

RESEARCH

Open Access

# Complete mitochondrial genome sequences of three rhombosoleid fishes and comparative analyses with other flatfishes (Pleuronectiformes)

Shu-Ying Wang, Wei Shi, Xian-Guang Miao and Xiao-Yu Kong\*

## Abstract

**Background:** *Peltorhamphus novaezeelandiae*, *Colistium nudipinnis*, and *Pelotretis flavilatus* belong to the family Rhombosoleidae of Pleuronectiformes. Their high phenotypic similarity has provoked great differences in the number and nomenclature of the taxa that depend primarily on morphological features. These facts have made it necessary to develop molecular markers for taxonomy and phylogenetic studies. In this study, the complete mitogenomes (mtDNA) of the three rhombosoleid fishes were determined for the comparative studies and potential development of molecular markers in the future.

**Results:** The lengths of the complete mitogenome of the three flatfishes are 16,889, 16,588, and 16,937 bp in the order mentioned above. The difference of lengths mainly results from the presence of tandem repeats at the 3'-end with variations of motif length and copy number in the control regions (CR). The gene content and arrangement is identical to that of the typical teleostean mtDNA. Two large intergenic spacers of 28 and 18 bp were found in *P. flavilatus* mtDNA. The genes are highly conserved except for the sizes of *ND1* (which is 28 bp shorter than the two others), *ND5* (13 bp longer), and *tRNA<sup>Glu</sup>* (5 bp longer) in *P. flavilatus* mtDNA. The symbolic structures of the CRs are observed as in other fishes, including ETAS, CSB-F, E, D, C, B, A, G-BOX, pyrimidine tract, and CSB2, 3.

**Conclusions:** Comparative genomic analysis within rhombosoleids revealed that the mitogenomic feature of *P. flavilatus* was significantly different from that of the two others. Base composition, gene arrangement, and CR structure were carried on in the 17 mitogenomes. Apart from gene rearrangement in two tongue soles (*Cynoglossus semilaevis* and *Cynoglossus abbreviatus*), the gene order in 15 others is identical to that of the typical fish mitogenomes. Of the 16 studied mitogenomes, 15 species (except for *Zebrias zebrius*) have tandem repeats at the 3', 5', or both 3'- and 5'-ends of the CRs. Moreover, the motif length and copy number intraspecies or interspecies are also variable. These phenomena fully indicate the diversity of repeats in flatfish mtDNA and would provide useful data for studies on the structure of mitogenomes in fishes.

**Keywords:** *Peltorhamphus novaezeelandiae*; *Colistium nudipinnis*; *Pelotretis flavilatus*; mtDNA

## Background

Flatfishes share a common asymmetrical body and bottom-dwelling mode of life. Their high phenotypic similarity has provoked great confusion in the number and nomenclature of taxa depending on the relevance assigned to morphological features (Chapleau 1993; Cooper and Chapleau 1998; Hoshino 2001). These facts have made it necessary to

develop molecular markers to figure out controversial aspects of flatfish systematics.

Generally, mitochondrial DNA (mtDNA) in vertebrata consists of 13 protein-coding genes, 2 rRNA genes, 22 tRNA genes, 1 origin of replication on the light strand (*O<sub>L</sub>*), and a single large control region (CR). Most genes are encoded by the heavy (H-) strand while only the *ND6* gene and eight tRNA genes are encoded by the light (L-) strand (Boore 1999). Due to its simple structure, the lack of recombination, multi-copy status in a cell, maternal inheritance, and high evolutionary rate, the mtDNA

\* Correspondence: xykong@scsio.ac.cn

Key Laboratory of Tropical Marine Bio-resources and Ecology, Marine Biodiversity Collection of South China Sea, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

has been extensively used for population genetic study and phylogenetic analysis (Miya et al. 2003; Inoue et al. 2010; Shi et al. 2011).

To date, several genes of mtDNA have been used as molecular markers in the establishment of phylogenetic relationships among flatfishes, such as rRNA genes (Azevedo et al. 2008), the cytochrome b gene (Borsa and Quignard 2001), the control region (Tinti et al. 1999), and their combinations (Infante et al. 2004). Nevertheless, it has been shown that the use of limited sequence data and markers with different evolving rates may cause errors in inferences of the evolutionary relationships among taxa. In this sense, complete mitochondrial genomes have demonstrated their ability in resolving persistent controversies over higher level relationships of teleost (Miya et al. 2003; Kawahara et al. 2008; Inoue et al. 2010; Shi et al. 2011). Currently, the complete mtDNA sequences of more than 1,000 fish species have been determined (as of 21 April 2013, <http://www.ncbi.nlm.nih.gov/>), including 14 species from 6 families in Pleuronectiformes.

*Peltorhamphus novaezeelandiae* (common sole), *Colistium nudipinnis* (turbot), and *Pelotretis flavilatus* (lemon sole) belong to the family Rhombosoleidae of Pleuronectiformes (Nelson 2006). These fishes are primarily in a South Pacific group, occurring mostly around Australia and New Zealand. Up to now, there has been controversy over the taxonomic status of the rhombosoleids. Regan (1910) separated Pleuronectidae into three subfamilies: Pleuronectinae, Samarinae, and Rhombosoleinae, and Hubbs (1945) also admitted this opinion. Then, Chabanaud (1946) recommended a familial ranking for Rhombosoleinae based on their three highly important morphological characteristics. However, this recommendation was not widely accepted. Only some researchers agreed with the classification (Chapleau and Keast 1988; Cooper and Chapleau 1998; Guibord 2003; Nelson 2006), while the others kept using a subfamilial ranking (Sakamoto 1984; Li and Wang 1995; Schwarzhans 1999; Evseenko 2004).

Previously, there were no reports on the complete mitogenome of rhombosoleid fishes, and only a few mitogenomic fragments are available. In the present study, the complete mitochondrial sequences of the three rhombosoleid fishes were determined for the first time. The genomic features of these mitogenomes were analyzed and compared with other flatfish mtDNAs. The results of this study could provide useful data for the studies of mitogenome structures in fishes and the development of molecular markers to explore the classification issues within Rhombosoleidae and Pleuronectiformes in the future.

## Methods

### Sampling, DNA extraction, PCR, and sequencing

Fish samples were obtained from the Sydney fish market, Australia, and preserved in 75% ethanol. The ethical

approval is not required because the specimen used in the present study was common marine captured economic fishes, and all fish specimens had died when we obtained them and they were sourced from commercial fisheries. Those species were not involved in the endangered list of IUCN. Total genomic DNAs were extracted from muscular tissues with DNA extraction kit (TIANGEN Biotech, Beijing, China) by following the manufacturer's protocol.

The primers used to amplify the contiguous (Table 1), overlapping segments of complete mitochondrial genomes of the three rhombosoleid fishes were designed by aligning and comparing with previously reported mitogenomic sequences of flatfishes or other references (Palumbi et al. 1991; Kong et al. 2009; Shi et al. 2011). The PCR was performed in a 25- $\mu$ l reaction volume containing 0.2 mM dNTP, 0.5  $\mu$ M of each primer, 1.0 U Taq polymerase (Takara, Dalian, China), 2.5  $\mu$ l of 10 LA PCR Buffer II (Mg<sup>2+</sup> Plus), and approximately 50 ng DNA template. The PCR cycling included an initial denaturation at 95°C for 3 min, 35 cycles of a denaturation at 95°C for 30 s, an annealing step at 48°C for 40 s, elongation at 68°C to 72°C for 1 to 4 min, and a final extension at 72°C for 10 min. The PCR products were detected in 1.0% agarose gels, purified with the Takara Agarose Gel DNA Purification Kit (Takara Bio Inc., Beijing, China) and used directly as templates for the cycle sequencing reactions in both directions (with ABI 3730 DNA sequencer, Life Technologies Biotechnology Corporation, Shanghai, China). Fragments that could not be directly sequenced were inserted into the pMD20-T vector (Takara), transformed into *E. coli* DH5 $\alpha$  for cloning and then sequenced. The new primers were designed for walking sequencing. An overlap of more than 30 bp between the two adjacent sequences was used to ensure the correct assembly and integrity of the complete sequences.

The sequenced fragments were assembled into complete mitochondrial genomes using CodonCode Aligner (vers. 3, CodonCode Corporation, Dedham, MA, USA) and BioEdit (Hall 1999). Annotation and boundary determination of protein-coding genes and rRNA genes were performed using NCBI-BLAST (<http://blast.ncbi.nlm.nih.gov>). Alignments with previously published mitogenomic sequences of flatfishes and other closely related bony fishes were carried out to ensure accuracy (Table 2). The transfer RNA genes and their secondary cloverleaf structures were identified using tRNAscan-SE 1.21 (Lowe and Eddy 1997), with cutoff values set to 1 when necessary. The secondary structures of *tRNA*<sup>Ser</sup>(AGC) and *tRNA*<sup>Cys</sup> were further constructed by RNA structure (Reuter and Mathews 2010) and confirmed by examination of their anticodons.

To compare the mitogenome characteristics among Pleuronectiformes, the complete mitogenome sequences of 14 other flatfishes from 6 families were retrieved from GenBank (Table 2).

**Table 1 PCR primers for amplification of the complete mitogenome of three rhombosoleid species**

Forward	Sequence (5'-3')	Reverse	Sequence (5'-3')
Pleur-Z15	ATTAAAGCATAACHCTGAAGATGTTAAGAT	Pleur-F6746	GCGGTGGATTGTAGACCCATARACAGAGGT
Pleur-Z2625	GTTTACGACCTCGATGTTGGATCAGGACAT	Pleur-F13413	TAGCTGCTACTCGGATTTGCACCAAGAGT
Pleur-Z10818	TTYGAAGCAGCCGCMTGATACTGACAYTT	H15149	AAACTGCAGCCCCCTCAGAATGATATTTGTCTCTCA
L14734	CGAAGCTTGATATGAAAAACCATCGTTG	Pleur-F17147	TAGTTTARTGCGAGAATCCTAGCTTTGGG
Pleur-Z17054	GYCGGTGGTTARAATCCTCCCTACTGCT	Pleur-F11089	TTTAACCAAGACCRGGTGATTGGAAGTC
Pleur-Z13347	AAGGATAACAGCTCATCCGTTGGTCTTAGG	Pleur-F2753	TAGATAGAAACTGACCTGGATTACTCCGGT
Z-Ser	CTCGCAGCAATGAACACT	16SBR	CCGGTCTGAACCTCAGATCACGT
Z-6188	GGTGAAATCCCTTAGTCCC	F12-F-ATP6	ATGTAAAGGCAGCGGTAG
R-COI-8010	CCMCGACGCTACTCTGACTA	F12-F-12S	TGTCTATCACTGCTGGGTT
F18-Z-Ile	CTTGCCCTGGTTGTATGA	F-49	GGCCCATCTTAACATCTTC
F18-Z-Arg	CCCCAAATAAACCTGAC	F18-F-COII	CTATCCGAGCCTGAACAA
F18-Z-ND4	GCTTTGCCTACTGTCAT	F18-F-ND4	TCCCACATCCGTCGTCAT
F20-Z-49	CAGCCCTCACAAGACACT	F-14170	ATTCCTCCTCTTTGTGGG
Z-6487	AGCAGACACTCTAATTAAGC	F18-F-Cytb	CGTCCCTCCAGTTGCTCT
F12-Z-Gln	GAGATCAAAACTCTTAGTGC	F-5196	CTAAATGGTTGGGGTATGG
F20-Z6883	TCGGCTCACTTATTCCC	F20-F7749	ACGAGTGGAGGACATCTT
F20-Z-13347-2	GCCCTCCTCGTAACTTGA	F20-F-17054-2	CCCTCACCTCAATAAGA

## Results and discussion

### Genome organization

The lengths of the *P. novaezeelandiae*, *C. nudipinnis*, and *P. flavilatus* mitogenomes are 16,889, 16,588, and 16,937 bp, respectively (GenBank accession no. JQ639065, JQ639063, KC554065; note, the order of the following data is the same as these). Their gene arrangements are identical to those of typical teleost species (Saitoh et al. 2000; Yue et al. 2006). All the three mitogenomes contain 13 protein-coding genes, two rRNA genes, 22 tRNA genes, one *O<sub>L</sub>*, and one CR. All genes are encoded by the H-strand except for the *ND6* and eight tRNA genes, which were encoded on the L strand (Table 3). Base compositions of the three mitogenomes are 26.6% to 28.2%, 27.1% to 28.1%, 16.7% to 17.6%, and 27.7% to 28.1% for A, C, G, and T, respectively, with a bias on AT content (54.3% to 56.1%). A total of 34-, 28-, and 66-bp intergenic spacers are found in the three genomes. The majority of spacer lengths range from 1 to 7 bp except for the two larger spacers in *P. flavilatus*. One is 28 bp between *tRNA<sup>Leu</sup>(UUA)* and *ND1*, and the other is 18 bp between *tRNA<sup>Asp</sup>* and *COII*, which is a polymeric C fragment. A total of 37-bp, 37-bp, and 39-bp overlaps, respectively, were observed. The four notable over-lapping positions (*ATP8* and *ATP6*, *ATP6* and *COIII*, *ND4L* and *ND4*, *ND5* and *ND6*) were also observed as reported in other vertebrate species (Kong et al. 2009) (Table 3).

### Protein-coding genes

The sizes of the 13 protein-coding genes are 11,441 bp in *P. novaezeelandiae* and *C. nudipinnis* but are 11,424 bp in

*P. flavilatus*. Comparison of the length of each gene reveals that the genes are highly conserved in size except for the *ND1* gene (28 bp shorter) and the *ND5* gene (13 bp longer) in *P. flavilatus* mtDNA. The start codons are identical in the three species. Eleven genes use the ATG, whereas *COI* starts with GTG, and *ND3* with ATA, which has rarely been found in fish mitogenomes to date (other examples include *Albula glossodonta*, *Monopterus albus*, *Petrosirtes breviceps*, *Solea senegalensis*, and *Cynoglossus semilaevis*) (Miya et al. 2003; Miya et al. 2001; Inoue et al. 2004; Kong et al. 2009). Ten of the 13 genes use the same stop codons. The *ND5* gene ends with TAA in *P. novaezeelandiae* and *P. flavilatus* and TAG in *C. nudipinnis*; the *COII* and *cytb* end with TAA in *P. novaezeelandiae*, but with T in *C. nudipinnis* and *P. flavilatus* (Table 3).

The base compositions of the 13 protein-coding genes are T > C > A > G. The proportions of the four bases have no apparent bias at the first codon position but have significant difference at the second and third positions. The percentage of T at the second position is up to 40.9% to 41.1%, but that of G is only 13.6% to 13.8%. In particular, G at the third position is only 9.6% to 12.5%, which is in agreement with previous reports (Saitoh et al. 2000; Miya et al. 2003; Oh et al. 2007). There is a slight difference in codon usage among three rhombosoleids. The most frequently used amino acid is leucine (16.5% to 17.2%), while cysteine (0.7% to 0.8%) is the least frequently used. The level of homology of genes between the three rhombosoleid species ranges from 63% (*ND2* gene) to 85% (*COII* gene). The similarity between *C. nudipinnis* and *P. flavilatus* is generally higher

**Table 2 Information of the mitogenomes sequences of 17 flatfishes used in present study**

Species	Accession number	Genome		Protein-coding gene			Non-coding region					
		Length (bp)	AT (%)	Length (bp)	AT skew	GC skew	Control region				O <sub>L</sub> region	
							Length	Repeat site	Copy number	Size	Length	Block
<i>Cynoglossus semilaevis</i>	EU366230	16,731	60.59	11,416	0.20	−0.57	982	5′	11.9	32	44	CCGGC
<i>Cynoglossus abbreviatus</i>	NC_014881	16,417	60.35	11,412	0.23	−0.61	661	5′	4.7	18	41	CCGGC
<i>Solea senegalensis</i>	AB270760	16,659	54.63	11,430	0.17	−0.63	1,017	3′	8.1	8	46	CCGGG
<i>Zebrias zebrinus</i> <sup>a</sup>	JQ700100	16,758	54.81	11,438	0.17	−0.64	1,078	-	-	-	50	CCAGC
<i>Peltorhamphus novaezeelandiae</i>	JQ639065	16,889		11,429	0.03		1,212	3′	6.1	7	43	CTGGC
								3′	39.3	7		
<i>Colistium nudipinnis</i>	JQ639063	16,588	54.33	11,435	−0.02	−0.45	919	3′	7	17	49	CTGGC
<i>Pelotretis flavilatus</i>	KC554065	16,937	56.08	11,427	0.10	−0.51	1,218	3′	2.9	145	49	CTGGC
<i>Verasper variegatus</i>	NC_007939	17,273	54.69	11,433	0.12	−0.57	1,572	3′	8.3	61	47	CCGGC
<i>Verasper moseri</i>	NC_008461	17,588	55.06	11,435	0.13	−0.56	1,889	3′	13.3	61	46	CCGGC
<i>Platichthys stellatus</i>	NC_010966	17,103	53.21	11,438	0.12	−0.50	1,400	3′	3.2	120	48	CCGGC
<i>Kareius bicoloratus</i> <sup>b</sup>	NC_003176	15,973	53.26	11,436	0.13	−0.51	-	-	-	-	46	CCGGC
<i>Hippoglossus stenolepis</i>	NC_009710	17,841	54.29	11,443	0.17	−0.55	2,135	3′	17.4	61	48	CCGGC
<i>Hippoglossus hippoglossus</i>	NC_009709	17,546		11,443	0.16		1,841	3′	3	11	48	CCGGC
								3′	12.4	61		
<i>Reinhardtius hippoglossoides</i>	NC_009711	18,017	54.9	11,443	0.12	−0.55	2,312	3′	20.3	61	48	CCGGC
<i>Paralichthys olivaceus</i>	AB028664	17,090	53.54	11,428	0.16	−0.55	1,400	3′	5.4	74	48	CCGGT
<i>Scophthalmus maximus</i>	EU419747		55.78	11,506		−0.56	1,658	3′	61.4	12	45	CCGGT
								3′	6.4	11		
<i>Psettodes erumei</i>	NC_020032.1		53.61	11,427		−0.71	1,601	5′	8.7	56	45	CCGGC
								3′	23.3	8		

<sup>a</sup>There was no repeat units in the control region of *Zebrias zebrinus*; <sup>b</sup>the control region of *Kareius bicoloratus* was incomplete. 3' or 5' indicates that the repeat region is at the 3'- or 5'-end of the control region, respectively.

than that between *P. novaezeelandiae* and each of these two fishes (Figure 1).

### Ribosomal and transfer RNA genes

Two rRNA genes are typically located between *tRNA<sup>Phe</sup>* and *tRNA<sup>Leu</sup>* (*UUA*) and separated by *tRNA<sup>Val</sup>* (Table 3). The lengths of *12S rRNA* genes are similar and those of *16S rRNA* have approximately 15 bp differences among three species (Table 3). The level of homology of rRNA genes is very similar in *12S rRNA* but slightly different in *16S rRNA* among the three rhombosoleids (Figure 1).

The 22 tRNA genes are interspersed between rRNA and protein-coding genes. Most of these tRNAs are of similar length as those in other fishes, except for *tRNA<sup>Glu</sup>*, which is 5 bp longer in *P. flavilatus* than those in the other fishes. The majority of tRNA genes could be recognized and folded into secondary structures by tRNAscan-SE, except for two genes that were identified by comparing with other flatfishes. One is the *tRNA<sup>Ser</sup>*(AGC) gene in *C. nudipinnis* and *P. flavilatus*, which is the common case in fishes, and another is the *tRNA<sup>Cys</sup>* gene, in which the dihydrouracil loop cannot be formed in any of the three

fishes mtDNA. The lengths are 7, 5, 4, and 5 bp for that of the amino acid arm, anticodon arm, and DHU and TΨC arm, respectively. Both the anticodon and TΨC loop are 7 nucleotides long, whereas DHU loop size varies from 5 to 11 nucleotides.

### Non-coding sequences

The O<sub>L</sub> is normally located between *tRNA<sup>Asn</sup>* and *tRNA<sup>Cys</sup>* in the WANCY region and is from 38 to 44 bp in size. These regions have the potential to fold into a stem-loop structure with 13 or 14 bp in the arms and 10 or 14 nucleotides in the loops. The highly conserved sequence motif 5'-GCCAG-3' is substituted by 5'-GC CGG-3' (Figure 2).

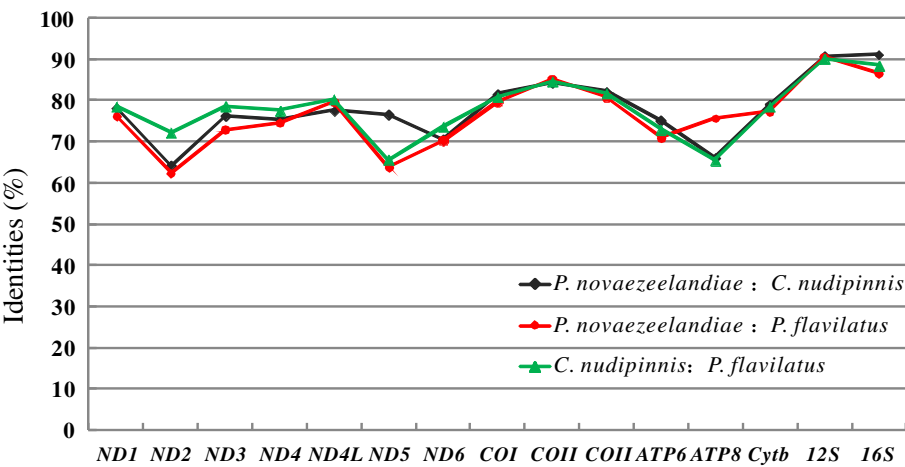
The control regions are commonly situated in the location between *tRNA<sup>Pro</sup>* and *tRNA<sup>Phe</sup>*. Their lengths are quite different and are 919, 1,212, and 1,218 bp in *C. nudipinnis*, *P. novaezeelandiae*, and *P. flavilatus*, respectively (Table 3). These differences mainly result from the presence of tandem repeats at the 3'-end, in which the motif length and copy number of tandem repeat are variable. There are two 7-bp motifs with 6 or 39 copies in *P. novaezeelandiae*, a

**Table 3 Features of the mitogenomes of three rhombosoleid fishes**

Gene	<i>Peltorhamphus novaezeelandiae</i>				<i>Colistium nudipinnis</i>				<i>Pelotretis flavilatus</i>			
	Position	Start	Stop	I-R*	Position	Start	Stop	I-R	Position	Start	Stop	I-R
<i>tRNA<sup>Phe</sup></i>	1 to 68	GAA		0	1 to 68	GAA		0	1 to 69	GAA		0
<i>12S rRNA</i>	69 to 1,010			0	69 to 1,010			0	70 to 1,016			0
<i>tRNA<sup>Val</sup></i>	1,011 to 1,081	TAC		0	1,011 to 1,081	TAC		0	1,017 to 1,087	TAC		0
<i>16S rRNA</i>	1,082 to 2,795			1	1,082 to 2,792			1	1,088 to 2,816			0
<i>tRNA<sup>Leu</sup>(UUA)</i>	2,797 to 2,870	TAA		0	2,794 to 2,867	TAA		0	2,817 to 2,890	TAA		28
<i>ND1</i>	2,871 to 3,845	ATG	TAA	5	2,868 to 3,842	ATG	TAA	5	2,919 to 3,866	ATG	TAA	6
<i>tRNA<sup>Ile</sup></i>	3,851 to 3,920	GAT		-1	3,848 to 3,917	GAT		-1	3,873 to 3,942	GAT		-1
<i>tRNA<sup>Gln</sup></i>	3,920 to 3,990	TTG		-1	3,917 to 3,987	TTG		-1	3,942 to 4,013	TTG		-1
<i>tRNA<sup>Met</sup></i>	3,990 to 4,059	CAT		0	3,987 to 4,055	CAT		1	4,013 to 4,081	CAT		1
<i>ND2</i>	4,060 to 5,109	ATG	TAA	-1	4,057 to 5,106	ATG	TAA	-1	4,083 to 5,132	ATG	TAA	-1
<i>tRNA<sup>Trp</sup></i>	5,109 to 5,179	TCA		1	5,106 to 5,176	TCA		1	5,132 to 5,202	TCA		1
<i>tRNA<sup>Ala</sup></i>	5,181 to 5,249	TGC		1	5,178 to 5,246	TGC		1	5,204 to 5,272	TGC		1
<i>tRNA<sup>Asn</sup></i>	5,251 to 5,323	GTT		0	5,248 to 5,320	GTT		0	5,274 to 5,346	GTT		0
<i>O<sub>L</sub></i>	5,324 to 5,361			-4	5,321 to 5,364			-4	5,347 to 5,390			-4
<i>tRNA<sup>Cys</sup></i>	5,358 to 5,417	GCA		-1	5,361 to 5,420	GCA		-1	5,387 to 5,446	GCA		-1
<i>tRNA<sup>Tyr</sup></i>	5,417 to 5,486	GTA		1	5,420 to 5,489	GTA		1	5,446 to 5,515	GTA		1
<i>COI</i>	5,488 to 7,038	GTG	TAA	0	5,491 to 7,041	GTG	TAA	0	5,517 to 7,067	GTG	TAA	0
<i>tRNA<sup>Ser</sup>(UCA)</i>	7,039 to 7,109	TGA		3	7,042 to 7,112	TGA		3	7,068 to 7,138	TGA		3
<i>tRNA<sup>Asp</sup></i>	7,113 to 7,181	GTC		6	7,116 to 7,184	GTC		6	7,142 to 7,210	GTC		18
<i>COII</i>	7,188 to 7,880	ATG	TAA	7	7,191 to 7,881	ATG	T-	0	7,229 to 7,919	ATG	T-	0
<i>tRNA<sup>Lys</sup></i>	7,888 to 7,961	TTT		1	7,882 to 7,955	TTT		1	7,920 to 7,993	TTT		1
<i>ATP8</i>	7,963 to 8,130	ATG	TAA	-10	7,957 to 8,124	ATG	TAA	-10	7,995 to 8,162	ATG	TAA	-10
<i>ATP6</i>	8,121 to 8,804	ATG	TAA	-1	8,115 to 8,798	ATG	TAA	-1	8,153 to 8,836	ATG	TAA	-1
<i>COIII</i>	8,804 to 9,589	ATG	TAA	-1	8,798 to 9,583	ATG	TAA	-1	8,836 to 9,621	ATG	TAA	-1
<i>tRNA<sup>Gly</sup></i>	9,589 to 9,660	TCC		-3	9,583 to 9,654	TCC		-3	9,621 to 9,691	TCC		-3
<i>ND3</i>	9,658 to 10,011	ATA	TAG	-2	9,652 to 10,005	ATA	TAG	-2	9,689 to 10,042	ATA	TAG	-2
<i>tRNA<sup>Arg</sup></i>	10,010 to 10,078	TCG		0	10,004 to 10,072	TCG		0	10,041 to 10,109	TCG		0
<i>ND4L</i>	10,079 to 10,375	ATG	TAA	-7	10,073 to 10,369	ATG	TAA	-7	10,110 to 10,406	ATG	TAA	-7
<i>ND4</i>	10,369 to 11,749	ATG	T-	0	10,363 to 11,743	ATG	T-	0	10,400 to 11,780	ATG	T-	0
<i>tRNA<sup>His</sup></i>	11,750 to 11,818	GTG		0	11,744 to 11,812	GTG		0	11,781 to 11,849	GTG		0
<i>tRNA<sup>Ser</sup>(AGC)</i>	11,819 to 11,885	GCT		4	11,813 to 11,879	GCT		4	11,850 to 11,916	GCT		4
<i>tRNA<sup>Leu</sup>(CUA)</i>	11,890 to 11,962	TAG		0	11,884 to 11,956	TAG		0	11,921 to 11,993	TAG		1
<i>ND5</i>	11,963 to 13,801	ATG	TAA	-4	11,957 to 13,795	ATG	TAG	-4	11,995 to 13,845	ATG	TAA	-4
<i>ND6</i>	13,798 to 14,319	ATG	TAG	0	13,792 to 14,313	ATG	TAG	0	13,842 to 14,363	ATG	TAG	1
<i>tRNA<sup>Glu</sup></i>	14,320 to 14,388	TTC		4	14,314 to 14,382	TTC		4	14,365 to 14,438	TTC		-2
<i>Cytb</i>	14,393 to 15,533	ATG	T-	0	14,387 to 15,527	ATG	T-	0	14,437 to 15,577	ATG	T-	0
<i>tRNA<sup>Thr</sup></i>	15,534 to 15,607	TGT		-1	15,528 to 15,599	TGT		-1	15,578 to 15,649	TGT		-1
<i>tRNA<sup>Pro</sup></i>	15,607 to 15,677	TGG		0	15,599 to 15,669	TGG		0	15,649 to 15,719	TGG		0
D-loop	15,678 to 16,889				15,670 to 16,588				15,720 to 16,937			

I-R\*, intergenic region; non-coding bases between the feature on the same line and the line below, with a negative number indicating an overlap.





**Figure 1** Sequence identities of 13 protein genes and two rRNA genes among the three rhombosoleids.

17-bp motif with seven copies in *C. nudipinnis*, and a 145-bp motif with three copies in *P. flavilatus* (Figure 3).

The AT contents of the CRs reach up to 64.4% to 66.7%, which are higher than those of the whole mtDNA sequences. The symbolic structures of the CRs are observed as in other fishes (Figure 3), including the extended termination associated sequence (ETAS, containing TAS-cTAS: TACAT-ATGTA), central conserved sequence blocks (CSB-EE,D), G-BOX (GTGGGGG), pyrimidine tract (poly-T), and conserved sequence blocks (CSB 2-3) (Nesbo et al. 1998; Manchado et al. 2007; Wang et al. 2013).

**Comparative analyses with other flatfishes**

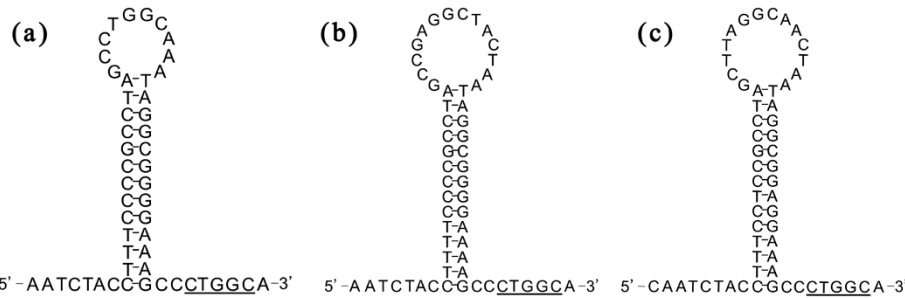
Up to now, 14 mitogenome sequences from other flatfishes had been determined (Table 2). To better understand their features of mtDNAs, a comparative analysis was carried out in several aspects.

First of all, the content of the 17 mitogenome sequences (including three from the present study) are the same, which consists of 37 genes, 1 *O<sub>L</sub>*, and 1 CR. However, the gene arrangements differ among them. Apart from the two tongue soles (*C. semilaevis* and *C. abbreviatus*), the orders of the 15 others are identical to that

of the typical fish mitogenomes. The organization of the tongue sole mitogenomes differed, in which the *tRNA<sup>Gln</sup>* gene is inverted from the light strand to the heavy strand (inversion), accompanied by shuffling of the *tRNA<sup>Ile</sup>* gene and long-range translocation of the control region downstream to a site between the *ND1* and the *tRNA<sup>Gln</sup>* genes.

The lengths of the 17 mitogenome sequences show apparent differences (from 15,973 bp of *Kareius bicoloratus* to 18,017 bp of *Reinhardtius hippoglossoides*). The reason for the short CR in *K. bicoloratus* mtDNA is due to the unfinished sequencing of CR and that for two other tongue soles (16,417 bp or 16,731 bp) is due to the rearrangement of the CR. The variations for the other fishes are primarily caused by the presence of the repeated arrays in control regions.

The proportions of three bases (A\T\C) have no obvious difference, ranging from 20% to 30%; however, that of G is remarkably lower, from 14.5% (*C. semilaevis*) to 17.7% (*Platichthys stellatus*). The AT compositions in the 15 species mitogenomes are generally approximately 50% (from 53.21% to 56.08%), but those of the two tongue soles reach up to 60.35% and 60.59% (Table 2).



**Figure 2** The stem-loop structures of *O<sub>L</sub>* in the mitogenomes of three rhombosoleids. The underlined sequences indicate the conserved sequence motif. (a) *P. novaezeelandiae*, (b) *C. nudipinnis*, and (c) *P. flavilatus*.

**Figure 3** Alignment of the control regions of *P. novaezeelandiae* (*P.nov*), *C. nudipinnis* (*C.nud*), and *P. flavilatus* (*P.fla*) mtDNA. The blocks CSB-A, B, C, D, E, F, poly-T, and CSB-2 and CSB-3 are shaded. CSB, conserved sequence block. The sequence in parentheses indicates the motif of the tandem repeat, and the arabic number indicates the copy number.

Based on the alignment of CR sequences of 14 flatfishes, the typical tripartite structure was found as those in *P. novaezeelandiae*, *C. nudipinnis*, and *P. flavilatus* mtDNA. The six blocks of CSB-F, E, D, C, B, and A were identified in the central conserved blocks domain, and the key sequences of each block are as follows: CSB-F: GTAAGAGCCTACCAACCGG, CSB-E: GGGTGAGGG ACAAAAATT -GTGGGGG, CSB-D: TATTCCTGGC ATTTGGTTCC-TACTTCAGGGCCAT, CSB-C: CTTA CATAAGTTAATG, CSB-B: CATACTGACTCGTTACCC AGCAAGCCGGGCGTTC; CSB-A: CTCCAGCGGGT AAGGGG. The G-box (GTGGGGG) is the most conservative in CSB-E. Simultaneously, a pyrimidine tract following the CSB-A was also identified (TTCTC-TTTT TT TTTTTCCTTTC). Two conserved sequence blocks of CSB-2 and CSB-3 at the 3'-end of the CR were identifiable; their sequences are CSB-2: AAAACCCCCC-TAC

CCCCCTAAA and CSB-3: CCTGAAAACCCCCCGG, respectively (see Additional file 1).

Generally, the variation of the control region is relatively greater than that of the other sequences in the mitogenome. However, there have also been some conserved structures in the CR as they are supposed to contain the functional structures, such as the origins of the H-strand (OH), the heavy strand promoter (HSP), and light strand promoter (LSP) of transcription transcripts (Shadel and Clayton 1997). So far, the sequences of some conserved blocks have not been defined. Generally, only the TAS, CSB-F, E, and D of the central conserved blocks are identified in most fishes, while the CSB-C, B, and A are greatly variable (Lee et al. 1995; Guo et al. 2003; Manchado et al. 2007; Zhang et al. 2010), and CSB-1,2,3 have been found only in some fishes (Lee et al. 2001; Liu 2002; Guo et al. 2003).

## Conclusions

We sequenced the complete mitogenomes of three rhombosoleid fishes in the Pleuronectiformes. Comparative genomic analysis within the rhombosoleids revealed that the genomic feature of *P. flavilatus* is apparently different from the other two.

The comparison of complete mitogenome sequences of 3 rhombosoleids with that of the other 14 flatfishes show some different features among them. Firstly, the genomic arrangement of the 15 mitogenomes is identical to that of a typical teleost, but the order of two tongue soles showed clear rearrangements. Secondly, the length heterogeneity is apparently large and up to 1,600 bp. The main reason for this case is due to the presence of repeat regions in the CRs. The 15 species have tandem repeats, which were distributed at all potential existing sites in the CR, including 3'-, 5'-, or both 3'- and 5'-ends of the CR. Moreover, the motif length and copy number in intraspecies or interspecies are also variable. Thirdly, six blocks of CSB-F, E, D, C, B, and A in the central conserved blocks domain and CSB-2 and CSB-3 in the conserved sequence blocks were identified. However, CSB-1 is not conserved in the flatfishes studied. These phenomena fully indicated the diversity of repeats in flatfishes and would provide useful data for further studies on the structure of mitogenomes in fishes.

Summarily, the complete mitogenomic sequences of rhombosoleids and rich molecular information were obtained in this study. It will contribute to figuring out the existing controversy, such as the taxonomic status and phylogenetic relations of rhombosoleids. Comparative genomics analysis within flatfishes conducted here may help better understand the evolution of mitogenomic structures and explore the phylogenetic relationships of the Pleuronectiforms.

## Data accession

Sequences were deposited in the NCBI [no. JQ639065, JQ639063, KC554065].

## Additional file

**Additional file 1: Alignment of the control regions of the 14 flatfishes mtDNA.**

## Competing interests

The authors declare that they have no competing interests.

## Authors contributions

SYW collected datasets, carried out experiments, and drafted the manuscript. WS and XGM carried out partial experiments. XYK directed the whole research work and revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We are very grateful to Mr. Bernard Yau for his help in collecting the rhombosoleids samples from Australia. This study is supported by the National Natural Science Foundation of China (30870283 and 31471979).

Received: 22 July 2013 Accepted: 21 November 2014

Published online: 17 December 2014

## References

- Azevedo MFC, Oliveira C, Pardo BG, Martinez P, Foresti F (2008) Phylogenetic analysis of the order Pleuronectiformes (Teleostei) based on sequences of 12S and 16S mitochondrial genes. *Genet Mol Biol* 31(1):284–292
- Boore J (1999) Animal mitochondrial genomes. *Nucleic Acids Res* 27:1767–1780
- Borsa P, Quignard JP (2001) Systematics of the Atlantic-Mediterranean soles *Pegusa impar*, *P. lascaris*, *Solea aegyptiaca*, *S. senegalensis*, and *S. solea* (Pleuronectiformes: Soleidae). *Can J Zool* 79(12):2297–2302
- Chabanaud P (1946) Sur le véritable caractère externe, par quoi les Soleiformes diffèrent des Pleuronectiformes. *Bull Mus Natl Hist Nat* 18:158–160
- Chapleau F (1993) Pleuronectiform relationships: a cladistic reassessment. *Bull Mar Sci* 52(1):516–540
- Chapleau F, Keast A (1988) A phylogenetic reassessment of the monophyletic status of the family Soleidae, with comments on the suborder Soleoidei (Pisces; Pleuronectiformes). *Can J Zool* 66(12):2797–2810
- Cooper JA, Chapleau F (1998) Monophyly and intrarelationships of the family Pleuronectidae (Pleuronectiformes), with a revised classification. *Fish Bull* 96(4):686–726
- Evseenko S (2004) Family Pleuronectidae Cuvier 1816 – righteye flounders. *Calif Acad Sci Annotated Checklists Fishes* 37:1–37
- Guibord AC (2003) Taxonomic revision of the Poecilopsettidae and phylogenetic analysis of the Rhombosoleidae and Pleuronectiformes (Acanthopterygii): Révision taxinomique des Poecilopsettidae et phylogénèse des Rhombosoleidae et des Pleuronectiformes (Acanthopterygii). *Univ Ottawa, Ottawa*
- Guo X, Liu S, Liu Y (2003) Comparative analysis of the mitochondrial DNA control region in cyprinids with different ploidy level. *Aquaculture* 224(1):25–38
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hoshino K (2001) Homologies of the caudal fin rays of Pleuronectiformes (Teleostei). *Ichthyol Res* 48(3):231–246
- Hubbs CL (1945) Phylogenetic position of the Citharidae, a family of flatfishes. *University of Michigan Press, Canton*
- Infante C, Catanese G, Manchado M (2004) Phylogenetic relationships among ten sole species (Soleidae, Pleuronectiformes) from the Gulf of Cadiz (Spain) based on mitochondrial DNA sequences. *Mar Biotechnol* 6(6):612–624
- Inoue JG, Miya M, Tsukamoto K, Nishida M (2004) Mitogenomic evidence for the monophyly of elopomorph fishes (Teleostei) and the evolutionary origin of the leptocephalus larva. *Mol Phylogenet Evol* 32(1):274–286
- Inoue JG, Miya M, Miller MJ, Sado T, Hanel R, Hatooka K, Aoyama J, Minegishi Y, Nishida M, Tsukamoto K (2010) Deep-ocean origin of the freshwater eels. *Biol Lett* 6(3):363–366
- Kawahara R, Miya M, Mabuchi K, Lavoue S, Inoue JG, Satoh TP, Kawaguchi A, Nishida M (2008) Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): a new perspective based on whole mitogenome sequences from 75 higher teleosts. *Mol Phylogenet Evol* 46(1):224–236
- Kong X, Dong X, Zhang Y, Shi W, Wang Z, Yu Z (2009) A novel rearrangement in the mitochondrial genome of tongue sole, *Cynoglossus semilaevis*: control region translocation and a tRNA gene inversion. *Genome* 52(12):975–984
- Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *J Mol Evol* 41(1):54–66



- Lee J-S, Miya M, Lee Y-S, Kim CG, Park E-H, Aoki Y, Nishida M (2001) The complete DNA sequence of the mitochondrial genome of the self-fertilizing fish *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae) and the first description of duplication of a control region in fish. *Gene* 280(1):1-7
- Li SZ, Wang HM (1995) *Fauna Sinica (Osteichthyes): Pleuronectiformes*. Science Press, Beijing
- Liu H (2002) Structure and evolution of fish mitochondrial control region, bitterlings as example. *Prog Nat Sci* 12:266-270
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25(5):955-964
- Manchado M, Catanese G, Ponce M, Funes V, Infante C (2007) The complete mitochondrial genome of the Senegal sole, *Solea senegalensis* Kaup. Comparative analysis of tandem repeats in the control region among soles. *Mitochondrial DNA* 18(3):169-175
- Miya M, Kawaguchi A, Nishida M (2001) Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. *Mol Biol Evol* 18(11):1993-2009
- Miya M, Takeshima H, Endo H, Ishiguro NB, Inoue JG, Mukai T, Satoh TP, Yamaguchi M, Kawaguchi A, Mabuchi K, Shirai SM, Nishida M (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol Phylogenet Evol* 26(1):121-138
- Nelson JS (2006) *Fishes of the World* (Fourth Edition). John Wiley, Inc, New York
- Nesbo 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(4):1907-1919
- Oh DJ, Kim JY, Lee JA, Yoon WJ, Park SY, Jung YH (2007) Complete mitochondrial genome of the rock bream *Oplegnathus fasciatus* (Perciformes, Oplegnathidae) with phylogenetic considerations. *Gene* 392(1-2):174-180
- Palumbi S, Martin A, Romano S, McMillan W, Stice L, Grabowski G (1991) The Simple Fool's Guide to PCR, Version 2.0, privately published document compiled by S. Palumbi. Dept. Zool Univ Hawaii Honolulu HI 96822
- Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* 41(3):353-358
- Regan CT (1910) The origin and evolution of the teleostean fishes of the order Heterosomata. *Ann Mag Nat Hist* 6:484-496
- Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11:129
- 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
- Sakamoto K (1984) Interrelationships of the family Pleuronectidae (Pisces: Pleuronectiformes). *Mem Fac Fish Hokkaido Univ* 31:95-215
- Schwarzhan W (1999) A comparative morphological treatise of recent and fossil otoliths of the order Pleuronectiformes, vol 2. F. Pfeil, Verlag
- Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem* 66(1):409-435
- Shi W, Kong XY, Wang ZM, Jiang JX (2011) Utility of tRNA genes from the complete mitochondrial genome of *Psetta maxima* for implying a possible sister-group relationship to the Pleuronectiformes. *Zool Stud* 50(5):665-681
- Tinti F, Colombari A, Vallisneri M, Piccinetti C, Stagni A (1999) Comparative analysis of a mitochondrial DNA control region amplified from three Adriatic flatfish species and molecular phylogenesis of Pleuronectiformes. *Mar Biotechnol* 1:20-24
- Wang Z-M, Shi W, Jiang J-X, Wang S-Y, Miao X-G, Huang L-M, Kong X-Y (2013) The complete mitochondrial genome of a striped sole *Zebrias zebrinus* (Pleuronectiformes: Soleidae). *Mitochondrial DNA* 24(6):633-635
- Yue G, Liew W, Orban L (2006) The complete mitochondrial genome of a basal teleost, the Asian arowana (*Sclerophages formosus*, Osteoglossidae). *BMC Genomics* 7:242
- Zhang Y, Kong X, Wang Z (2010) Characterization of mitochondrial control region sequence of *Psettodes erumei* and phylogenetic analysis of Pleuronectiformes. *J Trop Oceanograph* 29(6):71-78

doi:10.1186/s40555-014-0080-9

**Cite this article as:** Wang et al.: Complete mitochondrial genome sequences of three rhombosoleid fishes and comparative analyses with other flatfishes (Pleuronectiformes). *Zoological Studies* 2014 **53**:80.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)