African Journal of Biotechnology Vol. 7 (24), pp. 4415-4422, 17 December, 2008
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2008 Academic Journals

*Full Length Research Paper*

# Tandemly repeated sequence in 5'end of mtDNA control region of Japanese Spanish mackerel *Scomberomorus niphonius*

**Bo Nian Shui[1], Zhi Qiang Han[1], Tian Xiang Gao[1,2]*, and Zhen Qing Miao[1]**

[1]Fishery College, Zhejiang Ocean University, Zhoushan, 316004 P. R. China.
[2]The Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao, 266003 P. R. China.

Accepted 24 October, 2008

**Extensive length variability was observed in 5' end sequence of the mitochondrial DNA control region of the Japanese Spanish mackerel (*Scomberomorus niphonius*). This length variability was due to the presence of varying numbers of a 56-bp tandemly repeated sequence and a 46-bp insertion/deletion (indel). The structure and organization of this segment is similar to that of other teleost fish and vertebrates. However, extensive variation in the copy number of tandem repeats (0–5 copies) and the presence of a relatively large (46-bp) indel, are apparently uncommon in teleost fish control region sequences reported to date. The common occurrence of tandem arrays in fish control regions could be related to a stable secondary structure. Based on the frequency distribution of tandem repeat units, no significant geographic heterogeneity (*P*=0.904) among eight populations of Japanese Spanish mackerel was detected by the exact test. Long period of larval stage, strong dispersal ability of adults and long migrations might be responsible for the high genetic connections among populations and reducing the genetic heterogeneity.**

**Key words:** Japanese Spanish mackerel, mtDNA, tandem repeats, control region.

## INTRODUCTION

The mitochondrial genome (mtDNA) has been widely used as a marker for molecular genetic studies because of its compact size (16–17 kb), high rate of mutation, and exclusively maternal mode of inheritance (Brown et al., 1979; Harrison et al., 1989). The piscine mtDNA is a closed circular molecule containing a set of 37 genes specifying 13 proteins, two rRNAs, and 22 tRNAs encoded in both the heavy (H) and light (L) DNA strand reviewed by Meyer (1993). The control region (D-loop) is the only major non-coding segment in the vertebrate mitochondrial genome. It is also the most variable part of the mtDNA and evolves three to five times more rapidly compared with the rest of the mitochondrial genome. In fishes, the mitochondrial control region is located between the $tRNA^{Pro}$ gene and the $tRNA^{Phe}$ gene, and its primary function is usually believed to be the regulation of replication and transcription of the mitochondrial genome.

*Corresponding author. E-mail: gaozhang@ouc.edu.cn.

mtDNA length variability, now recognized as a common feature of animal mtDNA, has been observed in numerous fish species (Billington and Herbert, 1991; Rand, 1993; Lunt et al., 1998). In almost every case, length variability has been shown to be due to the presence of varying numbers of tandem repeats in the control region. Tandem repeated arrays are typically located in the parts of control region at, or proximal to, the 5' end of the D-loop DNA, where replication of mtDNA is initiated (Lee et al., 1995; Nesbø et al., 1998). The tandem repeats, varying in size from 10 bases to 1.5 kb, are highly polymorphic and exhibit copy number variation among individuals and frequently even within individuals (i.e., heteroplasmy). The detection of widespread mtDNA length variation and heteroplasmy is due in large part to the sensitivity of detection provided by the polymerase chain reaction (PCR).

Japanese Spanish mackerel (*Scomberomorus niphonius*), a member of the family Scombridae, is one of the most important fishery resources in China, Japan and Korea. Several preliminary studies related to *S. niphonius*

**Table 1.** Populations of Japanese Spanish mackerel (*Scomberomorus niphonius*) sampled, sample size, and distribution of repeat units documented in this study.

| Population | Date of collection | Sample size | No. of repeat unit | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 4 | 5 |
| Dongying | July 2007 | 20 | 19 | | 1 | | | |
| Weihai | April 2007 | 18 | 18 | | | | | |
| Ganyu | May 2007 | 21 | 17 | | 1 | 1 | 1 | 1 |
| Zhoushan | March 2008 | 19 | 17 | | | 2 | | |
| Ningbo | March 2007 | 22 | 16 | 1 | | 2 | 2 | 1 |
| Wenzhou | May 2008 | 16 | 14 | | 1 | 1 | | |
| Cheju Island | March 2008 | 14 | 10 | | 1 | 3 | | |
| Nagasaki | March 2008 | 22 | 22 | | | | | |
| Total | | 152 | 133 | 1 | 4 | 9 | 3 | 2 |

mainly covered general ecology and fishery biology. No study concerning the population genetic structure of this species is available to date (Seikai National Fisheries Research Institute, 2001). Studies of mtDNA sequence evolution in order Perciformes have revealed length variation in some fishes (Nesbø et al., 1998). However, up to present no tandem repeat sequence is reported in mackerel species.

In the present study, the occurrence of length variation was investigated in 152 individuals of Japanese Spanish mackerel from the Northwestern Pacific. Further, sequencing of the variable region was carried out to reveal the nature of the observed polymorphism in order to provide a basis for the use of the control region for subsequent studies of population-level variability in this commercially valuable species.

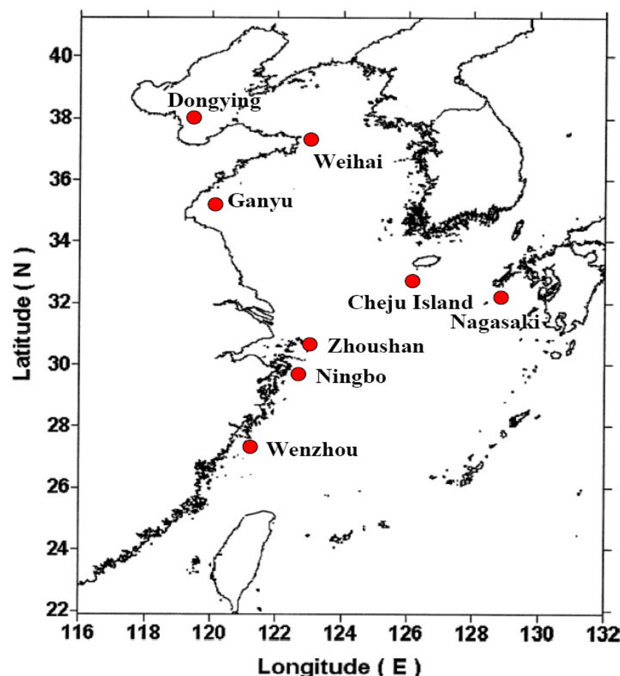



**Figure 1**. Sample sites for Japanese Spanish mackerel

## MATERIALS AND METHODS

### Sampling and sequencing

MtDNA length variation was analyzed in 152 individuals of Japanese Spanish mackerel sampled from eight sites in China, Japan and Korea (Figure 1, Table 1). Total genomic DNA was extracted using a standard proteinase K, phenol–chloroform method with ethanol precipitation.

The first hypervariable segment of the mtDNA control region was amplified with fish primers DL-S and DL-R. The primer sequences are DL-S: 5'- CCC ACC ACT AAC TCC CAA AGC -3' (forward) and DL-R: 5'- CTG GAA AGA ACG CCC GGC ATG -3' (reverse). Each PCR reaction was performed in a volume of 50 μL containing 20-50 ng template DNA, 5 μL of 10 × reaction buffer, 5 μL of $MgCl_2$ (25 mM), 1 μL of dNTPs (10 mM), 10 pM of each primer, and 2.5 units of Taq DNA polymerase (TaKaRa, Dalian, China) in an Eppendorf Mastercycler 5333 (Eppendorf, Hamburg, Germany). Initial denaturation was for 3 min at 94°C, followed 40 cycles of f 45 s at 94°C for denaturation, 45 s at 50°C for annealing, and 45 s at 72°C for extension; and a final extension at 72°C for 10 min. 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. We directly sequenced the different size of PCR fragment. The purified product was used as the template DNA for cycle sequencing reactions performed using BigDye Terminator Cycle Sequencing Kit (ver. 2.0, Applied Biosystems, Foster City, California), and sequencing was conducted on an ABI Prism 3730 (Applied Biosystems, Foster City, CA, USA) automatic sequencer with both forward and reverse primers. The primers used for sequencing were the same as those for PCR amplification.

### Sequence analysis

Sequences were edited and aligned using Dnastar software (DNASTAR, Inc., Madison, USA). Putative tRNA gene regions and conserved sequences blocks or feature segments were localized and verified through alignment with corresponding sequences of other fish species (Kong et al., 2007; Zhu et al., 2007; Lee et al., 1995; Guo et al., 2003). The potential secondary structure of repeat units was analyzed, and an estimation of minimum free energies was made using RNA structure 4.6. Pairwise tests for homogeneity of mtDNA variant frequency distribution between samples were

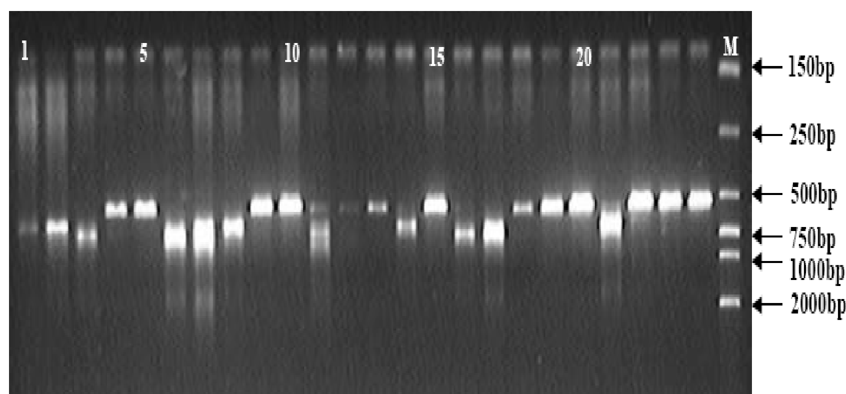



**Figure 2.** Japanese Spanish mackerel mtDNA length variants detected by the polymerase chain reaction.

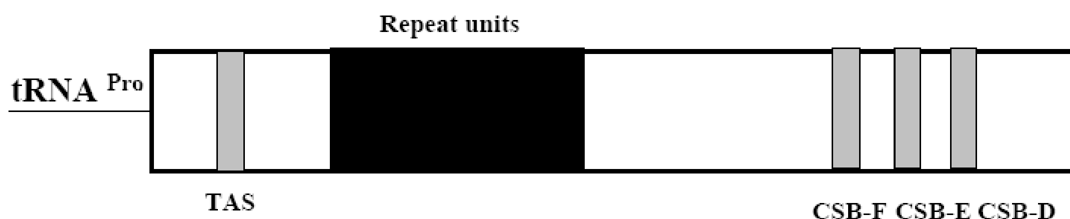



**Figure 3**. Overview of the Japanese Spanish mackerel mtDNA control region and flanking tRNA genes. Conserved sequence blocks and the repeat units are indicated by boxes.

performed using the exact test in Arlequin.

## RESULTS AND DISCUSSION

The segment at the 5' end of the control region was amplified and sequenced from 152 individuals of Japanese Spanish mackerel. The length of this segment ranged from 503 to 783 bp (including a 27 bp partial segment of the $tRNA^{Pro}$ gene and the first hypervariable segment of control region) (Figures 2 and 3). Conserved sequence elements reported in other vertebrate control regions were also observed in Japanese Spanish mackerel (Figures 3 and 4). A TAS (Termination Associated Sequences) motif-TACAT, proposed to act as a sequence-specific signal for the termination of D-loop synthesis, was found in the 5' part of the control region showing high similarity with other putative TAS elements in other fish species, such as *Takifugu fasciatus* (Shao et al., 2007), *Trachurus japonicus* (Zhu et al., 2007), lungfish (Zardoya and Meyer, 1996), percid fishes (Nesbø et al., 1998) and milkfish (Ravago et al., 2002). At the same time, a complementary TAS (cTAS) motif ATGTA was detected just two bp downstream of the tas motif in this study. In the central domain, Lee et al. (1995) reported a CSB-D in some teleost fish. Recently, Guo et al. (2003) identified three CSBs: CSB-D,CSB-E and CSB-F in the complete control region sequence from carps and zebrafish. Based on comparisons with the previous studies, three motifs (CSB-F, CSB-E,CSB-D) have been putatively identified, which are highly similar to the mammalian conserved sequence block-F, block-E and block-D (Southern et al., 1988), with the following sequences: TATTCCTGGCATTTGGTTCCTA (CSB-D), AGGGACAATAATTGTGGGGGT (CSB-E), ATGCAGTAAGAACCTACCA (CSB-F). These elements are apparently highly conserved, and have been observed in control region sequences of nine species of Carangidae. The GTGGG box, commonly found in most teleost fish, has also been identified. The best alignment of the sequences from the region is displayed in Figure 3. This shows that length variation observed in the study was mainly due to changes in the length of the repeat sequences. The repeat region is located at the 5' end of the D-loop near the $tRNA^{Pro}$ gene. Throughout this paper the repeat unit nearest the tRNA Pro gene is referred to as the first repeat. The remaining repeats are referred to as second, third, and consequent repeats. Variations in copy numbers of the tandem repeat account in part for the observed size variation. The shortest control region sequenced (503 bases) contains no copy of the tandem repeat, and the longest control region (783 bases) sequenced contains 5 tandem repeat copies. The size of the repeat unit is 56 bp. However, besides tandem repeat sequences, individuals with the same number of tandem repeats still exhibited length variability, which is accoun-

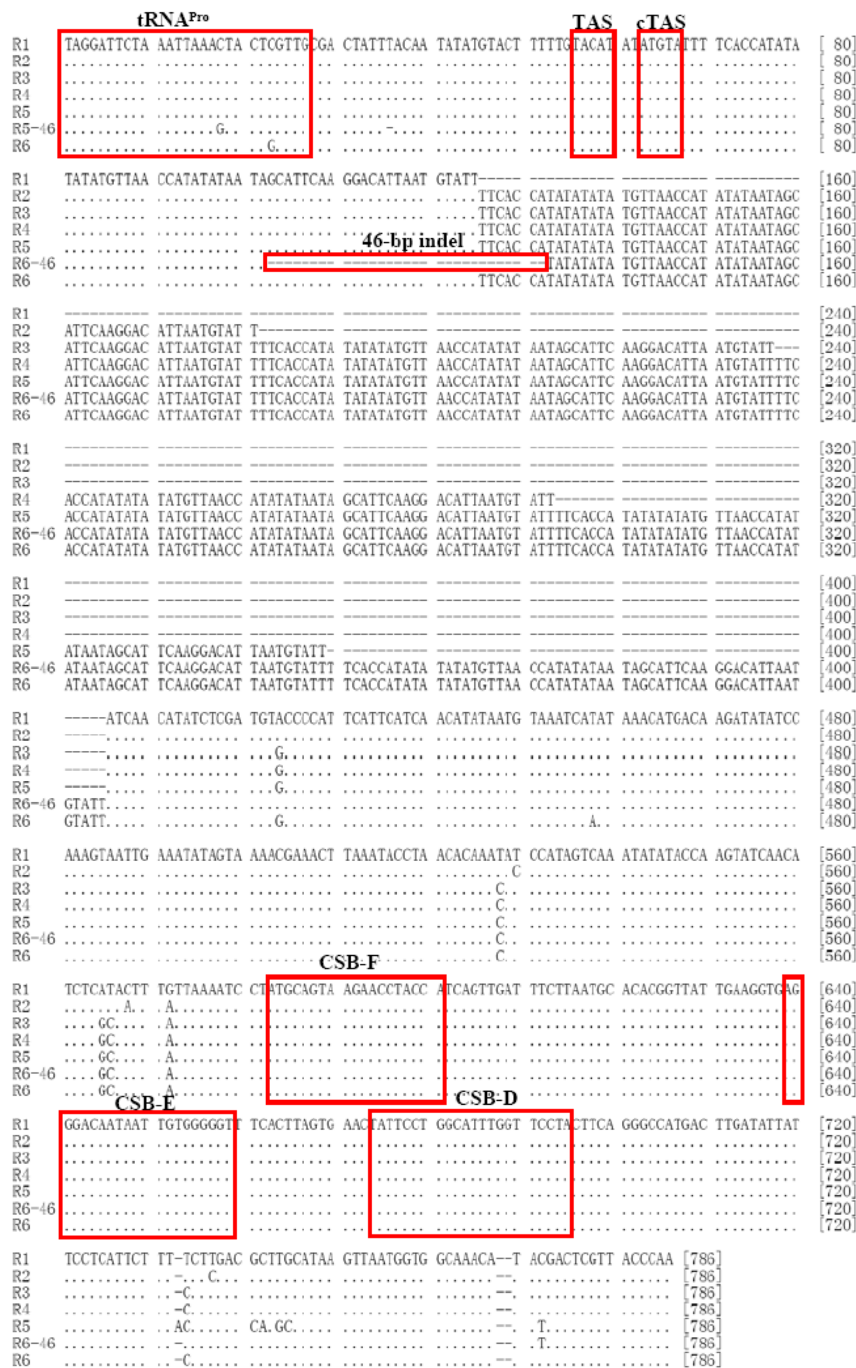
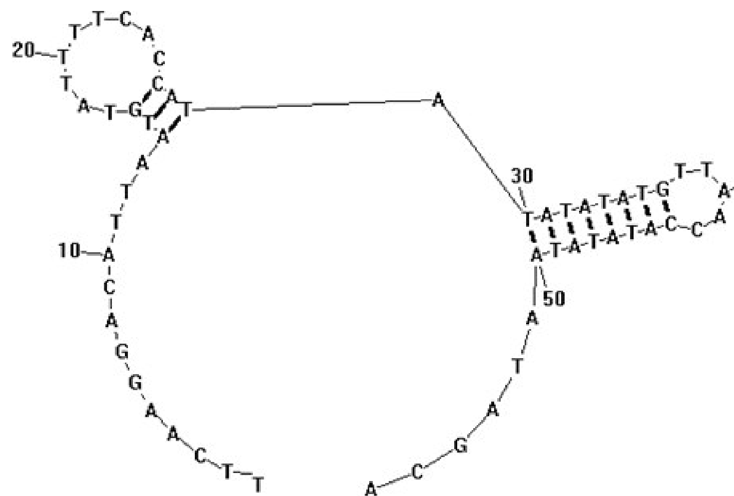
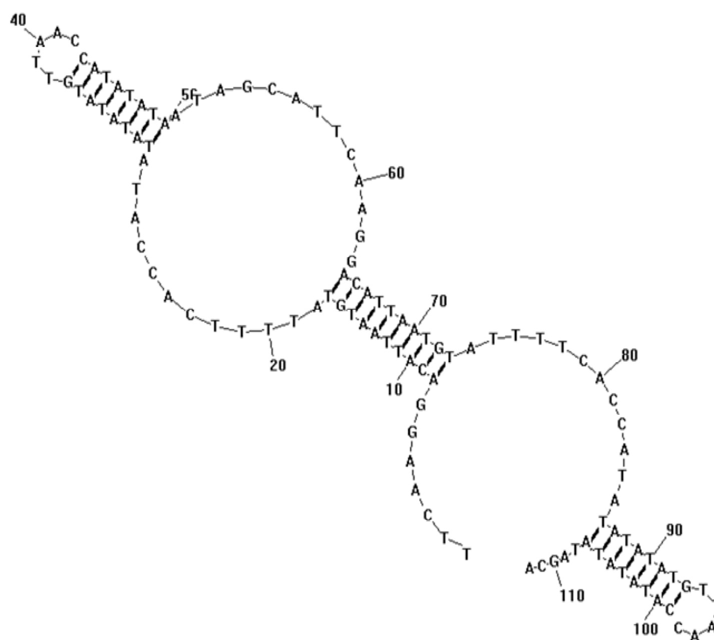


**Figure 4.** Sequence alignment of the 5'-end mitochondrial control region of Japanese Spanish mackerel. The orientation is from 5' to 3' on the light strand. Dots indicate nucleotide identity, dashes indicate deletions. Seven sequences are shown, including individuals with different tandem repeat units (0 to 5) and 46-bp indels.

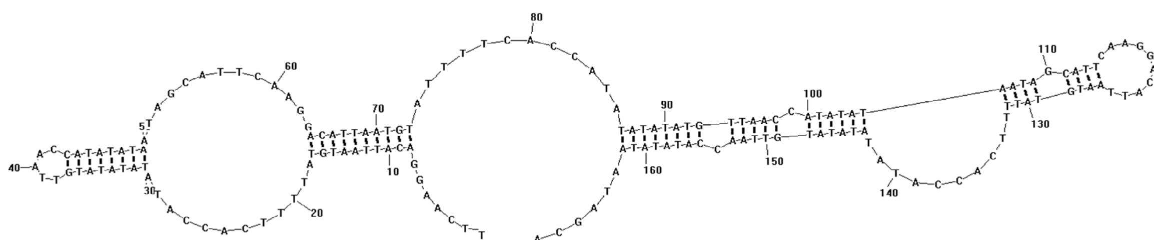
ted for by a 46-bp indel. This large indel is located before the first repeat unit (Figure 4). No large indels are further detected within the central domain. As already described for other species, the repeat elements in this species can also form distinctive secondary structures with free energies increasing linearly with repeat copy number: the 1-,2-,3-,4-,5- and 6-repeat arrays have folding energies of –4.6, -13.2,- 18.7,- 26.8,-32.3,-40.4 kcal/mole respectively as determined by the fold algorithm of RNA structure program (Figure 5). Free energies decrease linearly with increasing repeat copy number, and sequence variants exhibit varying folding stability.



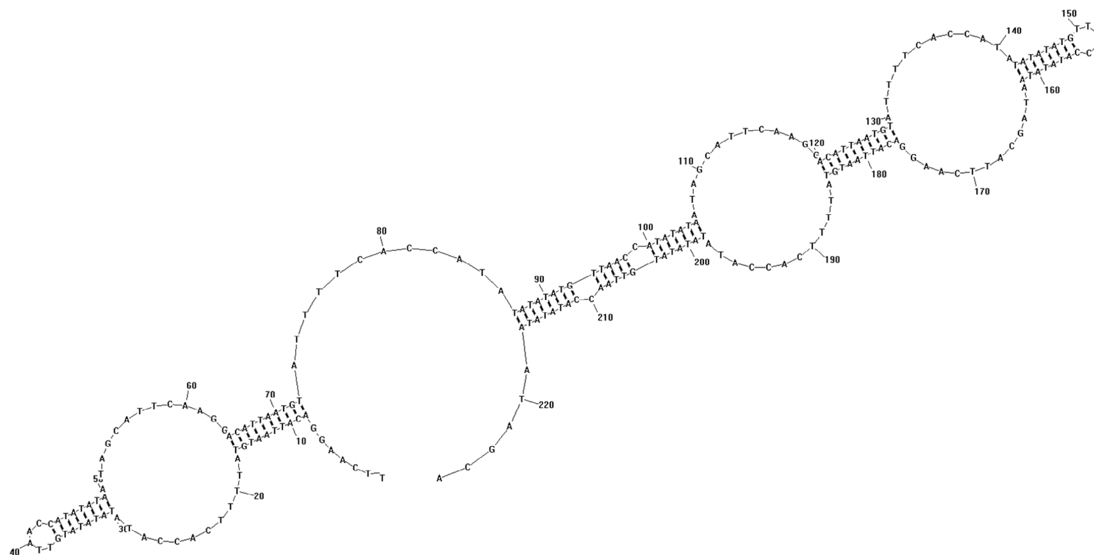
1 repeat



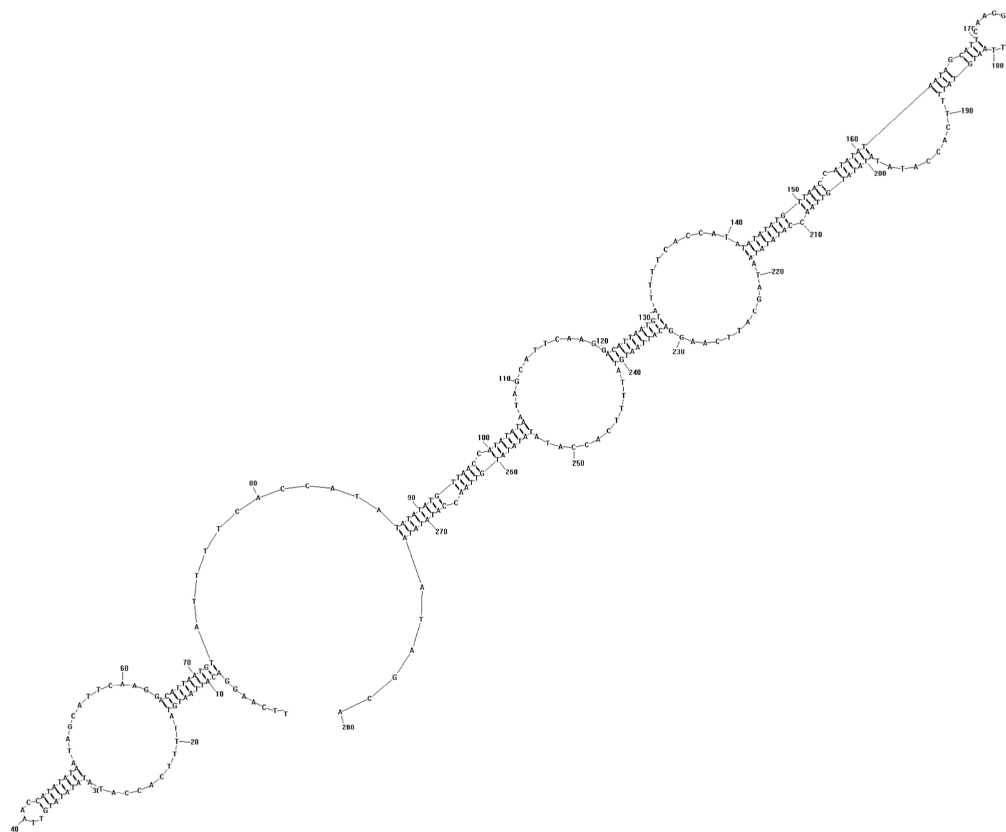
2 repeats



3 repeats



4 repeats



5 repeats

**Figure 5.** Predicted secondary structures of the repeat array with one, two, three, four and five copies of the repeat.

Table 1 gives the frequency distribution of tandem repeat among 8 Japanese Spanish mackerel populations. Among these populations, 19 individuals contain tandem repeat unit (Table 1). An analysis of the mtDNA variant

frequency distribution using the exact test revealed no significant geographic heterogeneity (P=0.904) among eight populations. The no significant geographic heterogeneity among populations of Japanese Spanish mackerel might indicate the high gene flow among populations in the Northwestern Pacific. Pelagic marine fishes generally show low levels of genetic differentiation among geographic regions due to higher dispersal potential during planktonic egg, larval or adult stages, coupled with an absence of physical barriers to movement between ocean basins or adjacent continental margins. Japanese Spanish mackerel has pelagic eggs, long period of larval stage and strong dispersal capacity of adults. Moreover, Japanese Spanish mackerel has a long distant annual migration cycle between the East China Sea and Yellow Sea (Seikai National Fisheries Research Institute, 2001). These biological characteristics might be the reason for facilitating the genetic connections among populations and reducing the genetic heterogeneity.

We have demonstrated the occurrence of a tandem array of 1–6 repeats in the control region of Japanese Spanish mackerel. However, length variability of the control region is only partly attributed to tandem repeat sequence. Individuals with the same number of tandem repeats still exhibited length variability (Figure 4), which is accounted for by a 46-bp indel. Large indels are apparently uncommon in the control region of teleost fish, with only two instances having been reported to date: an 86 bp indel in the control region of the trout *Oncho-rhynchus mykiss* (Nielsen et al., 1998), and a 40-bp indel in the central region of the rockfish *Sebastes capensis* (Rocha-Olivares et al., 1999). MtDNA length variations, caused by tandem repeats, have previously been identified in a number of fish species: several species of sturgeon (e.g., *Acipenser transmontanus*) cod (*Gadus morhua*); European sea bass (*Dicentrarchus labrax*), American shad (*Alosa sapidissima*), whitefish (*Coregonus lavaretus*) and *Coilia ectenes* (Brown et al., 1992; Bentzen et al., 1988; Árnason and Rand, 1992; Cecconi et al., 1995; Brazuzan, 2000; Yang et al., 2008). Like our study, tandem repeats were located in the 5'end of the control region. The precise mechanisms causing mtDNA length variation are not known; and several mechanisms have been proposed to account for the maintenance of length variability of mtDNA, such as intra- and intermolecular recombination (Rand and Harrison, 1989), unequal crossing over or gene conversion (Hoelzel et al., 1993), and strand slippage (Levinson and Gutman, 1987). Strand slippage is a more probable mechanism (Faber and Stepien, 1998), primarily because vertebrate mtDNA does not appear to recombine (Moritz and Brown, 1987). Moritz and Brown (1987) identified three general types of mtDNA size variation: 1) variation in the number of nucleotides in a homopolymer run, 2) variation in copy number of tandemly repeated elements and 3) duplication or deletion of large regions of the molecule. Strand slippage and mispairing has been suggested as a mechanism responsible for the first and second types of variation. The nature of the secondary structure in the repeat array was also an important component of the model presented by Buroker et al. (1990). Sequence analyses of tandem arrays in Japanese Spanish mackerel have revealed that these regions contain motifs capable of forming secondary structures and that these structures may contribute to the slippage and mispairing of strands across entire repeat units. The secondary structure of repeat unit in Atlantic cod was consistent with this conclusion (Árnason and Rand, 1992).

Sequence variability of tandem repeats between individuals has been observed in Japanese Spanish mackerel, and these sequence variation revealed little genetic structuring among 8 populations. The application of length variation as a population-level molecular marker remains to be critically assessed. PCR-based assays, due to their sensitivity and widespread applicability, are increasingly being used to detect length variation and heteroplasmy (Lunt et al., 1998). However, a number of methodological, theoretical, and interpretive factors reviewed in Lunt et al. (1998) need to be addressed before genetic patterns obtained from PCR-based VNTR frequencies can be interpreted with confidence.

## ACKNOWLEDGMENTS

The present study could not have been carried out without the willing help of those listed below in collecting specimens: Mr. Bo Chen, Mr. Ming-Hai Xu, Mr. Hong-Bo Zhao, Mr. Fu-Qing Wang and Dr. Kenji Asano. This work was supported by Zhejiang Provincial Natural Science Foundation of China (Y307496), State 863 High-Technology R&D Project of China (2006AA09Z418) and the National Key Basic Research Program from the Ministry of Science and Technology, P.R. China (2005CB422306).

### REFERENCES

Árnason E, Rand DM (1992). Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. Genetics. 132: 211-220.

Bentzen P, Legget WC, Brown GG (1988). Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (*Alosa sapidissima*). Genetics, 118: 509-518.

Billington N, Hebert PDN (1991). Mitochondrial DNA diversity in fishes and its implications for introductions. Can. J. Fish. Aquat. Sci. 48(Suppl. 1): 80-94.

Brazuzan P (2000). Tandemly repeated sequences in mtDNA control region of whitefish, *Coregonus lavaretus*. Genome 43: 584-587.

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

Brown JR, Beckenbach AT, Smith MJ (1992). Mitochondrial DNA length variation and heteroplasmy in populations of white sturgeon (*Acipenser transmontanus*). Genetics 132: 221-228.

Cecconi F, Giorgi M, Mariottini P (1995). Unique features in the mitochondrial D-loop region of the European sea bass *Dicentrarchus labrax*. Gene 160: 149-155.

Faber JE, Stepien CA (1998). Tandemly repeated sequences in the mitochondrial DNA control region and phylogeography of the pike-

perches Stizostedion. Mol. Phylogen. Evol. 10: 310-322.

Guo XH, Liu SJ, Liu Y (2003). Comparative analysis of the mitocndondrial DNA control region in cyprinids with different ploidy level. Aquaculture 224: 25-38.

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

Hoelzel AR, Hancock JM, Dover GA (1993). Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. J. Mol. Evol. 37:190-197.

Kong XY, Yu JZ, Zhou LS (2007). Comparative analysis of 5'-end sequence of the mitochondrial control region of six flatfish species (Pleuronectidae) from the yellow sea. Raffles B. Zool. 14: 111-120.

Lee WJ, Conroy J, Howell WH, Kocher TD (1995). Structure and evolution of teleost mitochondrial control regions. J. Mol. Evol. 41: 54-66.

Levinson G, Gutman GA (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4: 203-221.

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.

Meyer A (1993). Evolution of mitochondrial DNA in fish. In Biochemistry and Molecular Biology of Fish. Vol. 2. Edited by P.W. Hochachka and P. Mommsen. Elsevier Press, New York. pp. 1-38.

Moritz CTE, Brown WM (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. 18: 269-292.

Nesbø CL, Arab MO, Jakobsen KS (1998). Heteroplasmy, length and sequence variation in the mtDNA control regions of three Percid fish species (*Perca fluviatilis*, *Acerina cernua*, *Stizostedion lucioperca*). Genetics 148: 1907-1919.

Rand D (1993). Endotherms, ectotherms, and mitochondrial genome-size variation. J. Mol. Evol. 37: 281-295.

Rand DM, Harrison RG (1989). Molecular population genetics of mtDNA size variation in crickets. Genetics 121: 551-569.

Rocha-Olivares A, Rosenblatt RH, Vetter RD (1999). Cryptic species of rockfishes (Sebastes: Scorpaenidae) in the southern hemisphere inferred from mitochondrial lineages. J. Hered. 90: 404-411.

Shao AH, Zhu J, Shi QL, Chen K (2007). Characterization and phylogenetic analysis of control region of mitochondrial genome from *Takifugu fasciatus*. J. Fish. Sci. China 14: 352-360.

Seikai National Fisheries Research Institute (2001). Biological and ecological characteristics of valuable fisheries resources from the East China Sea and the Yellow sea –comparison between the Chinese and Japanese knowledges. Seikai National Fisheries Research Institute, Nagasaki.

Southern SO, Southern PJ, Dizon AE (1998). Molecular characterization of a cloned dolphin mitochondrial genome. J. Mol. Evol. 28: 32-40.

Yang JQ, Hu XL, Tang WQ, Lin XD (2008). mtDNA control region sequence variation and genetic diversity of *Coilia nasus* in Yangtze River Estuary and its adjacent waters. Chin. J. Zool. 43: 8-15.

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

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.