RESEARCH ARTICLE

Complete sequence of the mitochondrial genome of *Odontamblyopus rubicundus* (Perciformes: Gobiidae): genome characterization and phylogenetic analysis

TIANXING LIU, XIAOXIAO JIN, RIXIN WANG* and TIANJUN XU*

Laboratory for Marine Living Resources and Molecular Engineering, College of Marine Science, Zhejiang Ocean University, Zhoushan 316000, People's Republic of China

Abstract

Odontamblyopus rubicundus is a species of gobiid fishes, inhabits muddy-bottomed coastal waters. In this paper, the first complete mitochondrial genome sequence of *O. rubicundus* is reported. The complete mitochondrial genome sequence is 17119 bp in length and contains 13 protein-coding genes, two rRNA genes, 22 tRNA genes, a control region and an L-strand origin as in other teleosts. Most mitochondrial genes are encoded on H-strand except for *ND6* and seven tRNA genes. Some overlaps occur in protein-coding genes and tRNAs ranging from 1 to 7 bp. The possibly nonfunctional L-strand origin folded into a typical stem-loop secondary structure and a conserved motif (5′-GCCGG-3′) was found at the base of the stem within the $tRNA^{Cys}$ gene. The TAS, CSB-2 and CSB-3 could be detected in the control region. However, in contrast to most of other fishes, the central conserved sequence block domain and the CSB-1 could not be recognized in *O. rubicundus*, which is consistent with *Acanthogobius hasta* (Gobiidae). In addition, phylogenetic analyses based on different sequences of species of Gobiidae and different methods showed that the classification of *O. rubicundus* into *Odontamblyopus* due to morphology is debatable.

[Liu T., Jin X., Wang R. and Xu T. 2013 Complete sequence of the mitochondrial genome of *Odontamblyopus rubicundus* (Perciformes: Gobiidae): genome characterization and phylogenetic analysis. *J. Genet*. **92**, 423–432]

Introduction

In general, the typical genome organization of animal mitochondrial DNA (mtDNA) is a closed double-stranded circular molecule ranging in size from ~16 to 18 kb, and encoding 37 genes: 13 protein-coding genes (*CO1*, *CO2*, *CO3*, *Cytb*, *ND1*, *ND2*, *ND3*, *ND4*, *ND5*, *ND6*, *ND4L*, *ATPase6* and *ATPase8*), two ribosomal RNA genes (small and large ribosomal genes), 22 transfer RNA genes, and a large noncoding region commonly related to the initiation of transcription and replication, namely, the control region (CR) (Wolstenholme 1992; Boore 1999). Mitochondrial genomes have played an important role in studies of population genetics and reconstruction of phylogeny, on account of their quick evolutionary rate, maternal inheritance, high information content and lack of recombination (Avise 2000; Boore *et al*. 2005). So far, mitochondrial genome sequence and structure analysis have become a powerful approach for studying molecular evolution and phylogenetic relationships (Xu *et al*. 2011).

When compared with other teleost groups, Gobiidae is one of the largest families of fishes with more than 2000 species, in 210 genera, living in marine, freshwater and blackish habitats (Nelson 2006; Zander 2011). Most of adult gobiid fishes are relatively small, typically less than 10 cm. In addition, their economic value is low and hence they are not much studied. To date, most of the studies about Gobiidae mainly concentrated in the reproductive biology, early life-history and species survey, etc. The researches of molecular aspects are very few, especially in molecular evolution and phylogenetic relationships. Moreover, the phylogenetic relationships of gobiid fishes are still controversial. Giinthe (1861) divided Gobiidae into Gobiinae, Amblyopinae, Trypaucheninae and Callionymina according to the number of fins and tandem scales. Bleeker (1872) redefined four subfamily of Gobiidae, based on the analysis of the teeth, bones and the position of the eyes. Gosline (1995) considered that it was difficult to separate Eleotridae and Gobiidae based on morphology. In addition, morphological characteristics of gobiid fishes are

*For correspondence. E-mail: Rixin Wang, wangrixin1123@126.com; Tianjun Xu, tianjunxu@163.com.

Keywords. mitochondrial genome; phylogenetic analysis; Gobiidae; Bayesian analysis; *Odontamblyopus rubicundus*.

unconspicuous as a result of small body. Therefore, there are a lot of controversies in Gobiidae classification and it is desirable to adopt molecular methods to resolve these issues.

In this study, we present the first complete nucleotide sequence for the mitochondrial genome of the *O. rubicundus* and determined its mitochondrial genomic structure. This is the first species in *Odontamblyopus* for which the complete mitochondrial genome has been determined. Finally, we conducted phylogenetic analysis based on the mitochondrial sequence data with the main aim of investigating the phylogenetic position of *O. rubicundus* within family Gobiidae.

Materials and methods

Sample collection and DNA extraction

The specimens of *O. rubicundus* were collected from the Zhoushan (Zhejiang Province, China). A portion of the dorsal fin was sampled and preserved in 95% ethanol. Total DNA was extracted using the standard phenol–chloroform procedure as described previously.

PCR amplifications and sequencing

To obtain the complete mitochondrial genome of *O. rubicundus*, we conducted two-step PCR amplification. First, the long-PCR amplification was used to acquire two long fragments, which covered the whole mtDNA of *O. rubicundus*. Twenty-five μL of the long PCR mixture contained 5 μL 5× LA buffer (New England Biolabs, Beverly, USA), 3 μL dNTP (2.5 mM) and 1 μL (10 μM) of the forward and reverse primers (Miya and Nishida 2000; Kawaguchi *et al*. 2001), respectively, 1 μL of LA *Taq* DNA polymerase (New England Biolabs, Beverly, USA), 1 μL of the DNA template and 13 μL of sterile distilled H_2O. The PCR was performed in a PTC-200 with pre-denaturation at 94°C for 1 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 65°C for 10 min and a final extension at 65°C for 10 min. The long PCR products were purified using the DNA Gel Extraction Kit (Tiangen, Shanghai, China). The purified products obtained in the first step, with two pairs of homologous primers (table 1) and others (Miya and Nishida 1999) were used in the second PCR amplification. The PCR reactions were conducted in 25 μL reaction mixtures containing 2.5 μL 10× *Taq* Plus polymerase buffer (Tiangen), 2 μL dNTP (2.5 mM) and 1 μL of the forward and reverse primers (10 μM), respectively, 0.2 μL *Taq* DNA polymerase (5 U/μL) (Tiangen), 1 μL of the DNA template and 17.3 μL of sterile distilled H_2O. The PCR reaction consisted of pre-denaturation at 94°C for 5 min; 35 cycles of denaturation at 94°C for 30 s annealing at 56–60°C for 30 s and final extension at 72°C for 1–2 min. All of the PCR products were sequenced on an ABI3730xl DNA Analyzer (Applied Biosystems, Foster City, USA) by primer walking.

Table 1. Primers used for PCR amplification of the mtDNA of *O. rubicundus.*

Primer	Sequence 5′–3′
1-F	AAACCAAGGATAAGAAAGGAGC
1-R	TACATAAGGGACAGCAGAGAGG
2-F	TCTTTGCTCCTAACTACCT
2-R	TAAATCTTACCACCAATCG

Table 2. List I of species from Gobiidae used in phylogenetic analyses, with GenBank accession numbers.

Genera	Species	Accession number
Acanthogobius	*Acanthogobius hasta*	NC006131[a]
Acentrogobius	*Acentrogobius pflaumii*	NC018064[a]
Bathygobius	*Bathygobius coalitus*	AB429403
	Bathygobius cocosensis	AB374647
	Bathygobius cyclopterus	AB429402
	Bathygobius fuscus	AB429401
	Bathygobius hongkongensis	AB429400
Boleophthalmus	*Boleophthalmus pectinirostris*	NC016195
Callogobius	*Callogobius tanegasimae*	AB269898
Gillichthys	*Gillichthys mirabilis*	NC012906[a]
	Gillichthys seta	NC012908[a]
Glossogobius	*Glossogobius circumspectus*	NC018824[a]
	Glossogobius olivaceus	NC016664[a]
Gymnogobius	*Gymnogobius petschiliensis*	NC008743[a]
Luciogobius	*Luciogobius platycephalus*	NC019811[a]
Odontamblyopus	*Odontamblyopus rubicundus*	NC019647[a]
	Odontamblyopus lacepedii	FJ968551
	Odontamblyopus Rebecca	FJ968569
	Odontamblyopus sp.	FJ968564
Oxuderces	*Oxuderces dentatus*	NC016194[a]
Periophthalmus	*Periophthalmus argentilineatus*	AB257626
	Periophthalmus modestus	AB257624
Rhinogobius	*Rhinogobius brunneus*	AB674658
	Rhinogobius duospilus	AB190337
	Rhinogobius flumineus	AB190326
	Rhinogobius giurinus	AB190338
	Rhinogobius_sp.	AB190335
Scartelaos	*Scartelaos histophorus*	NC017888[a]
Sicyopterus	*Sicyopterus japonicas*	AB613024
Stiphodon	*Stiphodon alcedo*	NC018054[a]
	Stiphodon atropurpureus	AB613042
	Stiphodon elegans	AB613025
	Stiphodon imperiorientis	AB613026
	Stiphodon percnopterygionus	AB613059
Synechogobius	*Synechogobius ommaturus*	JX186192[a]
Taenioides	*Taenioides anguillaris*	FJ968577
	Taenioides cirratus	FJ968576
Tridentiger	*Tridentiger barbatus*	JX536694
	Tridentiger bifasciatus	NC015992[a]
Trypauchen	*Trypauchen vagina*	NC016693[a]
Valenciennea	*Valenciennea strigata*	AB429399

[a]This accession number represents the number of complete mtDNA sequence.

Table 3. List II of species from Gobiidae used in phylogenetic analyses, with GenBank accession numbers.

Genera	Species	Accession number		
Acanthogobius	*Acanthogobius hasta*		NC006131[a]	
Acentrogobius	*Acentrogobius pflaumii*		NC018064[a]	
Amblyeleotris	*Amblyeleotris fasciata*	HQ536758	HQ536784	HQ536686
	Amblyeleotris guttata	HQ536711	FJ796083	HQ536670
	Amblyeleotris gymnocephala	HQ536746	FJ796084	HQ536676
	Amblyeleotris periophthalma	HQ536723	FJ796085	HQ536651
	Amblyeleotris steinitzi	HQ536713	FJ796095	HQ536644
	Amblyeleotris yanoi	HQ536726	FJ796097	HQ536654
Arenigobius	*Arenigobius bifrenatus*	HQ909513	HQ909562	HQ909455
Asterropteryx	*Asterropteryx ensifera*	HQ909506	FJ796099	HQ536656
Bathygobius	*Bathygobius soporator*	HQ909514	HQ909563	HQ909456
Boleophthalmus	*Boleophthalmus pectinirostris*		NC016195[a]	
Bryaninops	*Bryaninops yongei*	HQ909516	HQ909565	HQ909458
Caffrogobius	*Caffrogobius saldanha*	HQ909508	HQ909560	HQ909450
Coryogalops	*Coryogalops anomolus*	HQ909517	HQ909566	HQ909459
Cryptocentrus	*Cryptocentrus cinctus*	HQ536719	FJ796109	HQ536678
	Cryptocentrus inexplicatus	HQ536745	FJ796113	HQ536677
	Cryptocentrus leptocephalus	HQ536731	HQ536777	HQ536660
	Cryptocentrus lutheri	HQ536776	HQ536791	HQ536706
	Cryptocentrus nigroocellatus	HQ536714	FJ796120	HQ909445
	Cryptocentrus sp. *A*	HQ536737	FJ796123	HQ536668
	Cryptocentrus strigilliceps	HQ909502	FJ796124	HQ536638
Ctenogobiops	*Ctenogobiops aurocingulus*	GU187238	FJ796135	HQ536640
	Ctenogobiops formosa	GU187229	FJ796129	HQ536636
	Ctenogobiops mitodes	GU187223	FJ796127	HQ536702
	Ctenogobiops tongaensis	GU187232	FJ796131	HQ536666
Elacatinus	*Elacatinus evelynae*	HQ909525	HQ909572	HQ909468
Exyrias	*Exyrias belissimus*	HQ909520	HQ909567	HQ909462
Favonigobius	*Favonigobius exquisitus*	HQ909522	HQ909569	HQ909465
Gillichthys	*Gillichthys mirabilis*		NC012906[a]	
	Gillichthys seta		NC012908[a]	
Glossogobius	*Glossogobius circumspectus*		NC018824[a]	
	Glossogobius olivaceus		NC016664[a]	
Gobiosoma	*Gobiosoma chiquita*	HQ909524	HQ909571	HQ909467
Gymnogobius	*Gymnogobius petschiliensis*		NC008743[a]	
Istigobius	*Istigobius decorates*	HQ909527	HQ909573	HQ909470
	Istigobius rigilius	HQ536742	FJ796137	HQ536672
Luciogobius	*Luciogobius platycephalus*		NC019811[a]	
Mahidolia	*Mahidolia mystacina*	HQ536728	FJ796119	HQ536696
Microgobius	*Microgobius microlepis*	HQ909537	HQ909583	HQ909477
Odontamblyopus	*Odontamblyopus rubicundus*		NC019647[a]	
Oplopomus	*Oplopomus oplopomus*	HQ536732	FJ796151	HQ909497
Oxuderces	*Oxuderces dentatus*		NC016194[a]	
Pomatoschistus	*Pomatoschistus minutus*	HQ909555	HQ909596	HQ536661
Psammogobius	*Psammogobius biocellatus*	HQ909507	HQ909559	HQ909449
Scartelaos	*Scartelaos histophorus*		NC017888[a]	
Sicyopterus	*Sicyopterus japonicas*		NC018826[a]	
Signigobius	*Signigobius biocellatus*	HQ536743	FJ796149	HQ536673
Stiphodon	*Stiphodon alcedo*		NC018054[a]	
Stonogobiops	*Stonogobiops xanthorhinica*	HQ536727	FJ796152	HQ536655
Synechogobius	*Synechogobius ommaturus*		JX186192[a]	
Tomiyamichthys	*Tomiyamichthys lanceolatus*	HQ536733	FJ796158	HQ536662
Tridentiger	*Tridentiger barbatus*		NC018823[a]	
	Tridentiger bifasciatus		NC015992[a]	
Trypauchen	*Trypauchen vagina*		NC016693[a]	
Valenciennea	*Valenciennea longipinnis*	HQ909504	FJ796154	HQ536648
	Valenciennea puellaris	HQ536710	FJ796157	HQ536635
Vanderhorstia	*Vanderhorstia ornatissima*	HQ536759	HQ536789	HQ536690
Zosterisessor	*Zosterisessor ophiocephalus*	HQ909554	HQ909595	HQ909496

[a]This accession number represents the number of complete mtDNA sequence.

Sequence analysis

The sequences of the PCR products obtained were edited in the Seqman program (DNAstar, Madison, USA) for assembling contiguous, overlapping sequences to obtain a complete mitochondrial genome sequence. The 13 protein-coding genes and two ribosomal RNA genes were noted using DOGMA software (Wyman *et al*. 2004) through reference sequences of Gobiidae which are available in GenBank (http://www.ncbi.nlm.nih.gov/genbank) to determine their respective gene boundaries. tRNA genes and their secondary structures were identified by means of tRNAscan-SE 1.21 software (Lowe and Eddy 1997). Base composition and codon usage were calculated using Mega 5.0 software (Tamura *et al*. 2011). Putative L-strand origin, CR and conserved motifs were identified via sequence homology and proposed secondary structure. The complete mitochondrial genome sequence of *O. rubicundus* had been deposited in GenBank, which can be accessed with accession number NC019647.

Phylogenetic analysis

Three group datasets (tables 2, 3 and 4) which were collected from the available sequences in GenBank were used in phylogenetic analysis. The first group (table 2) that was *ND5* sequences of Gobiidae was used to construct phylogenetic trees using NJ, ME and ML methods as implemented in Mega 5.0 (Tamura *et al*. 2011), *Beryx splendens* was regarded as outgroup. The second group (table 3) included *ND1*, *ND2* and *CO1* sequences of Gobiidae and the third group (table 4) that was the currently available complete mtDNA sequence from Gobiidae were respectively used for building phylogenetic trees using Bayesian inference (BI) analysis which is appropriate for relatively long sequence by MrBayes ver. 3.1.2 (Huelsenbeck and Ronquist 2001). The best-fit model GTR + I + G was selected for BI analysis by ModelTest 3.7 (Posada and Crandall 1998).

In the Bayesian phylogenetic analysis, two species (*Beryx splendens* and *Odontobutis platycephala*) served as outgroups. Four Markov chains were run for 5,000,000 generations by sampling the trees every 1000 generations. The 25% of trees were discarded as burn-in, and the remaining 75% sampled trees were used in constructing a 50% majority rule consensus tree and to calculate the posterior probability values.

Table 4. List III of species from Gobiidae used in phylogenetic analyses, with GenBank accession numbers.

Genera	Species	Accession number
Acanthogobius	*Acanthogobius hasta*	NC006131
Acentrogobius	*Acentrogobius pflaumii*	NC018064
Boleophthalmus	*Boleophthalmus pectinirostris*	NC016195
Gillichthys	*Gillichthys mirabilis*	NC012906
	Gillichthys seta	NC012908
Glossogobius	*Glossogobius circumspectus*	NC018824
	Glossogobius olivaceus	NC016664
Gymnogobius	*Gymnogobius petschiliensis*	NC008743
Luciogobius	*Luciogobius platycephalus*	NC019811
Odontamblyopus	*Odontamblyopus rubicundus*	NC019647
Oxuderces	*Oxuderces dentatus*	NC016194
Scartelaos	*Scartelaos histophorus*	NC017888
Sicyopterus	*Sicyopterus japonicas*	NC018826
Stiphodon	*Stiphodon alcedo*	NC018054
Synechogobius	*Synechogobius ommaturus*	JX186192
Tridentiger	*Tridentiger barbatus*	NC018823
	Tridentiger bifasciatus	NC015992
Trypauchen	*Trypauchen vagina*	NC016693

Results and discussion

Genome composition

The mitochondrial DNA of *O. rubicundus* is a closed double-stranded circular molecule of 17119 nucleotides (GenBank accession number: NC019647). Its mitogenome content, gene order and gene coding strand are given in figure 1 in electronic supplementary material at http://www.ias.ac.in/jgenet/. The mitogenome of *O. rubicundus* as in other vertebrates contains 13 protein-coding genes, two rRNA genes, 22 tRNA genes, a CR and an L-strand origin (O_L) (Cheng *et al*. 2010; Shi *et al*. 2012). Most of mitochondrial genes are encoded on H-strand except for *ND6* and seven tRNA genes (figure 1 in electronic supplementary material; table 5). The 22 tRNA genes scatter throughout the genome and range from 65 to 74 bp in size. The overall base composition is T, 27.2%; C, 28.5%; A, 28.7% and G, 15.7% (table 6). The content A + T with 55.9% is slightly rich than the G + C content, which is similar to other teleosts (Tzeng *et al*. 1992; Wang *et al*. 2007).

Protein-coding genes

Among the 13 protein-coding genes, 12 are encoded on heavy strand while only one (*ND6*) is encoded on the light strand (figure 1 in electronic supplementary material; table 5). Four overlaps occur in protein-coding genes: *ATPase8* and *ATPase6*, *ND4* and *ND4L*, *ATPase6* and *CO3*, which share four, seven and one nucleotides on the heavy strand, respectively. Four nucleotides are shared between *ND5* and *ND6*. In addition, there are four overlaps between other genes: *CO3* and $tRNA^{Gly}$ (one nucleotide), *ND3* and $tRNA^{Arg}$ (two nucleotides) (table 5).

ATG is the initiation codon of 11 out of 13 protein-coding genes (*ND1*, *ND2*, *CO2*, *ATPase8*, *CO3*, *ND3*, *ND4L*, *ND4*, *ND5*, *ND6* and *Cytb*), while the initiation codon of *CO1* is GTG, which is different from other bonyfish. In most of them, *CO1* gene uses ATG as the start codon. The initiation of ATPase 6 is ATA, which is rare in other bonyfish. The stop codon of *ND2*, *CO1*, *ND4L* and *ND5* is TAA. *ND1*, *ATPase8*, *ND3* and *ND6* end with TAG. The remaining five genes have incomplete stop codons, either TA or T, which

Table 5. Organization of the mitochondrial genome of *O. rubicundus*.

Gene	Position		Size (bp)		Codon		Intergenic	Strand
	From	To	Nucleotide	Amino acid	Initiation	Stop[a]	nucleotide[b]	
tRNA^{Phe}	1	68	68					H
12S rRNA	69	1016	948					H
tRNA^{Val}	1017	1088	72					H
16S rRNA	1089	2776	1688					H
tRNA^{Leu}(*UUR*)	2777	2850	74					H
ND1	2851	3825	975	324	ATG	TAG	4	H
tRNA^{Ile}	3830	3899	70				−1	H
tRNA^{Gln}	3899	3969	69				−1	L
tRNA^{Met}	3969	4037	71					H
ND2	4038	5084	1047	348	ATG	TAA		H
tRNA^{Trp}	5085	5155	71				2	H
tRNA^{Ala}	5158	5226	69				1	L
tRNA^{Asn}	5228	5300	73				36	L
tRNA^{Cys}	5337	5401	65					L
tRNA^{Tyr}	5402	5472	71				1	L
CO1	5474	7027	1554	517	GTG	TAA		H
tRNA^{Ser}(*UCN*)	7028	7098	71				3	L
tRNA^{Asp}	7102	7173	72				4	H
CO2	7178	7868	691	230	ATG	T–		H
tRNA^{Lys}	7869	7944	76				1	H
ATPase8	7946	8110	165	54	ATG	TAG	−4	H
ATPase6	8107	8789	683	227	ATA	TA-	−1	H
CO3	8789	9573	785	261	ATG	TA-	−1	H
tRNA^{Gly}	9573	9644	72					H
ND3	9645	9995	351	116	ATG	TAG	−2	H
tRNA^{Arg}	9994	10062	69					H
ND4L	10063	10359	297	98	ATG	TAA	−7	H
ND4	10353	11733	1381	460	ATG	T–		H
tRNA^{His}	11734	11802	69					H
tRNA^{Ser}(*AGY*)	11803	11870	68				4	H
tRNA^{Leu}(*CUN*)	11875	11947	73					H
ND5	11948	13786	1839	612	ATG	TAA	−4	H
ND6	13783	14304	522	173	ATG	TAG		L
tRNA^{Glu}	14305	14373	69				5	L
Cytb	14379	15519	1141	380	ATG	T–		H
tRNA^{Thr}	15520	15591	72					H
tRNA^{Pro}	15592	15661	70					L
CR	15662	11719	1458					H

[a]TA- and T– represent incomplete stop codons.
[b]Numbers correspond to the nucleotides separating adjacent genes. Negative numbers indicate overlapping nucleotides.

Table 6. Base composition of *O. rubicundus* mitochondrial genome.

Gene/region	Base composition (%)				
	T (U)	C	A	G	A + T (U)
Protein coding					
1st	21	27.7	25	26	48.7
2nd	41	27.7	18.1	13.5	49.1
3rd	28	32.1	33.3	6.9	61.3
Total	29.9	29.2	25.5	15.5	55.4
tRNA	24.9	24.7	31.1	19.3	56
sRNA	21.5	25	33.9	19.6	55.4
Control region	28.2	24.6	33.1	14.1	61.8
Overall	27.2	28.5	28.7	15.7	55.9

Table 7. Codon usage in *O. rubicundus* mitochondrial protein-coding genes.

Amino acid	Codon	Number	Amino acid	Codon	Number
Phe	TTT	136	Tyr	TAT	55
	TTC	100		TAC	53
Leu	TTA	100	His	CTA	30
	TTG	23		CAC	82
	CTT	181	Gln	CAA	91
	CTC	158		CAG	10
	CTA	162	Asn	AAT	38
	CTG	45		AAC	80
Ile	ATT	170	Lys	AAA	63
	ATC	94		AAG	10
Met	ATA	104	Asp	GAT	32
	ATG	53		GAC	42
Val	GTT	81	Glu	GAA	72
	GTC	47		GAG	21
	GTA	67	Cys	TGT	10
	GTG	30		TGC	20
Ser	TCT	63	Trp	TGA	104
	TCC	71		TGG	12
	TCA	49	Arg	CGT	4
	TCG	2		CGC	18
Pro	CCT	64		CGA	43
	CCC	96		CGG	7
	CCA	65	Ser	AGT	7
	CCG	1		AGC	47
Thr	ACT	54	Gly	GGT	37
	ACC	93		GGC	80
	ACA	141		GGA	81
	ACG	1		GGG	45
Ala	GCT	91	Stop	TAA	9
	GCC	144		TAG	4
	GCA	120		AGA	0
	GCG	0		AGG	0

are presumably completed as TAA by posttranscriptional polyadenylation (Ojala *et al*. 1981).

The base composition of the 13 protein-coding genes is 29.9% T, 29.2% A, 25.5% C and 15.5% G. The third codon position of the *O. rubicundus* has strong bias against G, which is consistent with other vertebrates (Broughton *et al*. 2001). Because of the hydrophobic character of the proteins (Naylor *et al*. 1995), pyrimidines are over rich (T + C = 68.7) at the second codon position. However, the purines are higher than pyrimidines at the first codon position (table 6).

The codon usage of 13 protein genes was identified (table 7). The mitochondrial DNA encodes 3813 amino acids including stop codons in all protein-coding genes. For amino acids with fourfold degenerate third position codons, T is the most frequent, followed by A and C for Leu and Val. Among Ser, Pro and Ala, C is mostly seen. However, for Gly, A is frequently used than other bases. Among twofold degenerate codons, T appears to be used more than C in pyrimidine codon family. Except for Gly, G is the least frequent third position nucleotide in all codon families. All these features are very similar to vertebrates (Lee and Kocher 1995; Xia *et al*. 2007).

Noncoding regions

There are two noncoding regions in the mitochondrial genome of *O. rubicundus*. One of them is an L-strand origin (O_L) which places in a cluster of five tRNA genes (WANCY) as in other vertebrates (Oh *et al*. 2007). The putative O_L, with 56 bp, is located between $tRNA^{Asn}$ and $tRNA^{Cys}$. And it has the potential to fold into a stable stem–loop secondary structure, with a stem formed by 12 paired nucleotides and a loop of 14 nucleotides (figure 1). As in the other gobiid fishes, the motif 5′-GCCGG-3′, which is found at the base of the stem with the $tRNA^{Cys}$, is likely to be involved in the transition from RNA to DNA synthesis (Hixson *et al*. 1986).

The CR, with 1458 bp, is located between $tRNA^{Pro}$ and $tRNA^{Phe}$ is the other noncoding region (Cheng *et al*. 2011a, b; Xu *et al*. 2011). As in other fishes, its base composition is rich in A and T (A + T= 61.8%). In general, the CR includes three domains, which are termination-associated sequence domain, the central conserved sequence block domain and the conserved sequence block domain, respectively (Sbisa *et al*. 1997). In mammals, the conserved sequences CSB-B, CSB-C, CSB-D, CSB-E and CSB-F are recognized in the central conserved sequence block domain (Southern *et al*. 1988). However, only CSB-F, CSB-E and CSB-D can be identified in fishes (Broughton and Dowling 1994). Compared to the recognition domain in gobiid fishes, the TAS, CSB-2 and CSB-3 could be detected in the CR of *O. rubicundus*. The central conserved sequence block domain (CSB-F, CSB-E and CSB-D) and CSB-1 could not be identified (figure 2). The lack of the central conserved sequence block domain and CSB-1 is confirmed in *Acanthogobius hasta* (II-Chan *et al*. 2004).

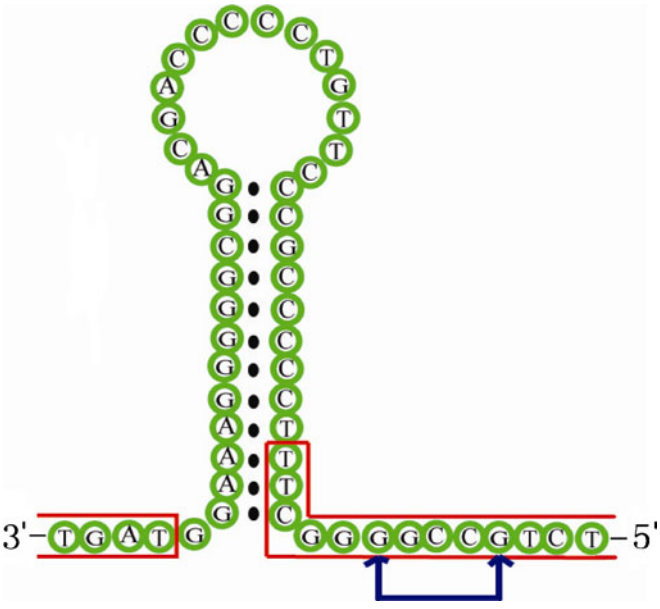

Figure 1. The stem–loop secondary structure of O_L from *O. rubicundus*. The motif 5′-GCCGG-3′ in the $tRNA^{Cys}$ gene is underlined with arrows.

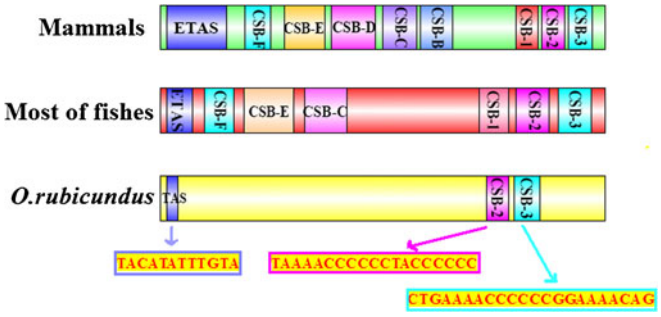

Figure 2. The structure pattern of CR of mammals, most of the fishes and *O. rubicundus*. In the CR of *O. rubicundus*, the putative conserved elements TAS, CSB2 and CSB3 are marked.

A TAS motif (TACATATTTGTA) is identified in the termination-associated sequence domain. In addition, only CSB-2 and CSB-3 are found in CSB domain at the 3′-end of the CR.

Ribosomal and transfer RNA genes

The 12S and 16S rRNA genes of *O. rubicundus* are 948 and 1688 nucleotides long, respectively (table 5). As in other vertebrates, they are located between $tRNA^{Phe}$ and $tRNA^{Leu}$ (UUR) genes, and are separated by the $tRNA^{Phy}$ (figure 1 in electronic supplementary material). Their base composition is T, 25.1%; C, 25%; A, 33.1% and G, 19.6% (table 6). The A + T content is 55.4%, which is similar to other bony fishes (Zardoya and Meyer 1997).

The mitochondrial genome of *O. rubicundus* encodes 22 tRNA genes, which are interspersed between the rRNA genes and protein-coding genes and range from 65 bp to 76 bp (figure 1 in electronic supplementary material; table 5). The structures of 20 tRNAs were predicted. The other two tRNA genes were identified on basis of their respective anticodons and secondary structures. There are two forms of $tRNA^{Ser}$ (UCN and AGN) among 22 tRNAs, so is $tRNA^{Leu}$ (UUR and CUN). $tRNA^{Ser}$ (AGN) is different from other tRNAs, which lacks the DHC stem (figure 2 in electronic supplementary material). Moreover, three tRNA clusters (HSL, IQM and WANCY) were determined in *O. rubicundus*, as in other vertebrates (Jin *et al*. 2012). The feature is a common phenomenon in vertebrate mitogenomes (Lee and Kocher 1995). Thirty-five pair mismatches are identified in the stem of 22 tRNAs, including seven pairs in the TΨC stems, 10 pairs in the anticodon stems, seven pairs in the DHU stems and 11 pairs in the amino acid acceptor stems.

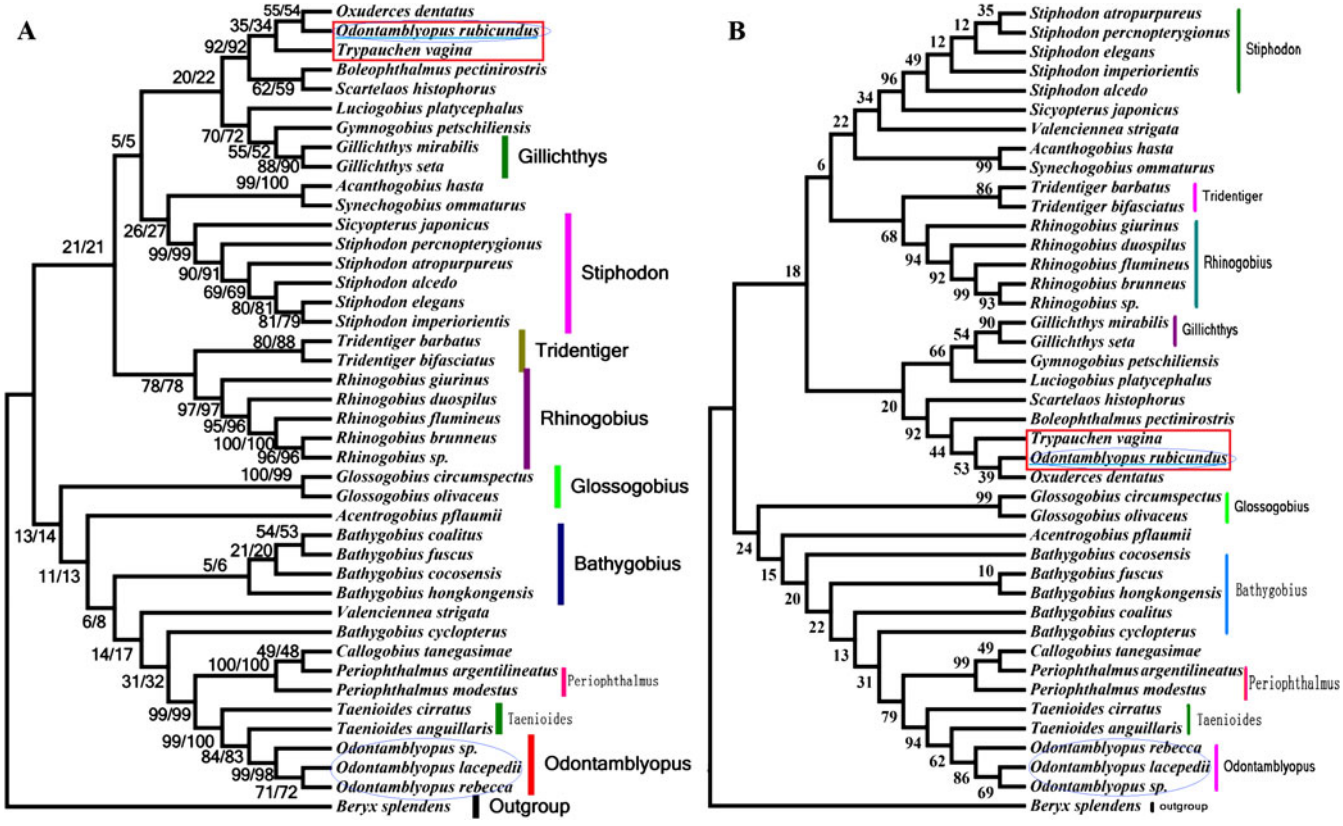

Figure 3. Phylogenetic tree of the Gobiidae based on partial *ND5* genes. Numbers above branches specify bootstrap percentages for ME (1000 replications) and NJ (1000 replications) (A), and ML (1000 replications) (B) analyses. Bootstrap values are given for each branch. *Beryx splendens* was used as outgroup.

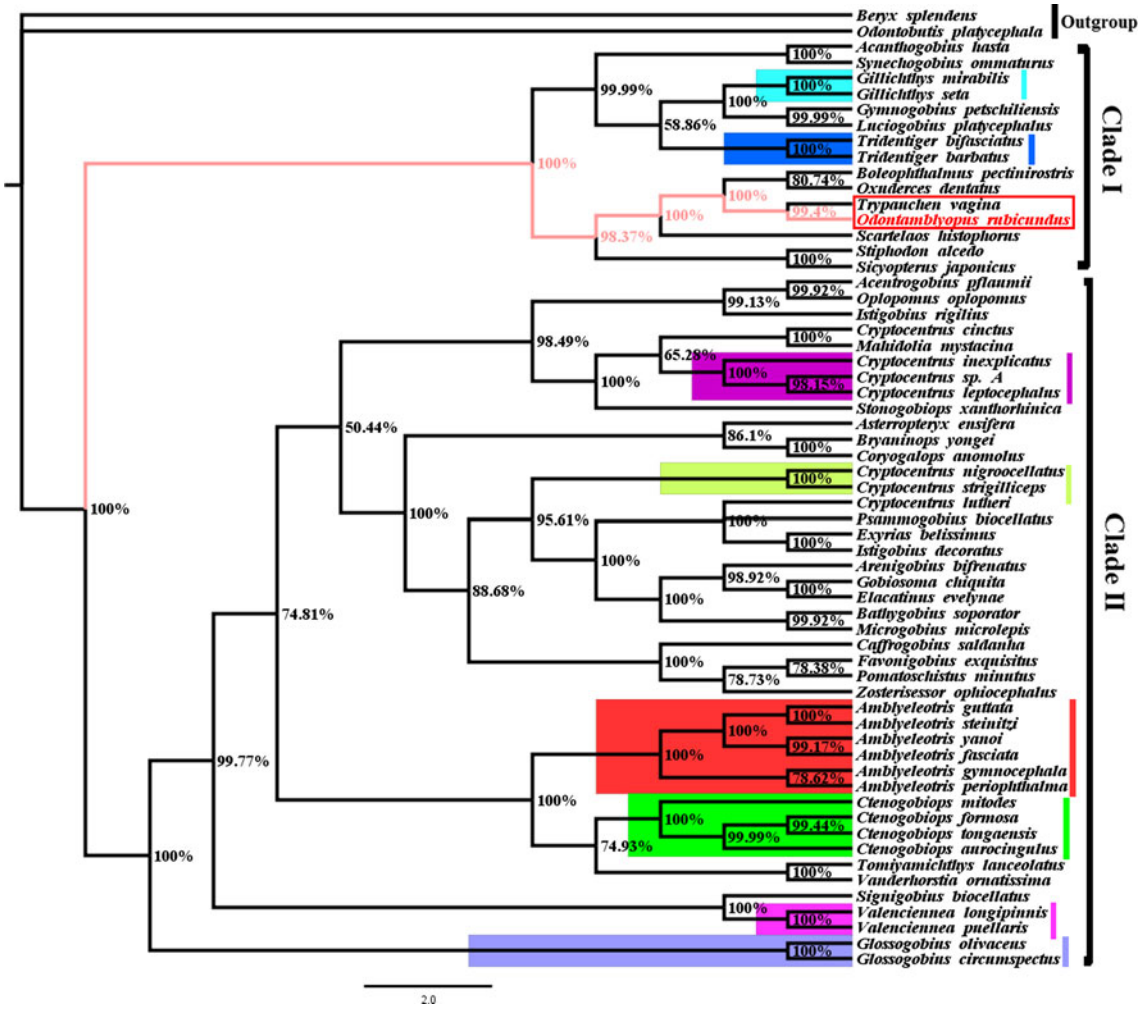

Figure 4. Bayesian tree reconstructed based on *ND1*, *ND2* and *CO1* from 59 species of Gobiidae. The values beside the nodes are Bayesian posterior probabilities. *B. splendens* and *O. platycephala* were used as outgroups, respectively.

The mismatches bases are mainly G–U and U–G. In addition, A–A, U–U, A–C, C–A, U–C and C–U mismatches are also observed (figure 2 in electronic supplementary material). The reason of these mismatches may be that mitochondrial DNA is not subjected to recombination process (Lynch 1997).

Phylogenetic analysis

The phylogenetic trees (figure 3) that based on ND5 using NJ, ME and ML are similar. The difference is only in detail and bootstrap value. In the trees based on ND5 using different methods, *O. rubicundus* cannot be merged into the *Odontamblyopus*, a result differing from those of morphological researches (Hamilton 1822). In our results, *O. rubicundus* is closely related to *Oxuderces dentatus* and *Trypauchen vagina*. However, the bootstrap value of phylogenetic trees is very poor because the sequences which were used for building phylogenetic tree are short and limited. So, to determine phylogenetic relationship, the second group data (table 3) and the third group data (table 4) were respectively used to construct phylogenetic trees using Bayesian inference analysis. In two phylogenetic trees (figures 4 and 5), based on the second and third data, *O. rubicundus* is also found to be closely related to *Trypauchen vagina*. So we believe that the classification of *O. rubicundus* into the genus *Odontamblyopus* based on morphology is debatable. However, to date, it is difficult to clear the phylogenetic position of *O. rubicundus* based on the available sequences in GenBank. So to clarify the taxonomic position of *O. rubicundus*, more sequences of *Odontamblyopus* will be necessary.

Phylogenetic analysis of ND1, ND2 and CO1 sequences using Bayesian inference analysis also reveals two distinct clades (figure 4). *G. petschiliensis* and *L. platycephalus* are in close relationship in clade I, as had been previously reported (Jeon *et al*. 2012). The phylogenetic trees based on the currently available complete mtDNA sequence from Gobiidae using Bayesian inference analysis also show two distinct clades (figure 5). The phylogenetic relationship of the species in figure 5 is completely consistent with figure 4.

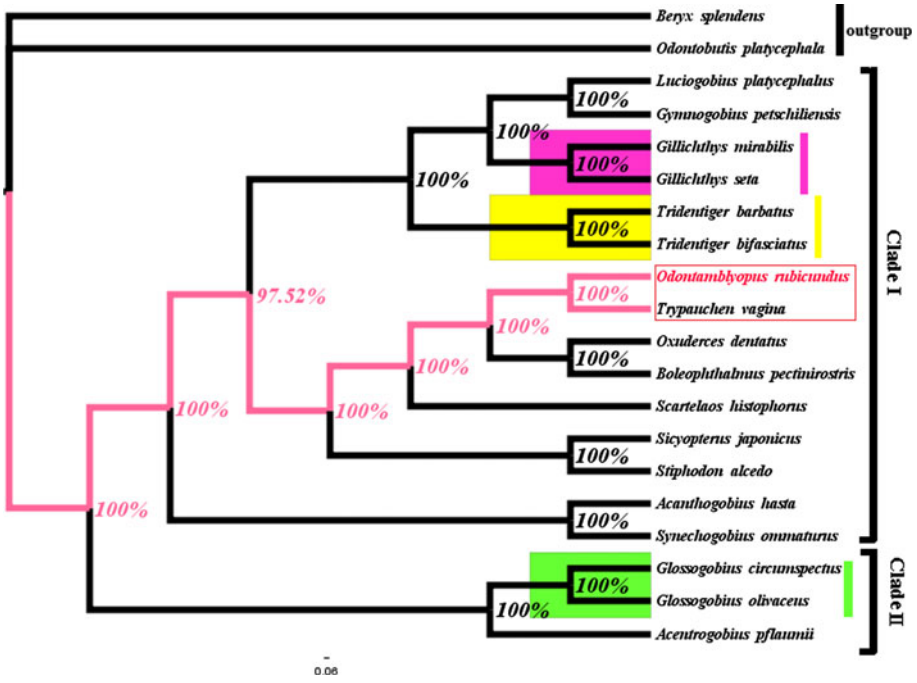

Figure 5. Bayesian tree reconstructed based on complete mtDNA sequence of 18 species of Gobiidae. The values beside the nodes are Bayesian posterior probabilities. *B. splendens* and *O. platycephala* were used as outgroups, respectively.

Acknowledgements

This study was supported by National Natural Science Foundation of China (31272661) and Zhejiang Provincial Natural Science Foundation (LY13C040001). We thank Prof. Shenglong Zhao for the help of fish sampled and identified by morphology.

References

Avise J. C. 2000 *Phylogeography*. Harvard Univ. Press, London, UK.

Bleeker P. 1872 Mémoire sur la fauna ichthyologique de Chine. *Ned Tijdschr. Dierk.* **4**, 128–129.

Boore J. L. 1999 Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767–1780.

Boore J. L., Macey J. R. and Medina M. 2005 Sequencing and comparing whole mitochondrial genomes of animals. *Methods Enzymol.* **395**, 311–348.

Broughton R. E. and Dowling T. E. 1994 Length variation in mitochondrial DNA of the *minnow cyprinella spiloptera*. *Genetics* **138**, 179–190.

Broughton R. E., Milam J. E. and Roe B. A. 2001 The complete sequence of the *zebrafish* (Danio rerio) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genome Res.* **11**, 1958–1967.

Cheng Y. Z., Xu T. J., Jin X. X. and Wang R. X. 2011a Complete mitochondrial genome of the yellow drum *Nibea albiflora* (Perciformes, Sciaenidae). *Mitochondrial DNA* **22**, 80–82.

Cheng Y. Z., Wang R. X. and Xu T. J. 2011b The mitochondrial genome of the spinyhead croaker *Collichthys lucida*: genome organization and phylogenetic consideration. *Mar. Genomics* **4**, 17–23.

Cheng Y. Z., Xu T. J., Shi G. and Wang R. X. 2010 Complete mitochondrial genome of the *miiuy croaker* Miichthys miiuy (Perciformes, Sciaenidae) with phylogenetic consideration. *Mar. Genomics* **3**, 201–209.

Giinthe A. 1861 On a collection of fishes sent by Capt. Dow from the Pacific coast of central America. *Proc. Zool. Soc. London* **3**, 370–376.

Gosline W. 1995 The osteology and relationships of certain Gobioid fishes, with particular reference to the genera Kraemeria and Microdesmus. *Pac. Sci.* **9**, 158–170.

Hamilton F. 1822 *An account of the fishes found in the river Ganges and its branches*, pp. 1762–1829. A. Constable and Company, Edinburgh, UK.

Hixson J. E., Wong T. W. and Clayton D. A. 1986 Both the conserved and divergent 5′-flanking sequences are required for initiation at the human mitochondrial origin of light strand replication. *J. Biol. Chem.* **261**, 2384–2390.

Huelsenbeck J. P. and Ronquist F. 2001 MRBAYES: Bayesian inference of phylogenetic human mitochondria. *Nature* **290**, 470–474.

II-Chan K., Young J. K. and Chang B. K. 2004 The complete mitochondrial genome of the javeline goby *Acanthogobius hasta* (Perciformes, Gobiidae) and phylogenetic considerations. *Gene* **366**, 147–153.

Jeon H. B., Choi S. H. and Suk H. Y. 2012 Exploring the utility of partial cytochrome c oxidase subunit 1 for DNA barcoding of gobies. *Anim. Syst. Evol. Divers.* **4**, 269–278.

Jin X. X., Wang R. X., Xu T. J. and Shi G. 2012 Complete mitochondrial genome of *Oxuderces dentatus* (Perciformes, Gobioidei). *Mitochondrial DNA* **23**, 142–144.

Kawaguchi A., Miya M. and Nishida M. 2001 Complete mitochondrial DNA sequence of *Aulopus japonicus* (Teleostei: Aulopiformes), a basal Eurypterygii: longer DNA sequences and higher-level relationships. *Ichthyol. Res.* **48**, 213–223.

Lee W. J. and Kocher T. D. 1995 Complete sequence of a sea lamprey (*Pettromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. *Genetics* **139**, 873–887.

Lowe T. M. and Eddy S. R. 1997 tRNAscan-SE: a program for improved detection of transfer RNA genes in genome sequence. *Nucleic Acids Res.* **25**, 955–964.

Lynch M. 1997 Mutation accumulation in nuclear, organelle, and prokaryotic genomes: transfer RNA genes. *Mol. Biol. Evol.* **14**, 914–925.

Miya M. and Nishida M. 1999 Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. *Mar. Biotechnol.* **1**, 416–426.

Miya M. and Nishida M. 2000 Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. *Mol. Phylogenet. Evol.* **17**, 437–455.

Naylor G. J., Collins T. M. and Brown W. M. 1995 Hydrophobicity and phylogeny. *Nature* **373**, 555–556.

Nelson J. S. 2006 *Fishes of the world*, 4th edition, pp. 1–601. John Wiley, Hoboken, USA.

Oh D. J., Kim J. Y., Lee J. A., Yoon W. J., Park S.Y. and Jung Y. H. 2007 Complete mitochondrial genome of the rock bream *Oplegnathus fasciatus* (Perciformes, Oplegnathidae) with phylogenetic considerations. *Gene* **392**, 174–180.

Ojala D., Montoya J. and Attardi G. 1981 tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**, 470–474.

Posada D. and Crandall K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Sbisa E., Tanzariello F., Reyes F., Pesole G. and 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.

Shi G., Jin X. X., Zhao S. L., Xu T. J. and Wang R. X. 2012 Complete mitochondrial genome of *Trypauchen vagina* (Perciformes, Gobioidei). *Mitochondrial DNA* **23**, 151–153.

Southern S. O., Southern P. J. and Dizon A. E. 1988 Molecular characterization of a cloned dolphin mitochondrial genome. *J. Mol. Evol.* **28**, 32–40.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M. and Kumar S. 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Bio. Evol.* **28**, 2731–2739.

Tzeng C. S., Hui C. F., Shen S. C. and Huang P. C. 1992 The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variation among vertebrates. *Nucleic Acids Res.* **20**, 4853–4858.

Wang X. Z., Wang J., He S. P. and Mayden R. L. 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.

Wolstenholme D. R. 1992 Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* **141**, 173–216.

Wyman S. K., Jansen R. K. and Boore J. L. 2004 Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20**, 3253–3255.

Xia J., Xia K., Gong J. and Jiang S. 2007 Complete mitochondrial DNA sequence, gene organization and genetic variation of control regions in *Parargyrops edita*. *Fish Sci.* **73**, 1042–1049.

Xu T. J., Cheng Y. Z., Sun Y. N., Shi G. and Wang R. X. 2011 The complete mitochondrial genome of bighead croaker, *Collichthys niveatus* (Perciformes, Sciaenidae): structure of control region and phylogenetic considerations. *Mol. Biol. Rep.* **38**, 4673–4685.

Zander C. D. 2011 Morphological adaptation to special environments of gobies. In *The biology of gobies* (ed. R. A. Patzner, J. L. van Tassell, M. Kovăcić and B. G. Kapoor), pp. 345–366. Science Publishers, New York, USA.

Zardoya R. and Meyer A. 1997 The complete DNA sequence of the mitochondrial genome of a 'living fossil', the coelacanth (*Latimeria chalumnae*). *Genetics* **146**, 995–1010.

Received 11 March 2013, in revised form 19 May 2013; accepted 17 June 2013
Published on the Web: 20 November 2013