGT	8
	CCC	129		CGC	18
	CCA	44		CGA	42
	CCG	4		CGG	10
Thr	ACT	54	Ser	AGT	10
	ACC	140		AGC	42
	ACA	111	Stop	AGA	1
	ACG	13		AGG	0
Ala	GCT	56	Gly	GGT	28
	GCC	182		GGC	110
	GCA	93		GGA	72
	GCG	14		GGG	33

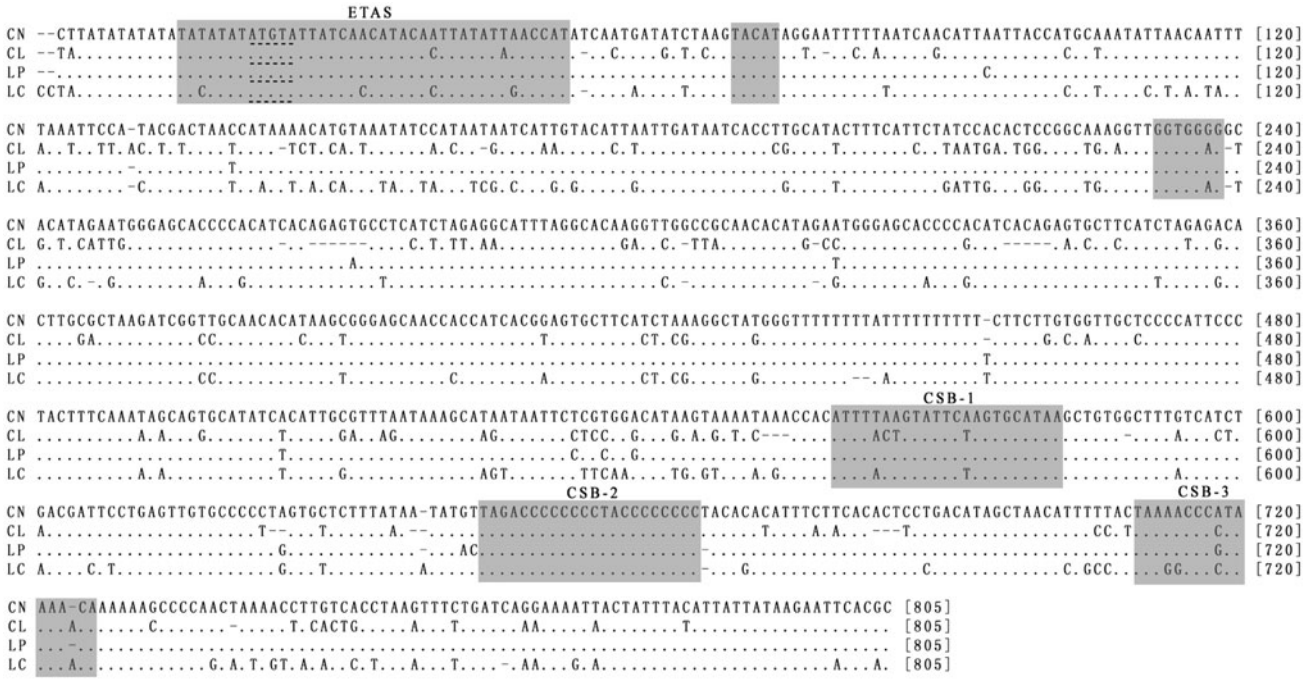

Fig. 1 Alignment of complete sequences of the mtDNA control regions of *C. niveatus*, *C. lucida*, *L. crocea* and *L. polyactis* (represented by CN, CL, LC, and LP, respectively). The ETAS, CSB-1, CSB-2, and CSB-3 are *shadowed* and *marked*

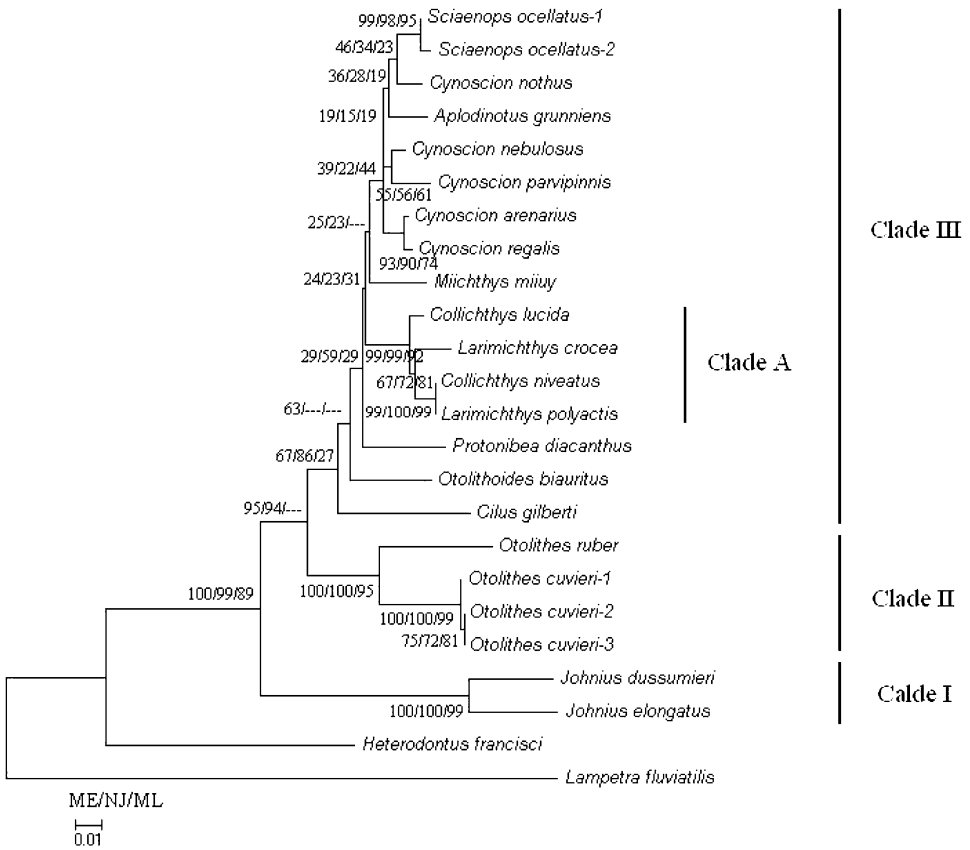

Fig. 2 Phylogenetic tree of the Sciaenidae based on partial 16S rRNA gene. Branch lengths and topology are from the Minimum Evolution analysis. *Numbers* above branches specify bootstrap percentages for ME (1000 replications), NJ (1000 replications), and ML (1000 replications) analyses. *Heterodontus francisci* and *Lampetra fluviatilis*, belonging to the order Heterodontiformes and Petromyzoniformes, respectively, were used as the outgroup taxons

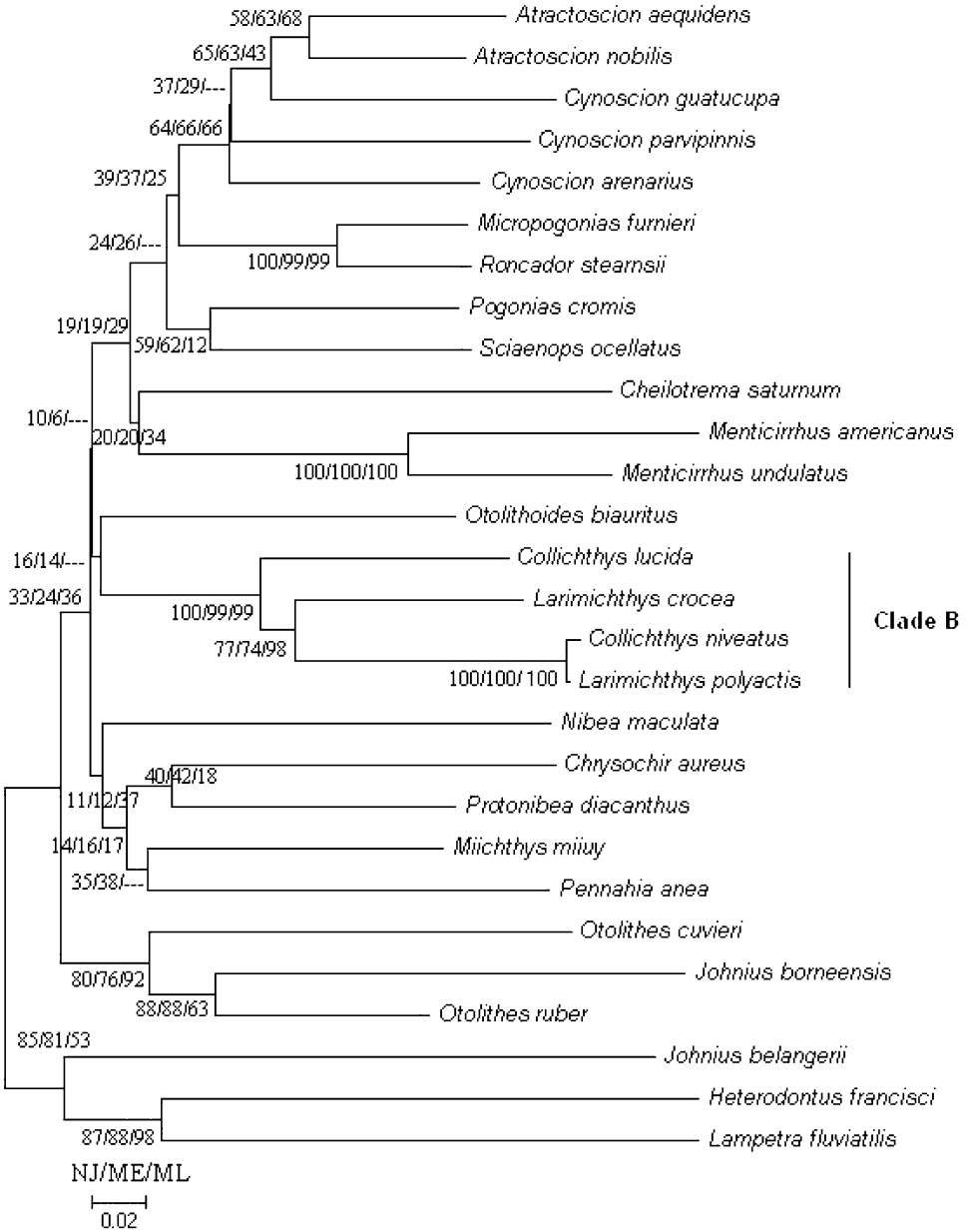

Fig. 3 Phylogenetic tree of the Sciaenidae based on partial COI gene. Branch lengths and topology are from the Neighbor Joining analysis. *Numbers* above branches specify bootstrap percentages for NJ (1000 replications), ME (1000 replications), and ML (1000 replications) analyses. *H. francisci* and *L. fluviatilis* were used as outgroups. Bootstrap values are given for each branch

(see Fig. 4 in Supplementary Material). The tRNA-Ser(AGY) found in the *C. niveatus* mitogenome had no discernible DHU stem, similar to that shown in the lamprey [19], bichir [28], and rock bream [29]. The 12S and 16S rRNA genes of *C. niveatus*, *C. lucida, L. crocea*, *L. polyactis*, and *M. miiuy* are 949 bp/1698 bp, 947 bp/1696 bp, 947 bp/1693 bp, 950 bp/1697 bp, and 946 bp/1693 bp, respectively. The 12S rRNA genes in these species are all similar in size to its counterparts in *Etheostoma radiosum* (AY341348), *Chaetodon auripes* (AP006004), and *Lutjanus rivulatus* (AP006000). However, the 16S rRNA genes of *C. niveatus*, and its coordinal species are much shorter than those of *Chaetodontoplus septentrionalis* (AP006007) and *Centropyge loriculus* (AP006006), members of Pomacanthidae. All these molecular features from our study revealed marked similarities among these species.

The absence of the typical central conserved sequence block in the control region of *Collichthys*

Mitochondrial control region included the promoters for both strands, the heavy strand replication origin, and the displacement region [5]. It is also a unique and highly variable area in the mitochondrial genome noted for its non-protein coding and a faster rate of evolution. Southern et al. [39] first recognized the conserved sequences CSB-B, CSB-C, CSB-D, CSB-E and CSB-F in the central conserved sequence block domain in mammals. However, only

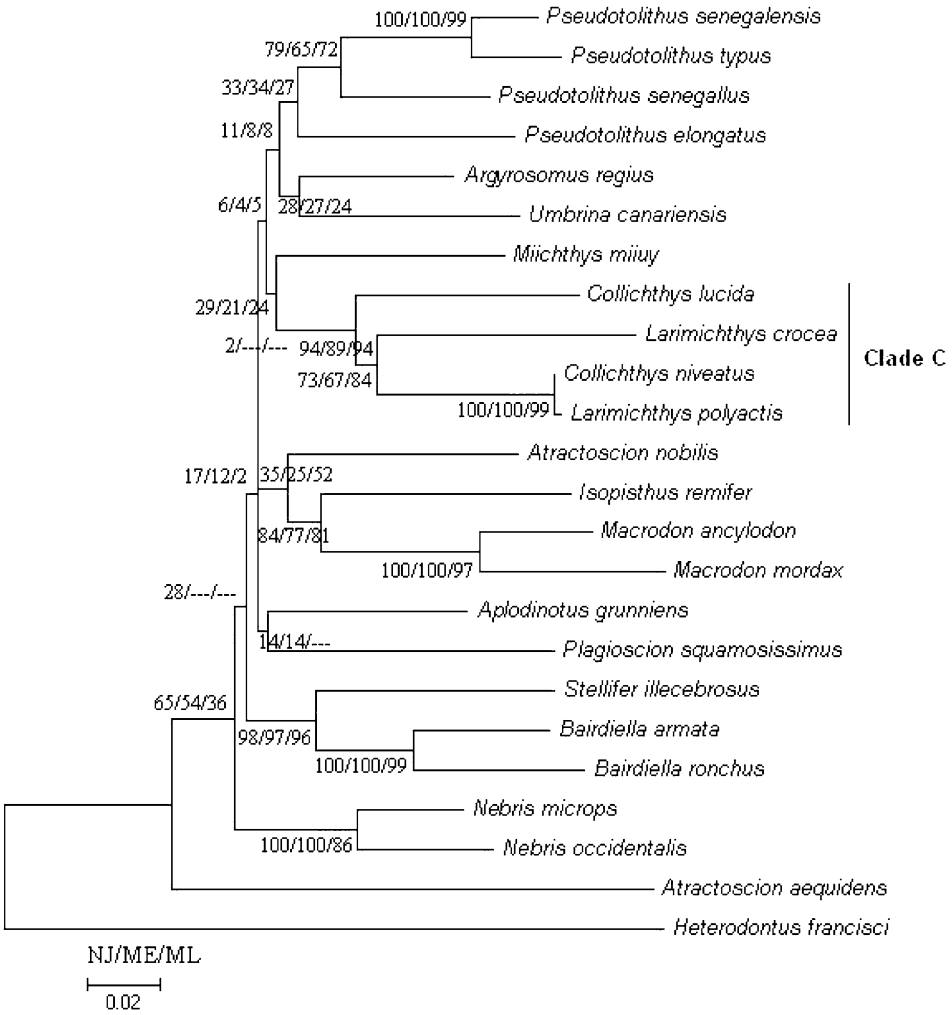

Fig. 4 Phylogenetic tree of the Sciaenidae based on partial Cytb gene. Branch lengths and topology are from the Neighbor Joining analysis. *Numbers* above branches specify bootstrap percentages for NJ (1000 replications), ME (1000 replications), and ML (1000 replications) analyses. *H. francisci* were used as an outgroup taxon

Table 6 Mitochondrial genomes in Sciaenidae reported

Species (accession number)	Size (bp)	A + T (%)	sRNA (12S + 16S)		Control region		13 protein-coding genes A + T (%)		
			Length (bp)	A + T (%)	Length	A + T (%)	1st	2nd	3rd
Larimichthys crocea (EU339149)	16466	53	947/1693	53.4	795	61.3	46.1	58.5	51.7
Larimichthys polyactis (NC_013754)	16470	52.6	950/1697	53.1	799	62.5	45.9	58.2	51.3
Collichthys niveatus (HM219223)	16469	52.6	949/1698	53.1	799	63.1	45.9	58.2	51.3
Collichthys lucida (NC_014350)	16442	53.6	947/1696	53.5	771	62.7	46.8	58.3	53.8
Miichthys miiuy (NC_014351)	16493	51.9	946/1693	52.8	845	63.3	45.8	58.2	47.7

CSB-F, CSB-E and CSB-D could be identified in fishes [2, 56, 58]. CSB-F is the mark to differentiate the central conserved sequence block domain from the termination associated sequence domain. CSB-D is highly conserved in fish and may function in the regulation of H-strand replication and the initiation of the D-loop structure and perhaps be involved in mitochondrial metabolism [10, 32]. Although termination associated sequence domain accumulated base substitutions, insertions and deletions at a substantially higher rate than the central domain, however, absence of several conserved motifs can also be observed in fishes [16, 31]. Cui et al. [6] reported the lack of the central domains observed in *L. crocea*, where the conserved blocks TAS and CSB-1, CSB-2, and CSB-3 were easily recognized with no CSB-F, CSB-E, and CSB-D. The same phenomenon was further confirmed in its congeneric species *L. polyactis*. In this study, the lack of the central domain were also identified in *Collichthys* species,

however, we found that the consensus sequences of CSB-F, CSB-E, and CSB-D with ATGTAGTA——GAGAC-CACC, AGGG———GTGGGG, and TTAT-CT-GG-AT CTG-T-AA, respectively, typically present in bagridae [58] can be identified in *M. miiuy* (see Fig. 2 in Supplementary Material) and *C. acoupa* (see Fig. 3 in Supplementary Material), and also in other Percoidei fishes [33]. Such variation imply a rapid evolution of the structure in the control region, which may provide information for elucidating the evolutionary origin of *Collichthys* and *Larimichthys* within the family Sciaenidae and it cannot be ruled out that the control region of these two species may not be functional due to the lack of central domains.

Phylogenetic relationships of Sciaenidae and the position of *C. niveatus*

Phylogenetic trees obtained using different methods based on the same data set are nearly identical, even though they are somewhat difference in detail and bootstrap value. Based on our analyses, *Johnius* was found to be distantly related to other Sciaenidae fishes, which was consistent with previous studies using different phylogenetic methods and also agree with the traditional morphological classification. Chen [4] proposed several NJ trees based on 16S rRNA sequence data using different methods that resulted in an independent clade composed of *Miichthys*, *Collichthys* and *Larimichthys*, which supported Zhu et al. [59]. However, the bootstrap values for this topology were extremely poor. Cui et al. [6] employed partial 16S rRNA and Cytb genes to build a phylogeny of 11 Sciaenidae fishes, where the support for the clade grouping *Miichthys*, *Collichthys* and *Larimichthys* was also poor. Phylogenies proposed in this study based on partial COI and 16S rRNA genes respectively showed *Miichthys* cannot be merged into the *Collichthys- Larimichthys* clade, on the other hand, the phylogeny based on Cytb placed the *Miichthys*, *Collichthys* and *Larimichthys* together to form an independent clade, which was consistent with the previous studies, however, this clade appeared ambiguous for having poor bootstrap value, being 29/21/24 (Fig. 4). The monophyly of Pseudosciaeniae was not supported and the relationship between *Miichthys* and *Collichthys- Larimichthys* clade deserve to be further studied.

Collichthys and *Larimichthys* were genera of Sciaenidae according to morphological researches [59]. In recent years, there have been several phylogenetic studies based on molecular data [4, 23, 45, 46], however, the results revealed unstable phylogenies in which there were some disagreements in the limit of the relationships within the members of *Collichthys* and *Larimichthys*. All phylogenetic trees proposed by this study produced completely identical and well-supported *Collichthys-Larimichthys* clades (Figs. 2, 3, and 4). In these clades, *C. niveatus* is found to be most closely related to *L. polyactis*, the two species were grouped with *L. crocea* and then grouped with *C. lucida*. *Collichthys* and *Larimichthys* should be merged to one genus by our phylogenetic analyses.

Our results is against with the traditional classification and the proposed phylogenetic position of *C. niveatus* within the Sciaenidae and the relationships among Sciaenidae species based on the findings of the present study should be accepted with caution, complete understanding of Sciaenidae relationships awaits assembly of additional DNA sequence data (e.g. whole mitochondrial genomes and multiple nuclear loci) and toxon sampling, and corroborating morphological evidence.

Acknowledgments This work was supported by Open Foundation from Ocean Fishery Science and Technology in the Most Important Subjects of Zhejiang (20100209, 20100118).

References

1. Amnuay J, Pradit S, Rafael Z (2007) The complete mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*), and the phylogenetic relationships among Siluriformes. Gene 387:49–57
2. Broughton RE, Dowling TE (1994) Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics 138:179–190
3. Chang YS, Huang FL, Lo TB (1994) The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. J Mol Evol 38:138–155
4. Chen QM (2007) Molecular phylogeny of the Sciaenidae in China. Jinan University, GuangZhou (in Chinese with an English abstract)
5. Clayton DA (1982) Replication of animal mitochondrial DNA. Cell 28:693–705
6. Cui ZX, Liu Y, Li CP, You F, Chu KH (2009) The complete mitochondrial genome of the large yellow croaker, *Larimichthys cracea* (Perciformes, Sciaenidae): unusual features of its control region and the phylogenetic position of the Sciaenidae. Gene 432:33–43
7. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791
8. FishBase (2000) http://www.fishbase.org/home.htm. Accessed in May
9. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704
10. Guo XH, Liu SJ, Liu Y (2003) Comparative analysis of the mitochondrial DNA control region in cyprinids with different ploidy level. Aquaculture 224:25–38
11. Habib M, Lakra WS, Mohindra V, Khare P, Barman AS, Singh A, Lal KK, Punia P, Khan AA (2010) Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of Channa marulius (Channidae: Perciformes). Mol Biol Rep. doi:10.1007/s11033-010-0175-2
12. Hu J, Zhang DX, Hao JS, Huang DY, Cameron S, Zhu CD (2010) The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini):

sequence, gene organization and a unique tRNA translocation event. Mol Biol Rep 37:3431–3438
13. Hurst CD, Bartlett SE, Davidson WS, Bruce IJ (1999) The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar*. Gene 239:237–242
14. Johansen S, Guddal PH, Johansen T (1990) Organization of the mitochondrial genome of Atlantic cod, Gadus morhua. Nucleic Acids Res 18:411–419
15. Kim IC, Kweon HS, Kim YJ, Kim CB, Gye MC, Lee WO, Lee YS, Lee JS (2004) The complete mitochondrial genome of the javeline goby (*Acanthogobius hasta*) (Perciformes, Gobiidae) and phylogenetic considerations. Gene 336:147–153
16. Kim IC, Lee JS (2004) The Complete Mitochondrial Genome of the Rockfish *Sebastes schlegeli* (Scorpaeniformes, Scorpaenidae). Mol Cells 17:322–328
17. Lavoue S, Miya M, Inoue JG, Saitoh K, Ishiguro NB, Nishida M (2005) Molecular systematics of the gonorynchiform fishes (Teleostei) based on whole mitogenome sequences: implications for higher-level relationships within the Otocephala. Mol Phylogenet Evol 37:165–177
18. Lakra WS, Goswami M, Gopalakrishnan A (2009) Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: Perciformes, Sciaenidae) based on 16S rRNA and cytochrome c oxidase subunit I mitochondrial genes. Mol Biol Rep 36:831–839
19. Lee WJ, Kocher TD (1995) Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. Genetics 139: 873–887
20. Li MD, Zhang SJ, Zhang YZ (2003) A catalogae of Chinese fishes Percoidei (in Part). Mar Sci Bull 22:45–54
21. Liu Y, Cui ZX (2009) The complete mitochondrial genome sequence of the cutlassfish *Trichiurus japonicus* (Perciformes: Sciaenidae): genome characterization and phylogenetic considerations. Mar Genom 2:133–142
22. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genome sequence. Nucleic Acids Res 25:955–964
23. Meng ZN, Zhuang ZP, Ding SX, Jin XS, Su YQ, Tang QS (2004) Molecular phylogeny of eight Sciaenid species (Perciformes: Sciaenidae) in the China Sea based on mitochondrial 16S rRNA sequence. Prog Nat Sci 14:514–521 (in Chinese with an English abstract)
24. Miya M, Takeshima H, Endo H, Ishiguro NB, Inoue JG, Mukai T, Satoh TP, Yamaguchi M, Kawaguchi A, Mabuchi K, Shirai SM, Nishida M (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Mol Phylogenet Evol 26:121–138
25. Miya M, Kawaguchi A, Nishida M (2001) Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. Mol Biol Evol 18: 1993–2009
26. Naylor GJ, Collins TM, Brown WM (1995) Hydrophobicity and phylogeny. Nature 373:555–556
27. Nelson JS (2006) Fishes of the World, 4th edn. Wiley, New York
28. Noack K, Zardoya R, Meyer A (1996) The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: ancient establishment of the consensus vertebrate gene order. Genetics 144:1165–1180
29. Oh DJ, Kim JY, Lee JA, Yoon WJ, Park SY, Jung YH (2007) Complete mitochondrial genome of the rock bream *Oplegnathus fasciatus* (Perciformes, Oplegnathidae) with phylogenetic considerations. Gene 392:174–180
30. Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474
31. Peng Z, Wang J, He S (2006) The complete mitochondrial genome of the helment catfish *Cranoglanis bouderius* (Siluriformes: Cranoglanididae) and the phylogeny of otophysan fishes. Gene 376:290–297
32. Pesole G, Gissi C, Chirico DE, Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol 48:427–434
33. Ponce M, Infante C, Jiménez-Cantizano RM, Pérez L, Manchado M (2008) Complete mitochondrial genome of the blackspot seabream, *Pagellus bogaraveo* (Perciformes: Sparidae), with high levels of length heteroplasmy in theWANCY region. Gene 409:44–52
34. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818
35. Rodrigues R, Schneider H, Santos S, Vallinoto M, Saint-Paul U, Sampaio I (2008) Low levels of genetic diversity depicted from mitochondrial DNA sequences in a heavily exploited marine fish (*Cynoscion acoupa*, Sciaenidae) from the northern coast of Brazil. Genet Mol Biol 38:487–492
36. Saitoh K, Miya M, Inoue JG, Ishiguro NB, Nishida M (2003) Mitochondrial genomics of ostariophysan fish: perspectives on phylogeny and biogeography. J Mol Evol 56: 464–472
37. Sambrook J, Russell DW (eds) (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New York, NY
38. Sbisa E, Tanzariello F, Reyes F, Pesole G, Saccone C (1997) Mammalian mitochondrial D-loop region structure analysis: identification of new conserved sequences and the functional and evolutional implications. Gene 205:125–140
39. Southern SO, Southern PJ, Dizon AE (1988) Molecular characterization of a cloned dolphin mitochondrial genome. J Mol Evol 28:32–40
40. Takashima Y, Morita T, Yamashita M (2006) Complete mitochondrial DNA sequence of Atlantic horse mackerel *Trachurus trachurus* and molecular identification of two commercially important species *T. trachurus* and *T. japonicus* using PCR-RFLP. Fish Sci 72:1054–1065
41. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Bio Evol 24:1596–1599
42. Taniguchi N (1963) Comparative osteology of the sciaenid fishes from Japan and its adjacent waters-II. Vertebrae. Jpn J Ichthyol 16:153–156
43. Taniguchi N (1969) Comparative osteology of the sciaenid fishes from Japan and its adjacent waters-I. Neurocranium. Jpn J Ichthyol 16:55–67
44. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882
45. Tian LX, Liang B, Zhang SY, Zhao SL, Wang RX (2004) Phylogenetic relationships of 7 Sciaenidae species based on cytochrome b gene sequences. J Oceanogr Taiwan Strait 23:436–443 (in Chinese with an English abstract)
46. Tong X, Du B, Yu DH, Gong SY, Guo YH, Huang GJ, Li LH (2007) Sequence analysis of mitochondrial 16S rRNA gene fragment in Chups croaker (*Nibea coibor*). Mar Fish Res 28:85–91 (in Chinese with an English abstract)
47. Tzeng CS, Hui CF, Shen SC, Huang PC (1992) The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variation among vertebrates. Nucleic Acids Res 20:4853–4858
48. Waldbieser GC, Bilodeau AL, Nonneman DJ (2003) Complete sequence and characterization of the channel catfish mitochondrial genome. DNA Seq 14:265–277

49. Wan JR, Sun S (2006) The category composition and abundance of ichthyoplankton in the ecosystem of the Yellow Sea and the East China Sea. Acta Zoologica Sinica 52:28–44
50. Wang XZ, Wang J, He SP, Mayden RL (2007) The complete mitochondrial genome of the Chinese hook snout carp *Opsariichthys bidens* (Actinopterygii: Cypriniformes) and an alternative pattern of mitogenomic evolution in vertebrate. Gene 399:11–19
51. Xia J, Xia K, Gong J, Jiang S (2007) Complete mitochondrial DNA sequence, gene organization and genetic variation of control regions in *Parargyrops edita*. Fish Sci 73:1042–1049
52. Xue Y, Jin XS, Zhang B, Liang ZL (2005) Feeding habits of three sciaenid fishes in the southern Yellow Sea. J Fish China 2:178–187 (in Chinese with an English abstract)
53. Yamanoue Y, Miya M, Matsuura K, Yagishita N, Mabuchi K, Sakai H, Katoh M, Nishida M (2007) Phylogenetic position of tetraodontiform fishes within the higher teleosts: Bayesian inferences based on 44 whole mitochondrial genome sequences. Mol Phylogenet Evol 45:89–101
54. Yang R, Wu XB, Yan P, Su X, Yang BH (2010) Complete mitochondrial genome of *Otis tarda* (Gruiformes; Otididae) and phylogeny of Gruiformes inferred from mitochondrial DNA sequences. Mol Biol Rep 37:3057–3066
55. Zardoya R, Garrido-Pertierra A, Bautista JM (1995) The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. J Mol Evol 41:942–951
56. Zeng QL, Liu HZ (2001) Study on mitochondrial DNA control region of the *Ictiobus cypriellus*. J Hubei Univ (Natural Science Edition) 23:261–264 (in Chinese with an English abstract)
57. Zhang P, Zhou H, Liang D, Liu YF, Chen YQ, Qu LH (2005) The complete mitochondrial genome of a tree frog, *Polypedates megacephalus* (Amphibia: Rhacophoridae), and a novel gene organization in living amphibians. Gene 346:133–143
58. Zhang Y, Zhang E, He SP (2003) Studies on the structure of the control region of the Bagridae in China and its phylogenetic significance. Acta Hydrobiol Sin 27:463–467 (in Chinese with an English abstract)
59. Zhu YD, Lo YL, Wu HL (1963) A study on the classification of the Sciaenoid fishes of China, with description of new genera and species. Shanghai Science and Technology Press, China