DNA Sequence, June 2008; 19(3): 246–255

informa
healthcare

FULL LENGTH RESEARCH PAPER

Sequence and organization of the complete mitochondrial genomes of spotted halibut (*Verasper variegatus*) and barfin flounder (*Verasper moseri*)

CHONGBO HE[1], JIABO HAN[1], LONGLI GE[1,2], ZUNCHUN ZHOU[1], XIANGGANG GAO[1], YUNLEI MU[1], WEIDONG LIU[1], JIE CAO[1,2], & ZHANJIANG LIU[3,†]

[1]*Liaoning Key Laboratory of Marine Fishery Molecular Biology, Liaoning Open Laboratory of Applied Marine Biotechnology, Liaoning Ocean and Fisheries Science Research Institute, 50 Heishijiao Street, Dalian 116023, People's Republic of China,* [2]*College of Life Science, Dalian Fisheries University, Dalian 116023, People's Republic of China, and* [3]*The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Auburn University, Auburn, AL 36849, USA*

(Received 28 January 2007; revised 28 June 2007; accepted 10 July 2007)

Abstract
In this work, the mitochondrial genomes for spotted halibut (*Verasper variegatus*) and barfin flounder (*Verasper moseri*) were completely sequenced. The entire mitochondrial genome sequences of the spotted halibut and barfin flounder were 17,273 and 17,588 bp in length, respectively. The organization of the two mitochondrial genomes was similar to those reported from other fish mitochondrial genomes containing 37 genes (2 rRNAs, 22 tRNAs and 13 protein-coding genes) and two non-coding regions (control region (CR) and WANCY region). In the CR, the termination associated sequence (ETAS), six central conserved block (CSB-A,B,C,D,E,F), three conserved sequence blocks (CSB1-3) and a region of 61-bp tandem repeat cluster at the end of CSB-3 were identified by similarity comparison with fishes and other vertebrates. The tandem repeat sequences show polymorphism among the different individuals of the two species. The complete mitochondrial genomes of spotted halibut and barfin flounder should be useful for evolutionary studies of flatfishes and other vertebrate species.

Keywords: *Flounder, halibut, mtDNA, mitochondrial genome, flatfish, phylogenetic analysis*

Introduction

The study of mitochondrial DNA (mtDNA) has always been a hot field in molecular biology. The keen interest comes from the practical applicability of mtDNA to population genetic and evolutionary studies. With their small size, high abundance in the cell, maternal inheritance and rapid evolutionary rate (Curole and Kocher 1999), mtDNA markers have been successfully applied in a large variety of organisms for the study of the molecular relationships among individuals, populations and species, and the authentication of processed fish products (Chow and Kishino 1995; Mackie et al. 1999). The gene content and organization of vertebrate complete mitochondrial genomes is quite conserved, although gene rearrangements have been described in some species (Boore 1999). mtDNA has been particularly useful as molecular markers in teleost fish that account for over half of the vertebrate species (Yue et al. 2006). This is because of the lack of common nuclear markers across a very large number of species, while mtDNA is conserved enough to allow cross species amplification using a common set of PCR primers. In fact, mitochondrial genomes have demonstrated adequacy

Correspondences: C. He, Liaoning Ocean and Fisheries Science Research Institute, 50 Heishijiao Street, Dalian 116023, People's Republic of China. Tel: 86 411 81787087. Fax: 86 411 84671027. E-mail: hechongbo@hotmail.com
†Tel: 1 334 844 4054. E-mail: liuzhan@auburn.edu

ISSN 1042-5179 print/ISSN 1029-2365 online
DOI: 10.1080/10425170701563303

in resolving persistent controversies over higher-level relationships of teleosts (Miya et al. 2001).

Flatfishes (Pleunectiformes) are highly specialized acanthopterygian fishes containing three suborders, nine families, 118 genera and approximately 538 species existing in the world (Nelson 1994), and in three suborders, eight families, 50 genera and 134 species in China (Wang 1958). Only two complete mitochondrial genomes of flatfish *Paralichthys olivaceous* (Saitoh et al. 2000) and *Platichthys bicoloratus* (Miya et al. 2001) have been reported in the peer-reviewed literature and can be retrieved from GenBank (status on 30 February 2006).

Spotted halibut (*Verasper variegatus*) and barfin flounder (*Verasper moseri*) are two commercially important species of the genus *Verasper*, family Pleuronectidae. In spite of some sporadic research on these species concerning their biological characteristics (Chen et al. 2002) and partial analysis of mitochondrial genes such as Cyt*b*, 16S rRNA, ND5 and COI (Gao et al. 2004), their mitochondrial genomes have not been characterized. The lack of molecular data and genome information about spotted halibut and barfin flounder has hindered the study of their biology. Polymorphic DNA markers are expected to be highly useful tools for the understanding of the biology of the flatfishes.

In this paper we describe the complete mitochondrial genomes of spotted halibut and barfin flounder and compare them with those of other fish species. These mitochondrial genome sequences will be useful for evolutionary studies, as well as for species stock identification of these species.

Materials and methods

Fish samples and DNA isolation

Twenty-four spotted halibut specimens were collected from the coast of Dalian, Northern coast of the Yellow Sea of China. One of the 24 was randomly chosen as the specimen for complete mitochondrial genome sequencing. One individual of barfin flounder was collected from the Hatchery of Liaoning Ocean and Fisheries Institute in Dalian, China for complete mitochondrial genome sequencing. A portion of the musculature was excised and immediately frozen in liquid nitrogen and kept at −70°C. Total genomic DNA was isolated from 500 mg of the tissue using genomic DNA extracting kit (Takara Co.). DNA was isolated following the manufacturer's protocols.

Primer design and PCR amplification

In order to amplify the entire mitochondrial genome of the two species, eight pairs of primers were designed in conserved regions identified by the alignment of mtDNA sequences from the available spotted halibut (GenBank accession no: DQ242488, DQ242490, DQ242494, DQ242492, AY671919), barfin flounder (GenBank accession no: DQ242489, DQ242491, DQ242495, DQ242493, AB207249) (Gao et al. 2004), and related fish species such as *P. bicoloratus* (AB028664, Saitoh et al. 2000; AP002951, Miya et al. 2001), using FastPCR v5.1 software (http://www.biocenter.helsinki.fi/bi/Programs/fastpcr. htm). All the PCR primers were so designed that there are at least 100 bp overlapping between each fragment

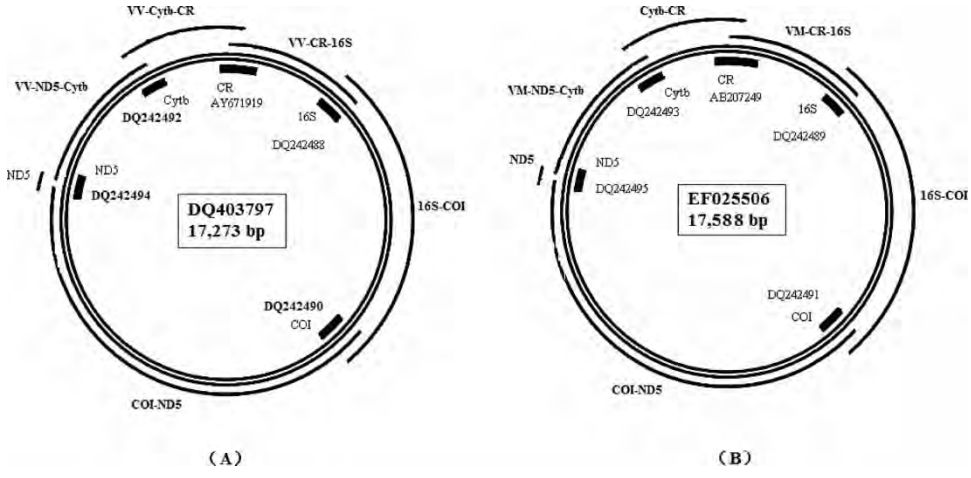

Figure 1. Strategy applied for completing the spotted halibut and barfin flounder mtDNA sequences. (A): represents spotted halibut; (B): represents barfin flounder. VV and VM denote *V. variegatus*, and *V. moseri*, respectively. Dark bars represent available partial mtDNA sequences of these two species from GenBank: DQ242488, DQ242490, DQ242494, DQ242492, DQ242489, DQ242491, DQ242495, DQ242493 (Gao et al. 2004), AY671919 and AB207249. CR indicates control regions, 16S denotes 16S ribosomal RNA genes; COI denotes cytochrome oxidase subunits I; ND5, NADH dehydrogenase subunits 5; Cyt*b*, cytochrome *b*. The arch lines indicate the fragments amplified by PCR using corresponding primers (16S_COI, COI_ND5, ND5_Cyt*b*, ND5, Cyt*b*_CR) and species specific primers (VV-CR-16S,VM-CR-16S, VV-ND5_Cyt*b*, VM-ND5_Cyt*b*), respectively.

Table I. Primers and PCR conditions for amplification of *V. variegatus* (VV) and *V. moseri* (VM).

PCRs	Forward primer 5′–3′	Reverse primer 5′–3′	Length (bp)	Amplification profile (35 cycles)
vm-CR_16S	gtccagtgttcatgcaatggat	aggatgtcctgatccaacatc	4300	96°C1′–60°C1′–72°C3′
vv-ND5_Cyt*b*	cgcattttctacatccagccag	gatgcgccrttggcatgratgct	1750	96°C1′–60°C1′–72°C1′
v-16S_COI	cgcctgtttaccaaaaacatcgcctc	ggtttcggtcygtyagtagyattg	4100	96°C1′–60°C1′–72°C3′
v-COI_ ND5	agccggaatagtggggacaggcc	ctggctggatgtagaaaatgcg	7300	96°C1′–60°C1′–72°C5′
v-ND5_Cyt*b*	cgcattttctacatccagccag	gatgcgccrttggcatgratgct	1750	96°C1′–60°C1′–72°C1′
v-ND5	actctagcaccatagtcgttgc	agtatggctttgaagaaggcgtg	270	96°C1′–60°C1′–72°C1′
v-Cyt*b*_CR	ctccctgccccctctaatatct	actgatgagtgtcgtgttcggt	1500	96°C1′–60°C1′–72°C2′
vv-CR_16S	acgcagtgttcatacgatacgc	aggatgtcctgatccaacatc	4100	96°C1′–60°C1′–72°C4′

Note: vm-CR_16S for *V. moseri* only; vv—for *V. variegatus*; v—for both *V. moseri* and *V. variegatus*.

(Figure 1) to ensure proper sequence assembly, especially in the consideration of low sequence qualities at the beginning of sequencing reactions. The sequences of primers and PCR amplification conditions are shown in Table I.

Six PCR fragments representing the respective entire mitochondrial genomes of spotted halibut and barfin flounder were amplified by normal or long and accurate (LA) PCR using Ex- and LA-*Taq* polymerase (Takara Co.), respectively, according to the manufactor's instructions (using the above primers). Briefly the reactions were carried out with a 25-μl reaction volume containing 16.75 μl sterilized distilled water, 2.5 μl deoxynucleoside triphosphate (dNTP 10 mM), 2.5 μl of 10 × buffer, 1 μl $MgCl_2$ (50 mM), 0.5 μl each primer (10 mM), 0.25 μl (1 U) Ex- and LA-*Taq* DNA polymerase and 1 μl DNA template containing approximately 50 ng DNA. PCR was performed in an Eppendorf Mastercycler gradient thermal cycler using the following thermal profiles for 35 cycles: denaturation at 96°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1–5 min as specified in Table I, with initiation denaturation for 5 min. The resultant PCR products were electrophoresed on 0.7% agarose gel, then the amplified DNA bands were purified from the gel using GenElute Agarose Spin Column (Sigma). The purified DNA fragments were used for PCR product-direct sequencing.

Sequencing analysis

Double-stranded sequencing by PCR product-direct sequencing and primer walking was conducted for these mtDNA fragments using an ABI PRISM 3730 automated DNA sequencer (Applied Biosystems, Foster City, CA) with BigDye Terminator Cycle Sequencing Kit. Complete sequencing of the long PCR segments was achieved by primer walking methods using sequencing primers as listed in Table II. Both heavy (H) and light (L) strand of complete nucleotide sequences for the entire mtDNA genomes of these two species were unambiguously sequenced. The sequences were assembled by using Sequence Analysis v3.4.1 (Applied Biosystems), Seqman V5.05 (DNASTAR) and the Blast Two Sequences program of NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). The locations of the 13 protein-coding genes and the two rRNA genes were determined by alignment and comparison with sequences of other teleost mitochondrial genomes including *P. olivaceous* (AB028664) (Saitoh et al. 2000), *P. bicoloratus* (AP002951) (Miya et al. 2001). Detection of tRNA genes was done with the program tRNAscan-SE (Lowe and Eddy 1997). Only the $tRNA^{Ser\ (ACY)}$ gene was undetectable by the computer program, and was identified visually between the $tRNA^{His}$ gene and the $tRNA^{Leu\ (CUN)}$ gene. We inferred the structures of the control region (CR) based on homology with reported sequences from several fishes and other vertebrates.

Phylogenetic analysis

Phylogenetic analysis was performed using mitochondrial genome of twenty fish species from representatives of Pleuronectiformes, Perciformes, Cypriniformes, Siluriformes, Salmoniformes (Table III). Four other vertebrate species—*Bos taurus* (Ruminantia), *Gallus gallus gallus* (Galliformes), *Chinemys reevesi* (Cryptodira) and *Mantella madagascariensis* (Anura) were used as outgroups. From the mitochondrial sequence data, four different data sets were analysed: (i) concatenated protein-coding and rRNA nucleotide sequences; (ii) concatenated protein-coding nucleotide sequences; (iii) concatenated protein-coding amino acid sequences; (iv) each of the protein-coding amino acid sequences and (v) each of the protein-coding and rRNA nucleotide sequences. The sequence of *ND6* encoded by the L strand was excluded from the analysis, due to the deviating nucleotide and amino acid composition of this gene as compared to those encoded by the H-strand. Third codons of the 12 heavy strand encoded protein-coding genes were excluded from the analysis. Each of the datasets were aligned using clustalX (Thompson et al. 1997) and then analysed by neibor-joining (NJ), minimum evolution (ME) and maximum parsimony (MP) in MEGA 3.1 (Kumar et al. 2004) and bootstrap analysis was performed with 1000 replicates.

Table II. Working primers used for sequencing the mitochondrial genomes of *V. variegatus* (vv) and *V. moseri* (VM).

Primers	Sequences (5′–3′) H strand	Primers	Sequences (5′–3′) L strand
vm-CR_16S-H1	gtccagtgtacatgcaatg	vm-CR_16S-F1	ggatgtcctgatccaacatc
vm-CR_16S-H2	ctgccttccccaggatacc	vm-CR_16S-F2	tggtagtttccaataccctc
vm-CR_16S-H3	cactaatcaaccataagctc	vm-CR_16S-F3	tatcgcctggacttaggg
vm-CR_16S-H4	gcttagccacaccctcaagg	vm-CR_16S-F4	tgtttcttagctttcgtggg
vm-CR_16S-H5	ccgagcaccctgaactgaac	vm-CR_16S-F5	gatgatggtgtgtaggaagg
vm-CR_16S-H6	ggcctaaaagcagccaccttg	vm-CR_16S-F6	aactggtgagcttatggttg
vm-16S_COI-H1	cgcctgtttaccaaaaacatcgcctc	vm-16S_COI-F1	ggtttcggtctgttagtagtattg
vm-16S_COI-H2	ctagtacgaaaggaccgaaaagg	vm-16S_COI-F2	gtcccatagacagaggttcgag
vm-16S_COI-H3	cagcccaagaggccatctgactgg	vm-16S_COI-F3	gatcgcgagggtattgacagtgg
vm-16S_COI-H4	gggcttaggtacaacactcacg	vm-16S_COI-F4	cctacatggtcggggtatgg
vm-16S_COI-H5	ggttttattatctctgggcggc	vm-16S_COI-F5	gagtcctaggctgacttcgtagg
vm-16S_COI-H6	ctcgaccaatcacaaagacatcggc	vm-16S_COI-F6	gtctgaactcagatcacgtaggac
vm-COI_ ND5-H1	agccggaatagtggggacaggcc	vm-COI_ND5-F1	ctggctggatgtagaaaatgcg
vm-COI_ND5-H2	ctcacatcgtagcatactactcagg	vm-COI_ND5-F2	ggctgatagtgatgttgaagg
vm-COI_ND5-H3	tgaatgactttacggctgcccccc	vm-COI_ND5-F3	gttcacccataagattgggcagag
vm-COI_ND5-H4	ccccgtggaatctccaattcgtgttc	vm-COI_ND5-F4	ggtccccgcatttaagcgttctg
vm-COI_ND5-H5	ggtggtcacaagtgagcaattc	vm-COI_ND5-F5	ggctgaggagaaggcaaagtgtg
vm-COI_ ND5-H6	cgtcgtacgagagggcacgtttcaag	vm-COI_ND5-F6	gtggcaacaaagaatgtagcaccg
vm-COI_ ND5-H7	cctctcaacagtacttgccattg	vm-COI_ND5-F7	gttgttaagatagcaacaacggg
vm-COI_ND5-H8	ccaacactcatacttattccaacagcctg	vm-COI_ND5-F8	gtttagttgaggcatgtcgctagg
vm-COI_ND5-H9	ctaaaactaggaggatacggcatg	vm-COI_ND5-F9	ttattaaggcgtgatcgtggaagtg
vm-COI_ND5-H10	caaagtcccttctcccaccgagag	vm-COI_ND5-F10	gcaggacaatatcaagagagg
vm-COI_ND5-H11	caagcagtactctataaccgtg	vm-COI_ND5-F11	gttaggtctacggaggccccggc
vm-ND5_Cyt*b*-H1	cgcattttctacatccagccag	vm-ND5_Cyt*b*-F1	gatgcgccattggcatgaatgct
vm-ND5_Cyt*b*-H2	ccccggagttatgcatcgcc	vm-ND5_Cyt*b*-F2	gttttagcagacgagctcgcggag
vm-ND5_Cyt*b*-H3	cggattagaagcaactgcaac	vm-ND5_Cyt*b*-F3	ggtagcctatgggaacaaag
vm-Cyt*b*-CR-H1	ctccctgccccctctaatatct	vm-Cyt*b*-CR-F1	actgatgagtgtcgtgttcggt
vm-Cyt*b*-CR-H2	gctgcactaacatcccttgccctg	vm-Cyt*b*-CR-F2	ggagcggaggatagcgtaagc
vm-Cyt*b*-CR-H3	gcattagtagctcagtgtgtag	vm-Cyt*b*-CR-F3	ggagaacgtatccaacgaaagc
vv-CR-16S-H1	acgcagtgttcatacgatacgc	vv-CR-16S-F1	aggatgtcctgatccaacatc
vv-CR-16S-H2	tgtcttccccaggataccc	vv-CR-16S-F2	ccaataccctctttgattggggtc
vv-CR-16S-H3	ggtgtttggattatttcaagcatctc	vv-CR-16S-F3	tatcgcctggacttaggggga
vv-CR-16S-H4	cattgattgtaccacacactc	vv-CR-16S-F4	gctttcgtgggttcaggatg
vv-CR-16S-H5	gtaccgcaagggaaagctgaaag	vv-CR-16S-F5	gaggagatgatggcgtgtag
vv-CR-16S-H6	ctcgttaaccccacactggtgtgc	vv-CR-16S-F6	cttgttacatggatactatgtacg
vv-16S-COI-H1	cgcctgtttaccaaaaacatcgcctc	vv-16S-COI-F1	ggtttcggtctgttagtagtattg
vv-16S-COI-H2	caactatttatcggttgggaag	vv-16S-COI-F2	ggagcggaggatagcgtaagc
vv-16S-COI-H3	tgcctaactattggcagc	vv-16S-COI-F3	ggctagaaataagccggttaca
vv-16S-COI-H4	tattacctaaccgcccgcaa	vv-16S-COI-F4	tgtagttcttgagctgactcgc
vv-16S-COI-H5	ctactgcttttacctccgtcgca	vv-16S-COI-F5	cagccaagaaggggttcc
vv-16S-COI-H6	aacttggaggcgtacttgcc	vv-16S-COI-F6	tagcatctgcccgtccgaac
vv-COI-ND5-H1	agccggaatagtggggacaggcc	vv-COI-ND5-F1	ctggctggatgtagaaaatgcg
vv-COI-ND5-H2	cggaataatttctcacatcgtagc	vv-COI-ND5-F2	tggaccaagtaacgtaaagggc
vv-COI-ND5-H3	cgagaagttagcgcagtag	vv-COI-ND5-F3	gagggaataaccagctgta
vv-COI-ND5-H4	gaccaccgaatagtaatccccgta	vv-COI-ND5-F4	aagataaggtgccggatg
vv-COI-ND5-H5	cggtggtcacaagtgagcaattc	vv-COI-ND5-F5	cttcaaggcataatagggcgg
vv-COI-ND5-H6	acatttcagggacaccacactc	vv-COI-ND5-F6	atggtggagaatttggcgtaac
vv-COI-ND5-H7	cctctcaacagtacttgccattg	vv-COI-ND5-F7	gtaggtaaaggcttaggag
vv-COI-ND5-H8	attccaacagcctggttgctc	vv-COI-ND5-F8	ttgtgtgctttgggaggtgg
vv-COI-ND5-H9	gtacggcataatccgcataatgac	vv-COI-ND5-F9	gctgggagaactgttcag
vv-COI-ND5-H10	caccgagagaggctcgcagcaac	vv-COI-ND5-F10	gccagacctaggaagtgttgg
vv-COI-ND5-H11	cgctgcactccaagcagtgctctat	vv-COI-ND5-F11	cagtagcattgtaattcctgcggc
vv-ND5-Cyt*b*-H1	cgcattttctacatccagccag	vv-ND5-Cyt*b*-F1	gatgcgccgttggcatggatgct
vv-ND5-Cyt*b*-H2	gcttcttccccggggttatg	vv-ND5-Cyt*b*-F2	cagcgttagctgccgagccgtat
vv-ND5-Cyt*b*-H3	atacggctcggcagctaac	vv-ND5-Cyt*b*-F3	actagttaatgcggcgagctctatg
vv-Cyt*b*-CR-H1	ctccctgccccctctaatatct	vv-Cyt*b*-CR-F1	actgatgagtgtcgtgttcggt
vv-Cyt*b*-CR-H2	cttgctgcactagcatcccttgc	vv-Cyt*b*-CR-F2	tcgaaaacaaaagggcaagtacg
vv-Cyt*b*-CR-H3	gcagtgttcatacgatacgc	vv-Cyt*b*-CR-F3	acgaggaggaggaggacaactcc

Table III. Twenty fishes and four other vertebrate species used in the phylogenetic analysis.

Order	Family	Species	Accession numbers
Pleuronectiformes	Bothidae	*P. olivaceous*	AB028664
		P. bicoloratus	AP002951
	Pleuronectidae	*V. variegatus*	DQ403797*
		V. moserri	EF025506*
		T. japonicus	AP003092
Perciformes	Carangidae	*T. trachurus*	AB108498
		C. melampygus	AP004445
		C. armatus	AP004444
Cypriniformes	Cyprinidae	*C. carpio*	X61010
		C. cuvieri	AB045144
	Cobitidae	*C. lacustre*	M91245
		L. echigonia	AB054126
	Lctaluridae	*I. punctatus*	AF482987
Siluriformes	Bagridae	*P. tokiensis*	AB054127
	Callichthyidae	*C. rabauti*	AB054128
Salmoniformes	Bathylagidae	*B. ochotensis*	AP004101
	Salangidae	*S. microdon*	AP004109
		O. mykiss	L29771
	Salmonidae	*S. salar*	U12143
		C. lavaretus	AB034824
Ruminantia	Bovidae	*B. taurus*	AY526085
Galliformes	Phasianidae	*G. gallus gallus*	AP003322
Anura	Rhacophoridae	*M. madagascariensis*	AB212225
Cryptodira	Batagurinae	*C. reevesi*	AY676201

Results and discussion

Gene content and genome organization

The spotted halibut and barfin flounder mitochondrial genomes were sequenced and analyzed and the entire mitochondrial genome was 17,273 and 17,588 bp in length, respectively. Their sequences were deposited and in the GenBank with accession numbers of *DQ403797* and *EF025506*.

The overall base composition of the heavy strand (H-strand) is adenosine (A): 26.2%, cytosine (C) 30.5%, guanine (G) 16.1% and thymine (T) 27.2% for spotted halibut and A: 28.0%, C: 28.4%, G: 16.6% and T: 27.0% for barfin flounder. As observed in other fish species, nucleotide cytosine occurs most often and base guanine the least often. This base composition is very similar to those reported for Japanese flounder *Paralichthys olivaceus* (A: 27.4%, C: 29.6%, G: 16.8% and T: 26.1%) (Saitoh et al. 2000).

Overall, the structure of the spotted halibut and barfin flounder mitochondrial genomes are very similar to that of other teleost species (Saitoh et al. 2000; Miya et al. 2001; Yue et al. 2006). They contain two rRNA genes (12S and 16S rRNA), 22 tRNAs, and 13 protein-coding genes: seven subunits of the NADH dehydrogenase complex (*ND1-6* and *ND4L*), three subunits of the cytochrome *c* oxidase complex (*COX I–III*), one cytochrome *b* (Cyt*b*), two subunits of ATPase (*ATP6* and *ATP8*). The non-coding control regions situated between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes contain the heavy strand origin of replication (O_H). A smaller CR containing the putative light strand origin of replication (O_L) was found between $tRNA^{Asn}$ and $tRNA^{Cys}$ genes, the same order as reported in the other fish species. Among these genes only *ND6* and eight tRNAs are encoded in the L-strand while all the others are encoded by the H-strand (Table IV).

Eleven potential overlaps between genes have been observed in the mitochondrial genomes of the two species. The longest one (10 bp) involving the two ATPase genes appears to be common in most vertebrate mitochondrial genome, and its size in fish (7–10 bp) (Meyer 1993; Broughton et al. 2001) is smaller than that in mammals (40–46 bp). The second largest overlap is 7 bp long, (between *ND4* and *ND4L* genes), whereas the remaining nine were in the size range of 1–5 bp.

Mitochondrial control region

Spotted halibut and barfin flounder mtDNA's heavy strand CR, also known as D-loop, contain O_H and are 1572 bp and 1889 bp long, respectively, flanked by $tRNA^{Pro}$ and $tRNA^{Phe}$. The sizes of these CRs are longer than the typical teleost mitochondrial genomes. The CR can be divided into three different domains (Anderson et al. 1981; Brown et al. 1986) (Figure 2). Domain I contains the extended termination associated sequence (ETAS). It is 295 bp long and contains several "ACAT" motifs. This region was conserved among flatfishes (Lee et al. 1995; Yu et al. 2005), indicating its functional importance. Domain II, commonly known as the central conserved blocks (CSB-A,B,C,D,E,F), is highly conserved within flatfishes. In domain III, three conserved sequence blocks

Table IV. Characteristics of *V. variegatus* (vv) and *V. moseri* (vm) mitochondrial genomes.

Names of gene	Location vv (vm)	Size bp vv/vm	Differences of nt (%)	Transcripted from
Control region	15,702–17,273 (15,700–17,588)	1572/1889	17.96	H non-coding
tRNA-Asp	7145–7215 (7144–7214)	71/71	13.17	H
ND6	13823–14344 (13822–14343)	522/522	8.83	L
Cyt*b*	14,418–15,558 (14,416–15,556)	1141/1141	8.32	H
ND5	11,989–13,827 (11,988–13,826)	1839/1839	7.15	H
ATPase6	8146–8828 (8145–8828)	683/683	6.50	H
ND2	4069–5114 (4068–5113)	1046/1046	6.49	H
WANCY	5331–5366 (5330–5365)	36/36	6.47	L non-coding
ND4	10,394–11,774 (10,393–11,773)	1381/1381	6.28	H
COII	7223–7913 (7222–7912)	691/691	6.22	H
ND1	2880–3854 (2879–3853)	975/975	5.61	H
ATPase8	7988–8155 (7987–8154)	168/168	5.22	H
tRNA-Cys	5367–5431 (5367–5430)	65/65	5.09	L
ND3	9686–10034 (9685–10033)	349/349	4.90	H
tRNA-Phe	1–68 (1–68)	68/68	4.84	H
COIII	8829–9613 (8828–9612)	785/785	4.71	H
COI	5500–7056 (5499–7055)	1557/1557	4.20	H
ND4L	10,104–10,400 (10,103–10,399)	297/297	3.57	H
tRNA-Glu	14,345–14,413 (14,344–14,411)	69/68	3.07	L
tRNA-His	11,775–11,844 (11,774–11,843)	70/70	3.04	H
tRNA-Trp	5115–5186 (5114–5185)	72/72	2.94	H
tRNA-Leu (*CUN*)	11916–11988 (11915–11987)	73/73	2.90	H
tRNA-Asn	5258–5330 (5257–5329)	73/73	2.90	L
16S rRNA	1090–2805 (1089–2804)	1716/1716	1.56	H
12S rRNA	69–1017 (69–1016)	949/948	1.52	H
tRNA-Arg	10,035–10,103 (10,034–10,102)	69/69	1.50	H
tRNA-Met	4000–4068 (3999–4067)	69/69	1.49	H
t*RNA-ser* (*UGA*)	7060–7130 (7059–7129)	71/71	1.45	L
tRNA-Val	1018–1089 (1017–1088)	73/73	1.43	H
TRNA-Gly	9614–9685 (9613–9684)	72/72	1.43	H
tRNA-Thr	15,559–15,631 (15,557–15,629)	73/73	1.41	H
tRNA-Leu (*UAA*)	2806–2879 (2805–2878)	74/74	1.39	H
tRNA-Ser (*AGY*)	11,845–11,911 (11,844–11,910)	67/67	0	H
tRNA-Pro	15,631–15,701 (15,629–15,699)	71/71	0	L
tRNA-Ile	3860–3930 (3859–3929)	71/71	0	H
tRNA-Gln	3930–4000 (3929–3999)	71/71	0	L
tRNA-Ala	5188–5256 (5187–5255)	69/69	0	L
tRNA-Tyr	5432–5499 (5431–5497)	68/68	0	L
tRNA-Lys	7914–7986 (7913–7985)	73/73	0	H

*The genes are listed by descending differences of nt (%) to show the difference between the two species.

(CSB-1,2,3) were identified by similarity comparisons with fishes and other vertebrates (Southern et al. 1988; Broughton and Dowling 1994; Lee et al. 1995; Sbisa et al. 1997; Randi and Lucchini 1998; Zeng and Liu 2001; Liu 2002; Liu et al. 2002; Zhang et al. 2003; Guo et al. 2004; Yu et al. 2005).

Repeats in the heavy strand control region

The 3′ region along the H-strand just downstream of the CSB-3 contains an array of tandem repeats containing eight (for spotted halibut) and thirteen (for barfin flounder) complete copies of a 61 bp unit and a 16 bp truncate copy. The presence of tandem repeats made the CR of spotted halibut (1572 bp) and *V. moseri* (1889 bp) longer than any other reported fishes. Without the repeat cluster, the size of CR is about 1 kb in fish, similar to that in other vertebrates.

PCR analysis with primer pair encompassing the tandem repeat confirmed this cluster is variable in length. In 24 individuals of spotted halibut studied, four types of PCR products were amplified with different lengths. In order to verify the number of tandem repeat cluster, we sequenced these four types of PCR products and obtained four haplotypes with one repeat copy difference in each sequence. The maximum repeat number observed is eight from 12 individuals and the minimum repeat number is five from 3 individuals. This indicated that the length polymorphism was caused by variations in repeat numbers, similar to that in microsatellites. Of the 24 individuals of spotted halibut, the copy number of the tandem repeats varied from five to eight copies. Similarly, of the 17 individuals of barfin flounder, the copy number of the tandem repeats varied from 8 to 14 copies (data not shown).

PCR products from primer pairs outside of the repeat cluster revealed no detectable size variants in each species. These data indicated that the tandem repeat array is the major source of length variation

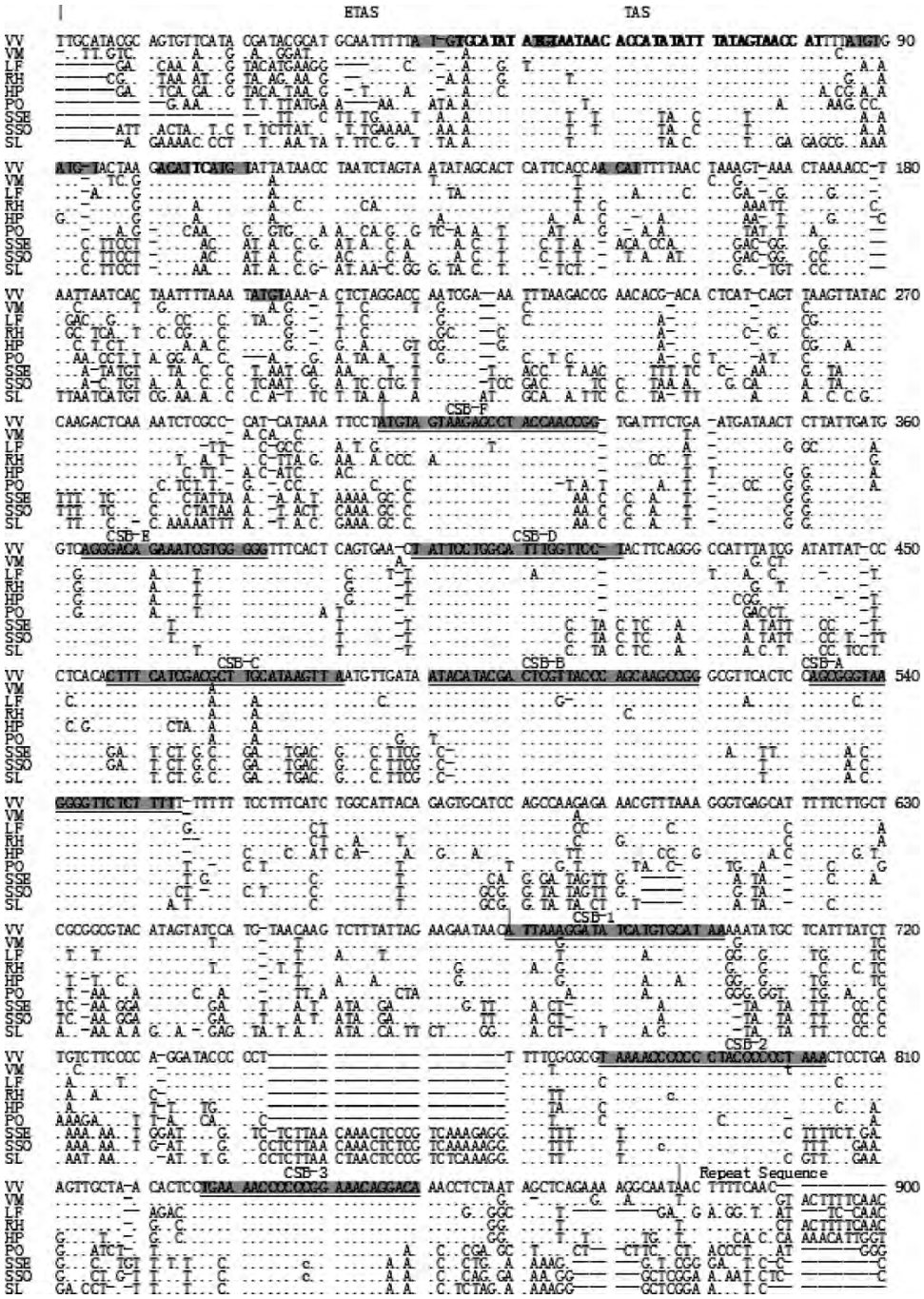

Figure 2. The structure of conserved sequence in CR of nine species of Pleuronectoidei. Vertical bars "|"above the sequence indicate the division of regions; the blocks show the conserved sequence boxes. In the ETAS, bold letters represent sequences of TAS, the sequences with shadow are conserved sequences and their reverse complementary sequences; CSB-F,-E,-D,-C,-B,-A and CSB-1,-2,-3 represent the consensus sequences of F,E,D,C,B,A in central conserved domain and consensus sequences of I, II, III in conserved sequences domain. This figure did not show the tandem repeat regions of the CR. VV (*V. variegatus*), VM (*V. moseri*), LF (*Limanda feruginea*), RH (*Reinhardtius hippoglossoides*), HP (*Heppoglossoides platessoides*), PO (*P. olivaceus*), SSE (*Solea senegalensis*), SSO (*Solea solea*), SL (*Solea lascari*).

in the mtDNA of spotted halibut. The presence of the tandem repeat sequences in mitochondrial CR appeared to be a characteristic of the pleuronectoid fishes such as spotted halibut, barfin flounder, winter flounder and bastard halibut (Lee et al. 1995; Saitoh et al. 2000). The analysis of repeated segment patterns in mitochondrial CR from vertebrate species including mammals, birds, reptiles, amphibians and fishes was conducted by comparing with the common CSB-3 sequence and it was found that all the pleuronectoid fishes and the amphibian anura species such as *Alytes obstetricans pertinax* (AY585337), *Bombina orientalis* (AY585338), *Bombina bombina* (AY458591), *M. madagascariensis* (AB212225),

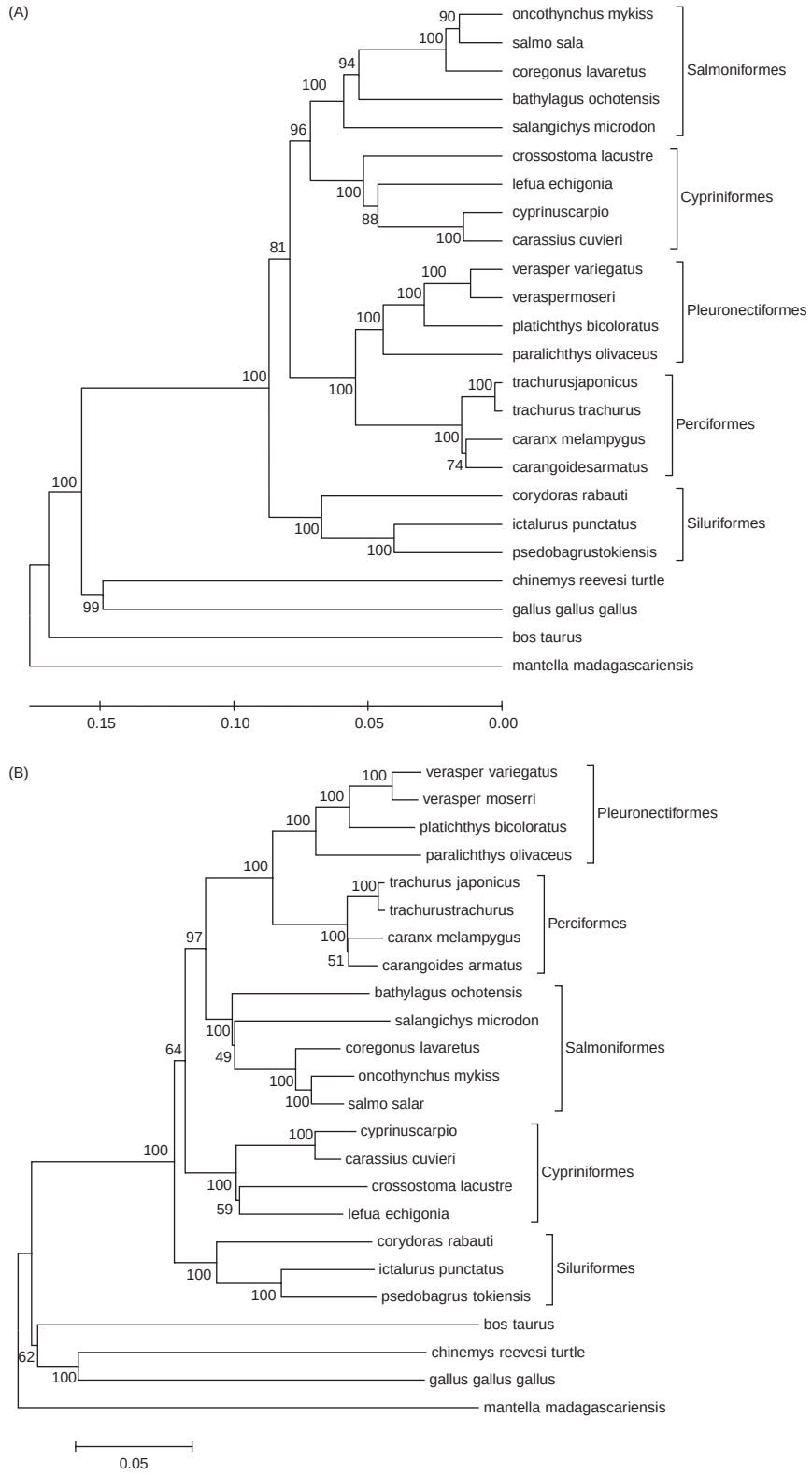

Figure 3. Phylogenetic position of the spotted halibut and barfin flounder relative to 18 other fish species and four vertebrates. (A): MP tree; (B): NJ tree. Legend all mitochondrial sequences are available in the GenBank database. The taxon positions and the accession numbers for these species were shown in Table III. *Oncothynchus mykiss* (L29771), *Salmo salar* (U12143), *Coregonus lavaretus* (AB034824), *Bathylagus ochotensis* (AP004101), *Salangichys microdon* (AP004109), *Crossostoma lacustre* (M91245), *Lefua echigonia* (AB054126), *Cyprinus carpio* (X61010), *Carassius cuvieri* (AB045144), *P. olivaceous* (AB028664), *P. bicoloratus* (AP002951), *V. variegatus* (DQ403797), *Verasper moserri* (EF025506), *Trachurus japonicus* (AP003092), *Trachurus trachurus* (AB108498), *Caranx melampygus* (AP004445), *Carangoides armatus* (AP004444), *Ictalmus punctatus* (AF482987), *Psedobagrus tokiensis* (AB054127), *Corydoras rabauti* (AB054128), *B. taurus* (AY526085), *G. gallus gallus* (AP003322), *M. madagascariensis* (AB212225) and *C. reevesi* turtle (AY676201).

Rhacophorus schlegelii (AB202078), *Buergeria buergeri* (AB127977), *Amphibia Anura_Rana nigromaculata* (AB043889) have similar repeated patterns in their mitochondrial control regions. The functions of these repeated structures are unknown at present.

Phylogenetic analysis

Phylogenetic trees constructed from various data sets using different methods (NJ, ME and MP) showed little variations, indicating that variation mainly originated from the type of data and not the method used (data not shown). Both the concatenated protein-coding nucleotide and amino acid sequence data sets allowed proper placement of the spotted halibut and barfin flounder and other fish species in the phylogenetic trees with high bootstrap support values (Figure 3). This suggested that complete mitochondrial genome can provide a reliable tool for phylogenetic studies of fish and other species (Yue et al. 2006). On the other hand, trees generated using each of the nucleotide and amino acid sequences data sets from 12 protein-coding genes and two rRNA (12S, 16S rRNA) genes were not consistent in placing the species in the proper clades with high bootstrap values. Of the 26 trees (figures not shown) constructed, the 16S rRNA gene was the most appropriate gene for reconstructing the more reliable phylogenetic trees in different tree-building methods or algorithms, followed by *ND2* and *ND5* genes. Genes *ATP8*, *ATP6*, *ND4L*, *ND3* and *ND1* were the least appropriate for reconstructing the reliable phylogenetic trees. This result is consistent with the results obtained by Russo et al. (1996) and Zardoya and Meyer (1996), with some exceptions. Russo et al reported that ND5, Cyt*b* and ND4 were the most reliable genes, and that COII, ND1 and ND4L were the poorest genes for phylogeny construction in vertebrates (using two whale species, cow, rat, mouse, opossum, chicken, frog, and three bony fish species). Zardoya and Meyer (1996) reported that ND4, ND5, ND2, **C**ytb, and COI were the best genes and ATP6, ND3, ATP8 and ND4L were the poorest genes for phenlogeny construction when investigating two known phylogenies of tetrapods and mammals. These results suggested that the different genes have different evolutionary rates, even the same gene in different organisms or lineages may have different evolutionary rates. The phylogenetic performance of these mitochondrial genes may depend on various factors such as the density of lineage creation events in time, the phylogenetic depth of the question, lineage-specific rate heterogeneity, and the completeness of taxa representation (Zardoya and Meyer 1996).

We also examined the differences between mtDNAs of spotted halibut and barfin flounder with those of other species by comparing the 13 protein-coding genes. The overall differences among these mitochondrial sequences are presented in Table V. The average amino acid sequence difference varied, at a rate of 3.16% within genera, 6.52% within family, and 8.99% within order.

Conclusion

This work generated the complete mitochondrial genome sequences for two flatfishes, the spotted halibut and barfin flounder. Although the mitochondrial genomes of spotted halibut and barfin flounder are similar to those of other teleosts and vertebrate mitochondrial genomes in terms of length, gene content, and gene order, they exhibit a number of interesting characteristics. Among them the most interesting is the presence of a long tandem repeat at the ends of domain III in the CR. As length polymorphism can be readily detected among individuals in the

Table V. Differences of amino acid sequences between the mtDNA protein-coding genes of *V. variegatus* (vv), *V. mosori* (vm) and other related species.

Names of mitochondrial gene	Difference of aa (%), vv/vm within genera	Names of mitochondrial gene	Difference of aa (%), vv/pb within family	Names of mitochondrial gene	Difference of aa (%), vv/po within order
ND2	8.79	ATP6	18.50	ND5	18.95
ATP8	6.06	ATP8	12.73	ATP8	16.36
Cyt*b*	5.23	ND5	11.44	ND2	12.64
ND5	4.16	ND2	6.61	ATP6	12.33
ATP6	3.92	ND3	6.03	ND3	9.48
ND3	3.28	COII	5.65	ND6	9.25
ND6	2.49	ND6	5.20	ND4	7.83
COII	1.95	ND4	4.35	Cyt*b*	7.11
ND4	1.72	Cyt*b*	4.21	ND1	7.10
ND1	1.51	COIII	3.83	COII	6.52
COIII	1.16	ND1	3.40	COIII	4.21
COI	0.82	COI	1.74	COI	4.07
ND4L	0.00	ND4L	1.02	ND4L	1.02
Mean	3.16	Mean	6.52	Mean	8.99

Note: vv, *Verasper variagatus*; vm, *V.moserri*; pb, *P. bicoloratus*; po, *P. olivaceous*. The genes are listed by descending difference of amino acids.

tandem repeat region. The tandem repeats could be potentially useful for the analysis of genetic diversity of populations, as well as phylogenetic and phylogeographic studies of the pleuronectidae family and possibly other members of pleuronectiformes order. As mtDNA has been demonstrated to be reliable for the reconstruction of phylogeny, the complete mitochondrial genomes of spotted halibut and barfin flounder should be useful for evolutionary studies, and for stock identification of these and their related species.

Acknowledgements

This work was supported by the Science and Technology Advance Project of the State Oceanic Administration and Natural Science Foundation of China (No. 30571410).

References

Anderson S, Bankier AT, Barrell BG. 1981. Sequence and organization of the human mitochondrial genome. Nature 290: 457–465.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780.

Broughton RE, Dowling TE. 1994. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics 138(1): 179–190.

Broughton RE, Milam JE, Roe BA. 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. Genome Res 11:1958–1967.

Brown GG, Gadaleta G, Pepe G, Saccone C, Sbisa E. 1986. Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. J Mol Biol 192: 503–511.

Chen SQ, Yu DX, Ma AJ, Liu XZ. 2002. A study on biological characteristics of spotted halibut, *Verasper variegatus* (Temminck et Schegel). Mod Fish Inform 17(10):25–27.

Chow S, Kishino H. 1995. Phylogenetic relationships between tuna species of the genus Thunnus (Scombridae: Teleostei): Inconsistent implications from morphology, nuclear and mitochondrial genomes. J Mol Evol 41:741–748.

Curole JP, Kocher TD. 1999. Mitogenomics: Digging deeper with complete mitochondrial genomes. Tree 14:394–398.

Gao TX, Chen SQ, Liu JX, Zhang YP. 2004. Comparative analysis of mitochondrial DNA sequences between *Verasper variegatus* and *V. moseri*. Gaojishu Tongxun 14(suppl):329–334.

Guo XH, Liu SJ, Liu Q, Liu Y. 2004. New progresses on mitochondrial DNA in fish. Acta Genet 31(9):983–1000.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163.

Lee WJ, Conroy J, Howell WH. 1995. Structure and evolution of teleost mitochondrial control regions. J Mol Evol 41(1):54–66.

Liu HZ. 2002. Structure and evolution of mitochondrial DNA control region in fish. Prog Nat Sci 12(3):266–270.

Liu HZ, Tzeng C, Teng H. 2002. Sequence variations in the mitochondrial DNA control region and their implications for the phylogeny of the Cypriniformes. Can J Zool 80:569–581.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964.

Mackie IM, Pryde SE, Gonzales-Sotelo C, Medina I, Pèrez-Martin R, Quinteiro J, Rey-Mendez M, Rehbein H. 1999. Challenges in the identification of species of canned fish. Trends Food Sci Technol 10:9–14.

Meyer A. 1993. Evolution of mitochondrial DNA in fishes. In: Hochachka PW, Mommsen TP, editors. Biochemistry and molecular biology of fishes. Amsterdam: Elsevier Science Publishers. p 1–38.

Miya M, Kawaguchi A, Nishida M. 2001. Mitogenomic exploration of higher teleostean phylogenies: A case study for moderatescale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. Mol Biol Evol 18:1993–2009.

Nelson J. 1994. Fishes of the world. 3rd ed., New York: John Wiley & Sons. p 624.

Randi E, Lucchini V. 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus Alectoris. J Mol Evol 47:449–462.

Russo CAM, Takezaki N, Nei M. 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol Biol Evol 13(3):525–536.

Saitoh K, Hayashizaki K, Yokoyama Y, Asahida T, Toyohara H, Yamashita Y. 2000. Complete nucleotide sequence of Japanese flounder (*Paralichthys olivaceus*) mitochondrial genome: Structural properties and cue for resolving teleostean relationships. J Hered 91(4):271–278.

Sbisa E, Tanzariello F, Reyes A, Pesole G. 1997. Mammalian mitochondrial D-loop region structural analysis identification of new conserved sequences and their functional and evolutionary implications. Gene 205(1–2):125–140.

Southern SO, Southern PJ, Dizon AE. 1988–1989. Molecular characterization of a cloned dolphin mitochondrial genome. J Mol Evol 28(1–2):32–42.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882.

Wang YK. 1958. Fish taxonomy. Shanghai: (in Chinese) Shanghai Sciences and Technology Press. p 532–540.

Yu J, Kong X, Yu Z, Zhou L. 2005. Comparative analysis of mitochondrial control region sequence from three flatfish species (Pleuronectidae). J Ocean Univ China 4(1):80–84.

Yue GH, Liew WC, Orban L. 2006. The complete mitochondrial genome of a basal teleost, the Asian arowana (*Scleropages formosus*, Osteoglossidae). BMC Genomics 7:242, http://www.biomedcentral.com/1471-2164/7/242.

Zardoya R, Meyer A. 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationgships among vertebrates. Mol Biol Evol 13(7):933–942.

Zeng QL, Liu HZ. 2001. Study on mitochondrial DNA control region of *Ictiobus cyprinllus*. J Hubei Univ (Nature Sci Ed. In Chinese) 23(3):261–264.

Zhang Y, Zhang E, He SP. 2003. Study on the structure of the control region of the bagridae in China and its phylogenetic significance. Acta Hydrobiol Sin 27(5):463–467.