View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by AJOL - African Journals Online

African Journal of Biotechnology Vol. 11(26), pp. 6904-6909, 29 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.3556
ISSN 1684–5315 ©2012 Academic Journals

Full Length Research Paper

Structure of mitochondrial DNA control region of *Argyrosomus amoyensis* and molecular phylogenetic relationship among six species of Sciaenidae

Chao Chen[1,2]*, Yan Lu Li[1,2], Lu Wang[1] and Guang Ye Gong[1]

[1]Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, 266071, China.
[2]College of Fisheries and Science, Shanghai Ocean University, Shanghai, 201306, China.

Accepted 17 February, 2012

The structure of the mitochondrial DNA (mtDNA) control region of *Argyrosomus amoyensis* was examined in this study. TAS, cTAS, CSB-D to CSB-F and CSB-1 to CSB-3 segments were detected in the species. The results indicated that the structures of these parts were different from that of most fishes. All the mtDNA control region sequences examined had missing tandem repeat sequences downstream of CSB-3, which were the same as most fishes'. In addition, part of the COI gene was used to analyze the phylogenetic relationships of six Sciaenids species. The phylogenetic tree results supported the classification by traditional morphology, and COI barcodes were useful for identifying these six species of Sciaenids.

Key words: Control region, structure, *Argyrosomus amoyensis*, COI, phylogenetic relationship, Sciaenidae.

INTRODUCTION

Argyrosomus amoyensis is a commercially important fish belonging to the Sciaenidae family. It is mainly distributed along the Chinese seaside from the south Yellow Sea to Taiwan Strait (Wu et al., 1998). In recent years, as a result of high levels of cultivation, the genetic diversity of *A. amoyensis* has reduced. Since measures are required to be taken to protect this species, more studies are needed. Hence a study of the genetic diversity and structure of the present wild populations of *A. amoyensis* is essential.

Due to its compact size (16 to 17 kb), high rate of mutation and exclusive maternal mode of inheritance (Brown et al., 1979; Harrison, 1989), the mitochondrial genome has been widely used as a marker in molecular genetic studies. As a closed circular molecule, the piscine mtDNA contains a set of 37 genes specifying 13 proteins, 2 rRNAs and 22 tRNAs encoded in both the heavy (H) and light (L) DNA strands (Meyer, 1993). The control region (D-loop) is the only significant non-coding segment in the vertebrate mitochondrial genome (Shui et al., 2008). Numerous studies have documented a variation of DNA sequence in this region. Despite the presence of several highly conserved sequences (Saccone et al., 1987), this region is known to exhibit some of the highest rates of evolution in the mitochondrial DNA. Most of these variations consist of nucleotide substitutions, and small insertions and deletions. However, a considerable variation of length has been observed in an ever-growing list of species (Lunt et al., 1998), which is caused by variation in the number of tandemly repeated sequences. These repeats can be observed near the 5′ end of the control region (Starner et al., 2004) or near the 3′ end (Broughton and Dowling, 1994), ranging from the size of four to hundreds of nucleotides, and they may have two to more than 100 copies (Lunt et al., 1998).

DNA barcoding is a new method of biological taxonomy based on the techniques of molecular biology and bioinformatics. Tautz et al. (2002) were the first to propose that DNA sequences could be used as a tool to identify a species or a taxon. Hebert et al. (2003) opined

*Corresponding author. E-mail: ysfrichenchao@126.com.

that the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon 'barcodes'; he established that the mtDNA cytochrome c oxidase subunit gene (COI) can serve as the core of a global bio-identification system for animals. Thus, a significant amount of work has been carried out to investigate the possibility of using the COI gene as a DNA barcode to identify species in many different animal groups. The results of these previous studies indicated that using COI gene sequences as DNA barcodes were generally effective, delivering more than 95% species-level resolution (Hebert et al., 2004a, b; Cooper et al., 2007; Ratnasingham and Hebert, 2007; Pyle et al., 2008; Ward et al., 2008).

In this study, we reported the sequence and structural characteristics of the control region of *A. amoyensis* with the intent that the variable sequences in the control region may provide useful information on their genetic diversity. The COI gene was also used to reveal the phylogenetic relationships among six species of Sciaenidae.

MATERIALS AND METHODS

Sample collection

We collected samples from two populations of *A. amoyensis* in this study: nine samples from Xiamen, Fujian Province, and six samples from Zhangzhou, Fujian Province. Muscle samples were preserved in 95% ethanol before DNA extraction. In this study, we amplified three COI sequences of *Larimichthys crocea* and two COI sequences of *Argyrosomus argentatus* to reveal the phylogenetic relationships among them. COI sequences for *Argyrosomus hololepidotus* (Genbank accession no: DQ107796, DQ107810 and DQ107811), *Argyrosomus inodorus* (Genbank no: HM007715, HM007716 and HM007717) and *Argyrosomus japonicus* (Genbank no: HM007718, HM007719 and HM007720) were used as outgroups in the phylogenetic relationship study.

DNA extraction, PCR and sequencing

Genomic DNA was isolated from muscle tissue by proteinase K digestion followed by a standard phenol-chloroform method (Sambrook, 1989). Each polymerase chain reaction (PCR) was performed in a volume of 25 µL containing 10 to 25 ng template DNA, 2.5 µL of 10×PCR buffer, 2.5 µL of $MgCl_2$ (25 mM), 0.5 µL of dNTPs (10 mM), 10 pM of each primer and 2.5 units of *Taq* DNA polymerase (TaKaRa Biotechnology Co, Ltd.) in an Eppendorf Mastercycler 5500 (Eppendorf, Hamburg, Germany).

The complete sequences of the D-loop were amplified with primers DI-S: 5′-CCCACCACTAACTCCCAAAGC-3′ (forward) (Han et al., 2008) and CR: 5′- GTGCGGATACTTGCATGTGT-3′ (reverse), which were designed for this study. The initial denaturation was done for 5 min at 94°C, followed by 35 cycles of 45 s at 94°C for denaturation, 45 s at 50°C for annealing, 1 min at 72°C for extension and a final extension at 72°C for 10 min. Initial denaturation was for 5 min at 95°C, followed by 35 cycles of 1 min at 95°C for denaturation, 1 min at 50°C for annealing, 1 m at 72°C for extension and a final extension at 72°C for 10 min. The COI gene fragment was amplified following the pattern of Ward et al. (2005).

All sets of PCR included a negative control reaction tube in which all reagents were included except template DNA. PCR product was separated on a 1.5% agarose gel and purified with the BioDev Gel Extraction System B (BioDev Technology (Beijing) Co, Ltd). Both strands of the purified products were sequenced using the BigDye Terminator cycle sequencing kit v2.0 (Applied Biosystems, Foster City, CA, USA); sequencing was conducted on an ABI Prism 3730 automatic sequencer (Applied Biosystems) with both forward and reverse primers used for PCR amplification.

Sequence analysis

Sequences were edited and aligned using DNAStar software (DNASTAR, Inc., Madison, USA) and refined manually. The nucleotide compositions and the average distances between species and within species were calculated with MEGA 4.0 (Tamura et al., 2007). The gamma distribution shape parameters for the rates of heterogeneity were calculated using the program Modeltest 3.7 (Posada and Crandall, 1998).

Phylogenetic analysis

The phylogenetic relationships among the six species of Sciaenidae were constructed using neighbor joining (NJ), maximum parsimony (MP), maximum likelihood (ML) and Bayesian methods. The optimal bases substitution model and the optimized parameters for the NJ analysis were estimated by Modeltest 3.7 (Posada and Crandall, 1998) via the hierarchical likelihood ratio tests (hLRTs). The best fit model for the COI fragments was HKY85. NJ and MP trees were constructed by PAUP*. ML trees were reconstructed by PHYML. Bayesian analysis was conducted by MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with the following parameters: Statefreqpr = dirichlet (1, 1, 1, 1), nst = 2, rates = gamma, ngen = 1000000, samplefreq = 100, nchains = 4, savebrlens = yes, burnin = 5000, burninfrac = 0.25.

RESULTS

Structure of the control region

The sequence length of the control region of 15 *A. amoyensis* individuals were all 961 bp. The content of A+T (59.6%) was higher than that of G+C (40.4%) in all the sequences.

When compared with the recognition sites reported in some fishes, three domains were detected: the extended termination associated sequence domain, the central conserved sequence block domain and the conserved sequence block domain. As in most fishes, no tandem repeat sequence was found in *A. amoyensis* at the 3′ end.

Moreover, one extended termination associated sequence domain was identified in the control region (Figure 1). A TAS (termination associated sequence) motif-TACAT was found at the 5′ end of the control region. A complementary TAS (cTAS) motif-ATGTA was also detected downstream of the TAS motif.

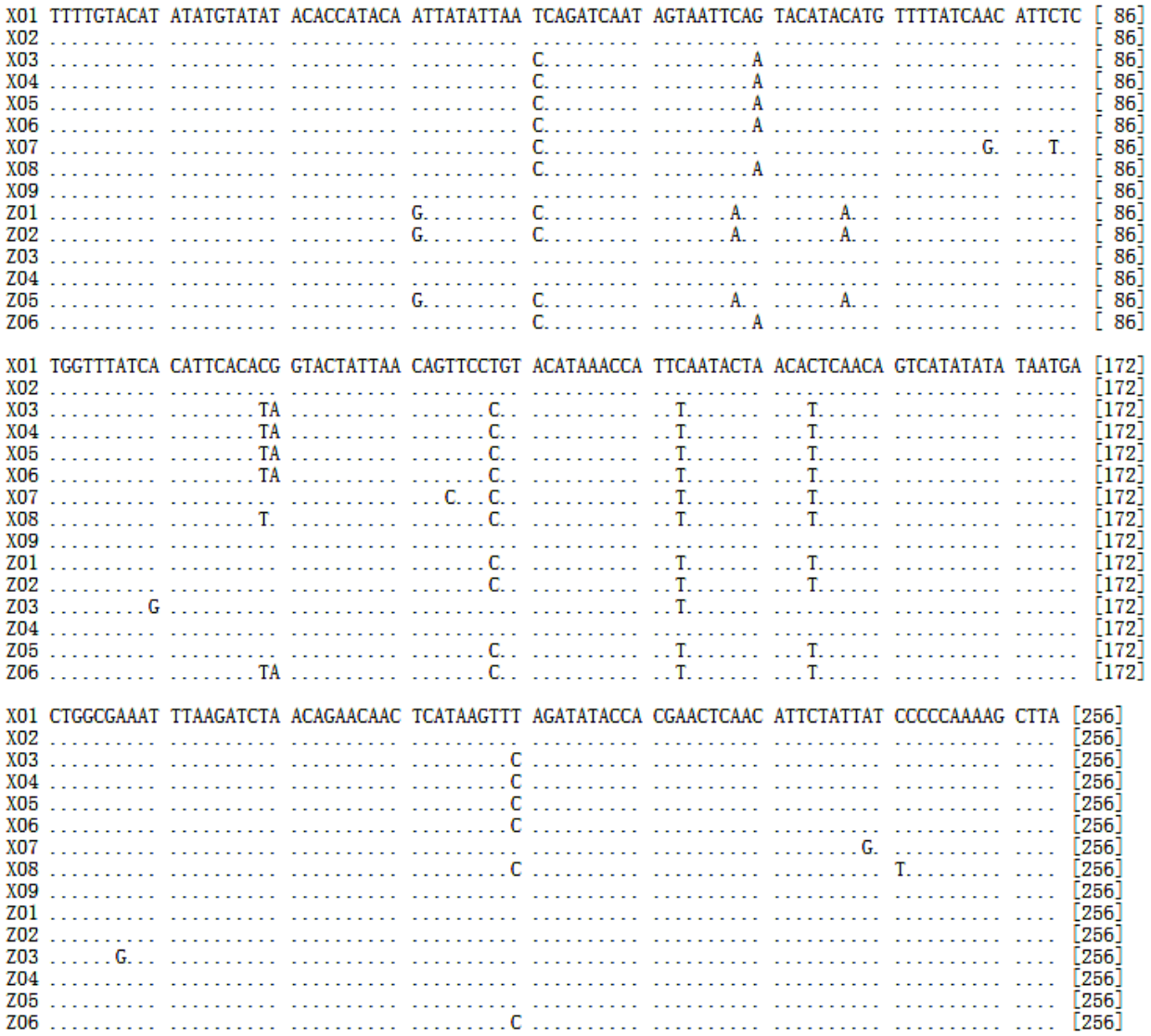

Figure 1. The sequence alignment of the ETAS of *A. amoyensis* (X indicates *A. amoyensis* individuals from Xiamen, while Z indicates *A. amoyensis* individuals from Zhangzhou).

Three of the six central conserved sequence blocks, including CSB-F, CSB-E and CSB-D were detected in the control region. The CSB-F distinguished the extended termination associated sequence domain from central conserved sequence block domain. The identical sequence of the CSB-F was: ATGTAGTAAGA-ACCGACCATCAGTTGATTTCTTAACGCACACGGTTATT GAA-GGTG (Figure 2). CSB-E was located downstream of the CSB-F, which was characterized by the box GTGGGG. The consensus sequence of the CSB-E was AGGGACAAATATCGTGGGGG. The CSB-D followed CSB-E and the consensus sequence was TATTCCTGG-CATTTGGTTCCTA. The CSB-C, CSB-B and CSB-A were not detected after the CSB-D. The sequences of the CSB-F, CSB-E and CSB-D were highly conserved and easily recognized in the central conserved sequence block domain.

In the conserved sequence block domain, three conserved sequence blocks - CSB1, CSB2 and CSB3 of *A. amoyensis* were found after the central conserved sequence block domain in the control region. CSB1 was the first domain of the conserved sequence block domain. The sequences of CSB2 and CSB3 were relatively more conserved than that of CSB1. Interestingly, there were inter-population CSB1 sequence differences, but no intra-population differences. As a result, CSB1 may be used as a tool to identify the wild populations of *A. amoyensis* (Figure 3).

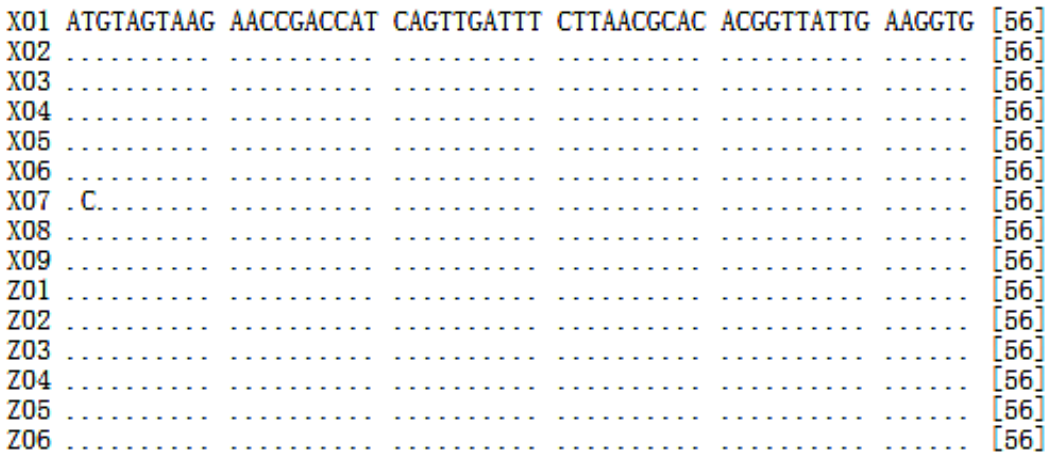

Figure 2. The sequence alignment of the CSB-F of *A. amoyensis* (X indicates *A. amoyensis* individuals from Xiamen, while Z indicates *A. amoyensis* individuals from Zhangzhou).

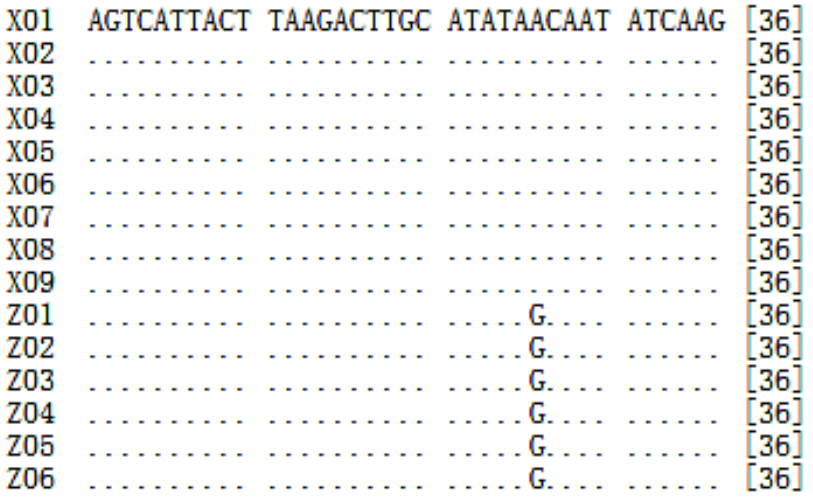

Figure 3. The sequence alignment of CSB 1 of *A. amoyensis* (X indicates *A. amoyensis* individuals of Xiamen, while Z indicates *A. amoyensis* individuals of Zhangzhou).

Phylogenetic relationship revealed by DNA barcoding

Part of the COI gene was used to study the genetic differences and the phylogenetic relationships among six species of Sciaenidae. The length of COI fragments used for the analysis was 652 bp.

Transition and transversion substitutions of the three fragments increased linearly against F84 distance (not given), indicating that basic changes at these sites were not saturated.

The NJ, MP, ML and Bayesian trees were constructed based on the COI sequence dataset. The trees constructed by the different methods were in most cases similar to each other. All the trees indicated that the six clades belong to the six species respectively with high bootstrap values for NJ, MP and ML trees and a high posterior probability for Bayesian tree (Figure 4).

DISCUSSION

As a unique and highly variable area in the mtDNA, the control region is noted for its non-protein-coding and faster rate of evolution. A TAS motif-TACAT which is proposed to act as a sequence-specific signal for the termination of D-loop synthesis was found at the 5′-end of

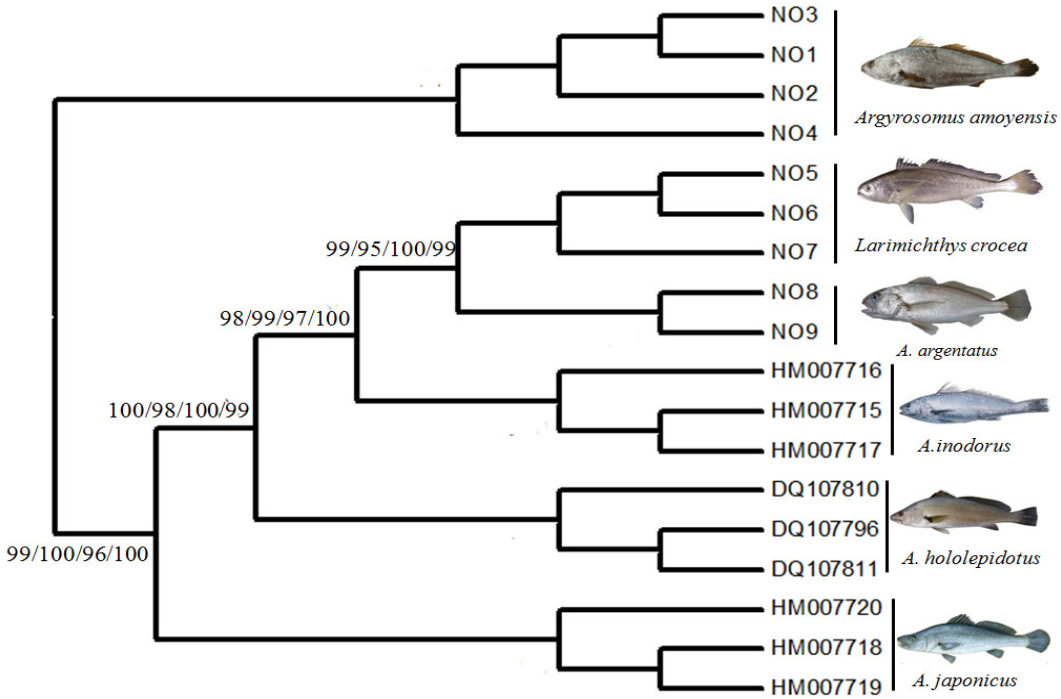

Figure 4. Phylogenetic trees of the six species based on DNA barcodes. The numbers on the branches from left to right are the bootstrap values of NJ, MP and ML trees and posterior probability of Bayesian tree. NJ, Neighbor joining; MP, maximum parsimony; ML, maximum likelihood.

the control region, showing high similarity with other putative TAS elements in other fish species such as *Trachurus japonicus* (Zhu et al., 2007) and lungfish (Zardoya and Meyer, 1996). A complementary TAS (cTAS) motif ATGTA was also detected downstream of the TAS motif in this study. In the central conserved sequence block domain, three of the six CSBs, including CSB-F, CSB-E and CSB-D were detected. Southern et al. (1988) was first to recognize the conserved sequences CSB-B, CSB-C, CSB-D, CSB-E and CSB-F in the central conserved sequence block domain in mammals. This phenomenon was very common in most fishes as only CSB-F, CSB-E and CSB-D could be identified in them (Broughton and Dowling, 1994). However, in 12 fishes of Pleuronectiformes, only CSB-A, CSB-B and CSB-C were found (He et al., 2007; Zhang et al., 2011). These results seem to indicate that the central conserved sequence block domain is relatively conserved in the Pleuronectiformes (He et al., 2007). As in other fishes, CSB-F is the feature that distinguishes the central conserved sequence block domain from the ETAS. Also, the sequences of CSB-F, CSB-E and CSB-D in *A. amoyensis* are highly conserved and consistent with those of other fishes (Zhang et al., 2003).

In association with the initiation of mtDNA duplication, CSB-1 is highly conserved in mammals (Sbisa et al., 1997), but variable in fishes (Liu, 2002; Guo et al., 2003; Zhang et al., 2003). In this study, it was also variable in the two populations of *A. amoyensis*. Since the sequences of CSB1 were the same within each of the populations but had one base-pair difference between the two populations, CSB1 may be used as a tool to identify the populations of *A. amoyensis*. The consensus sequence of CSB-2 and CSB-3 were same, being consistent with other fishes (He et al., 2007).

The structural or functional implication of the control region is mostly determined in these two domains, attributable to the highly conserved sequences in the central and conserved sequence block domains, (Zhao et al., 2006).

In the mtDNA control region of vertebrates, the ubiquitous CSB-3 can be used to determine whether there is any repeat sequence. Through the analyses of the mtDNA of mammals, reptiles, amphibians and fishes, it was established that in the control region of some reptiles, there also exists tandem repeat sequences, but they mainly occur between the ETAS region and the central conserved region such as is found in *Acipenser sinensis* (Zhang et al., 2000), *Phoca vitulina* (Arnason and Johnsson, 1992) and *Bornbina bombina* (He et al., 2007). In this study, we did not find out if all the mtDNA control region sequences of the individuals have tandem repeat sequences in the downstream of CSB-3 as is common with most vertebrates. Though some studies found out that among amphibians, after CSB-3, there are long tandem repeat sequences (He et al., 2007), we did

not get a similar result in this study.

In the present study, part of the COI gene was used to analyze the phylogenetic relationships of the six species. Both the genetic distances based on the segment and the phylogenetic trees based on the different methods indicated that the result supports traditional morphological classification of these six species. Two hundred and seven species of fish, mostly Australian marine fish, were sequenced (barcoded) for a 655 bp region of the COI gene by Ward et al. (2005). Our research, which is based on COI, is in agreement with Ward et al. (2005) who found average within species, genus, family, order and class genetic distances as 0.0039, 0.0993, 0.1546, 0.2218 and 0.2327 respectively. This study, therefore, supports the efficiency of COI barcodes in identifying the six species.

ACKNOWLEDGEMENT

This study was supported by international cooperation program of the Ministry of Science and Technology of China(S2012z10303) and Conservation base construction of large yello croaker program

REFERENCES

Arnason U, Johnsson E (1992). The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. J. Mol. Evol. 34: 493-505.

Broughton RE, Dowling TE (1994). Length variation in mitochondrial DNA of the minnow, *Cyprinella spiloptera*. Genetics. 138: 179-190.

Brown WM, George M, Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. 76: 1967-1974.

Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R, Ikonomi P (2007). Species identification in cell culture: A two-pronged molecular approach. *In Vitro* Cell Dev. Biol. An. 43: 344-351.

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.

Han ZQ, Gao TX, Yanagimato T, Sakurai Y (2008). Genetic population structure of *Nibea albiflora* in Yellow Sea and East China Sea. Fish. Sci. 74:44-552.

Harrison RG (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4: 6-11.

He CB, Cao J, Liu WD, Zhou ZC, Ge LL, Gao XG, Wang XM (2007). Structure analysis of mtDNA control region of spotted halibut (*Verasper Variegatus*) and its related species. Hereditas. 29: 829-836.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 270: 313-321.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004a). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly, *Astraptes fulgerator*. Proc. Natl. Acad. Sci. 101: 14812-14817.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004b). Identification of birds through DNA barcodes. Plos. Biol. 2: 1657-1663.

Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogeny. Bioinformatics. 17: 754-755.

Liu HZ (2002). The structure and evolution of the mtDNA control region in fish: Taking example for Acheilognathinae. Prog. Natural Sci. 12: 266-270.

Lunt DH, Whipple LE, Hyman BC (1998). Mitochondrial DNA variable number tandem repeats (VNTRs): Utility and problems in molecular ecology. Mol. Ecol. 7: 1441-1455.

Posada D, Crandall KA (1998). Modeltest: Testing the model of DNA substitution. Bioinformatics. 14: 817-818.

Pyle RL, Earle JL, Greene BD (2008). Five new species of the damselfish genus *Chromis* (Perciformes: Labroidei: Pomacentridae) from deep coral reefs in the tropical western Pacific. Zootaxa, 1671: 3-31.

Ratnasingham S, Hebert PDN (2007). BOLD: The barcode of life data system (www.barcodinglife.org). Mol. Ecol. Notes. 7: 355-364.

Ronquist F, Huelsenbeck JP (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed model. Bioinformatics, 19: 1572-1574.

Saccone C, Attimonelli M, Sbisa E (1987). Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. J. Mol. Evol. 26: 203-211.

Sambrook J (1989). Molecular Cloning: A Laboratory Manual, 2nd edition. Cold Spring Harbor Laboratory Press, New York.

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.

Shui BN, Han ZQ, Gao TX, Miao ZQ (2008). Tandemly repeated sequence in 5′ end of mtDNA control region of Japanese Spanish mackerel, *Scomberomorus niphonius*. 7: 4415-4422

Southern SO, Southern PJ, Dizon AE (1988). Molecular characterization of a cloned dolphin mitochondrial genome. J. Mol. Evol. 28:32-40.

Starner H, Pahlsson C, Linden M (2004). Tandem repeat polymorphism and heteroplasmy in the mitochondrial DNA control region of three spine stickleback (*Gasterosteus aculeatus*). Behavior. 141: 1357-1369.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599.

Tautz D, Arctander P, Minelli A, Thomas R H and VoglerA P (2002). DNA points the way ahead in taxonomy. Nature, 418:479.

Ward RD, Holmes BH, White WT, Last PR (2008). DNA barcoding of shared fish species from the North Atlantic and Australasia: Minimal divergence for most taxa, but *Zeus faber* and *Lepidopus caudatus* each probably constitute two species. Aquat. Biol. 3: 71-78.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005). DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360: 1847-1857.

Wu DX, Hong WS, Zhang QY (1998). Studies on early development of *Argyrosomus amoyensis*. J Oceanography Taiwan Strait. 17: 149-155.

Zardoya R, Meyer A (1996). The complete nucleotide sequence of the mitochondrial genome of the lungfish (*Protopterus dolli*) supports its phylogenetic position as a close relative of land vertebrates. Genetics. 142: 1249-1263.

Zhang SM, Wu QJ, Zhang YP (2000). Tandem repeats of Chinese sturgeon (*Acipenser sinensis*) and related species and its significance in evolution. Chin. J. Biochem. Mol. Biol. 16: 458-461.

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 Hydrobiolo. Sin. 27: 463-467.

Zhang Y, Zhang H, Gao TX, Miao ZQ (2011). Structure of mitochondrial DNA control region and molecular phylogenetic relationship among three flounders of genus *Pleuronectes*. Biochem. Syst. Ecol. 39: 627-634.

Zhao JL, Wang WW, LI SF, Cai WQ (2006). Structure of the mitochondrial DNA control region of the Sinipercine fishes and their phylogenetic relationship. Acta Genetica Sin. 33: 793-799.

Zhu SH, Zheng WG, Zou JX, Yang FC, Shen XQ (2007). Mitochondrial DNA control region structure and molecular phylogenetic relationship of Carangidae. Zool. Res. 28: 606-614.