Pakistan J. Zool., vol. 51(5), pp 1869-1880, 2019. **DOI**: http://dx.doi.org/10.17582/journal.pjz/2019.51.5.1869.1880

Sequencing Analysis of the Mitochondrial Genome of Japanese Sand Lance *Ammodytes personatus* Based on Next-Generation Sequencing Technology

Jingchen Chen, Zhaochao Deng and Zhiqiang Han*

Fishery College, Zhejiang Ocean University, Zhoushan, Zhejiang 316022, China

ABSTRACT

The next-generation sequencing technology is obviously faster than Sanger sequencing technology, and the cost is basically the same. Based on this, we used the next-generation sequencing technology to measure the whole mitochondrial genome sequence of *Ammodytes personatus*, and compared with Sanger sequencing results. Three samples belonging to two lineages of *A. personatus* mitochondrial DNA (mtDNA) were used in the study. Based on the complete mtDNA sequence, the genetic distance between the two lineages is 0.064. Results show that the total sequence lengths of the three samples of *A. personatus* mitochondrial genome are 16,537 bp (AP1), 16,536 bp (AP2) and 16,537 bp (AP3), where the base composition has biases of high A+T and low G+C contents, which are typical to the structural composition of vertebrates. Moreover, the two promoters (ATG, GTG) and four terminators (TAG, AGG, TAA, T or TA) in the 13 protein-coding genes are found. Except for tRNA-Ser$^{(AGN)}$, the second-structure of other tRNAs is the typical clover structure. The lengths of 12S rRNA and 16S rRNA are 945 and 1,698 bp (AP1, AP2) and 1,696 (AP3). A control region containing key sequence tags have three different domains, namely, terminating sequences (TAS1, TAS2), central conservatives (CSB-F, CSB-E and CSB-D) and conservative sequences (CSB1, CSB2 and CSB3). A conserved sequence region that controls the initiation of light-chain replication is identified in the non-coding region outside the control area. Based on the complete mitochondrial genome and COI gene, we can identify the phylogenetic relationship of *A. personatus* to other Perciformes species. Our results clearly demonstrate that the high-throughput sequencing method is the methodology of choice for generating complete mtDNA genome sequences.

Article Information
Received 26 September 2018
Revised 12 November 2018
Accepted 05 December 2018
Available online 12 July 2019

Authors' Contribution
ZH conceived and designed the work. ZD collected the specimens. JC performed the experiments, analysed the data and wrote the manuscript.

Key words
***Ammodytes personatus*, NGS technology, Mitochondrial DNA, Control region, Phylogenetic relationship.**

INTRODUCTION

The Japanese sand lance (*Ammodytes personatus*) is a common and commercially important species in the Northwestern Pacific. Moreover, populations of *A. personatus* are distributed along both sides of Japan's coast and the Yellow Sea. *A. personatus* is a cold water species that is abundant in shallow and near-shore areas in depths of up to 100 metres. It rises to the sea surface by night and buries itself in sand when the water temperature increases to 17 to 20°C during summer in southern and central Japan populations (Han *et al*., 2012).

Research on *A. personatus* has mainly focused on its morphology, physiology (Yamashita *et al*., 1986), ethology (Yamada *et al*., 1998) and population genetics (Okamoto *et al*., 1988). Some studies have used the Sanger sequencing technology to measure the complete sequence of *A. personatus* mitochondrial genome (Li *et al*., 2015; Gao *et al*., 2013; Satoh *et al*., 2016). High-throughput sequencing technology, also known as deep sequencing or next-generation sequencing technology, is capable of large-scale parallel sequencing. Compared with the sanger sequencing technology, next-generation sequencing technology is more time saving, which do not need to design specific primers. Sanger sequencing takes at least a week longer than the next-generation sequencing.

Vertebrate mitochondrial DNA (mtDNA) is self-replicating and has high mutation rate and evolution speed (Ingman *et al*., 2006). Substantial information on genetic evolution can be obtained by analysing the genetic structure and sequence characteristics of mtDNA. MtDNA is widely used in the fields of animal phylogeny, species classification (Pan *et al*., 2006; Liu *et al*., 2018), population genetic evolution and germplasm identification (Moritz *et al*., 1994). Two highly divergent mitochondrial lineages (lineages A and B) were detected in *A. personatus* based on the mtDNA control region (Han *et al*., 2012). Morphological population studies of *A. personatus*

* Corresponding author: d6339124@163.com
0030-9923/2019/0005-1869 $ 9.00/0

revealed northern and southern groups. Only one lineage (Lineage A) was detected in southern group and lineages A and B coexisted in northern group (Han *et al*., 2012). However, the genetic divergences for mtDNA protein coding genes of these two lineages were unknown. In this study we collected samples from Aomori, Japan and Qingdao, China, performed complete sequencing and identified the structural characteristics of *A. personatus* mitochondrial genome using next-generation sequencing technology. Moreover, this study further analysed the phylogenetic relationship of Ammodytidae fishes.

MATERIALS AND METHODS

Experimental materials and next-generation sequencing

One randomly selected *A. personatus* sample from Aomori, Japan and two samples from Qingdao, China were collected. The muscles were washed with deionised water and stored in ethanol at −20°C. DNA was extracted and sent to Biomarker Technologies Co., Ltd., Beijing. Illumina HiSeq accelerator 2500 platform was used to carry out the next-generation sequencing of two 150 bp length sequences at the sequencing depth of 50×.

Table I.- The complete mitochondrial genomes used for phylogenetic analyses in this study.

Species	Length (bp)	GenBank accession
Sillago sinica	16,572	NC_030373
Pholis fangi	16,523	NC_029842
Rhinogobius giurinus	16,520	NC_022692
Ammodytes personatus	16,537	AP006023
Ammodytes personatus	16,527	KF672362
Ammodytes personatus	16,537	JQ085861
Larimichthys polyactis	16,470	NC_013754
Larimichthys crocea	16,466	EU339149
Acanthogobius hasta	16,663	AY486321
Omobranchus elegans	16,517	KT284893

Data analysis

The mitochondrial genome sequences of various Perciformes species were used as the initial reference sequence (Table I) and the mitochondrial genome was annotated using GENEIOUS R11. Based on the comparison of the mitochondrial genome sequences of other Ammodytidae species, we performed the mtDNA protein-coding gene (PCGs) and D-loop annotation. MEGA6.0 was used to analyse the sequence base composition characteristics, codon usage and base offset of the mitochondrial genome in Ammodytidae species. tRNA analysis was performed using the online software tRNAscanSE 2.0 (http://lowelab.ucsc.edu/tRNAscan-SE/ index.html). Neighbour joining phylogenetic trees of the mitochondrial genome and the cytochrome coxidase 1 (COX1) gene of seven species of Perciformes was constructed using MEGA 6.0, wherein *Larimichthys crocea* and *Larimichthys polyactis* were considered as outgroups.

RESULTS

Mitochondrial genome composition

In this study, next-generation sequencing technique was used to construct a complete *A. personatus* mitochondria genome sequence with lengths of 16,536 bp (AP1), 16,537 bp (AP2) and 16,527 bp (AP3) (GenBank accession number MK163330-MK163332). A total of 13 PCGs, 22 tRNA genes, 2 rRNA genes and 1 control region (CR) were found (Fig. 1) , which accounted for 69.16% (AP1), 69.15% (AP2), 69.20% (AP3); 15.98%; 9.40% (AP1), 9.42% (AP2), 9.41% (AP3); and 5.26% of the whole mitochondrial length, respectively. The size, locus and sequence of mitochondrial genes in *A. personatus* were consistent with those in most vertebrates. In addition to the large difference in the sequence length of the CR and the relatively low homology, the length differences of the mitochondrial genome sequences of the protein coding gene, rRNA gene, tRNA gene and the mitochondrial genome of Perciformes fishes analysed in this study were small, whereas the homology was high. Specific site information is shown in Table II (AP3 shown in Table III).

Nucleotide composition

The A+T content (51.3%–58%) of Perciformes mitochondrial genome is higher than that of G+C (42%–48.7%), and its nucleotide composition has obvious A T bias (Supplementary Table I). In this study, the G+C content of *A. personatus* is on the average level of the Perciformes fish in the list, accounting for 47.3%–47.6% of the whole mitochondrial genome, 48.0%–48.5% of the protein encoded gene, 45.0% of the tRNA gene, 47.4%–47.6% of the rRNA gene and 39.5%–40.3% of CR. The base content of the whole mitochondrial genome of *A. personatus* ranging from high to low is as follows: C > T > A > G. The AT skewness base compositions are −0.017 (AP3) and −0.019 (AP1, AP2), and the GC skewness base compositions are −0.239 (AP3) and −0.243 (AP1, AP2). The A+T skew composition (-0.032 to 0.049) in the mitochondrial genome of other Perciformes species collected in this study was also higher than that of the G+C skew (-0.318 to -0.189), which indicated that more A and C than T and G were found in the mitochondrial genome

of Perciformes fish. The base composition of the genes in the Perciformes species showed that T and A content in the control and A and C content in the tRNA and rRNA genes were relatively high.

Protein-encoding genes

In this study, 13 protein-coding genes of *A. personatus* showed two initiation codons (GTG, ATG), similar to most teleost fish. COX1 uses GTG as the initiation codon, whereas the other 12 genes use ATG as the initiation codon. Relative to the stability of the initiation codon, the terminating codon changes more in the intraspecific level than the interspecific level. Four kinds of terminated codon of *A. personatus* are found in this study (AP1 shown in Table II; AP3 shown in Table III), in which TAA is the terminating codon of four genes, such as ND4L, COX2, ND5 and ATP8; TAG is the terminating codon of ND1 and ND6; AGG is the terminating codon of COX1; and ND2, ND3, COX3, ATP6, ND4 and CYTB genes use incomplete T and TA as terminating codons.

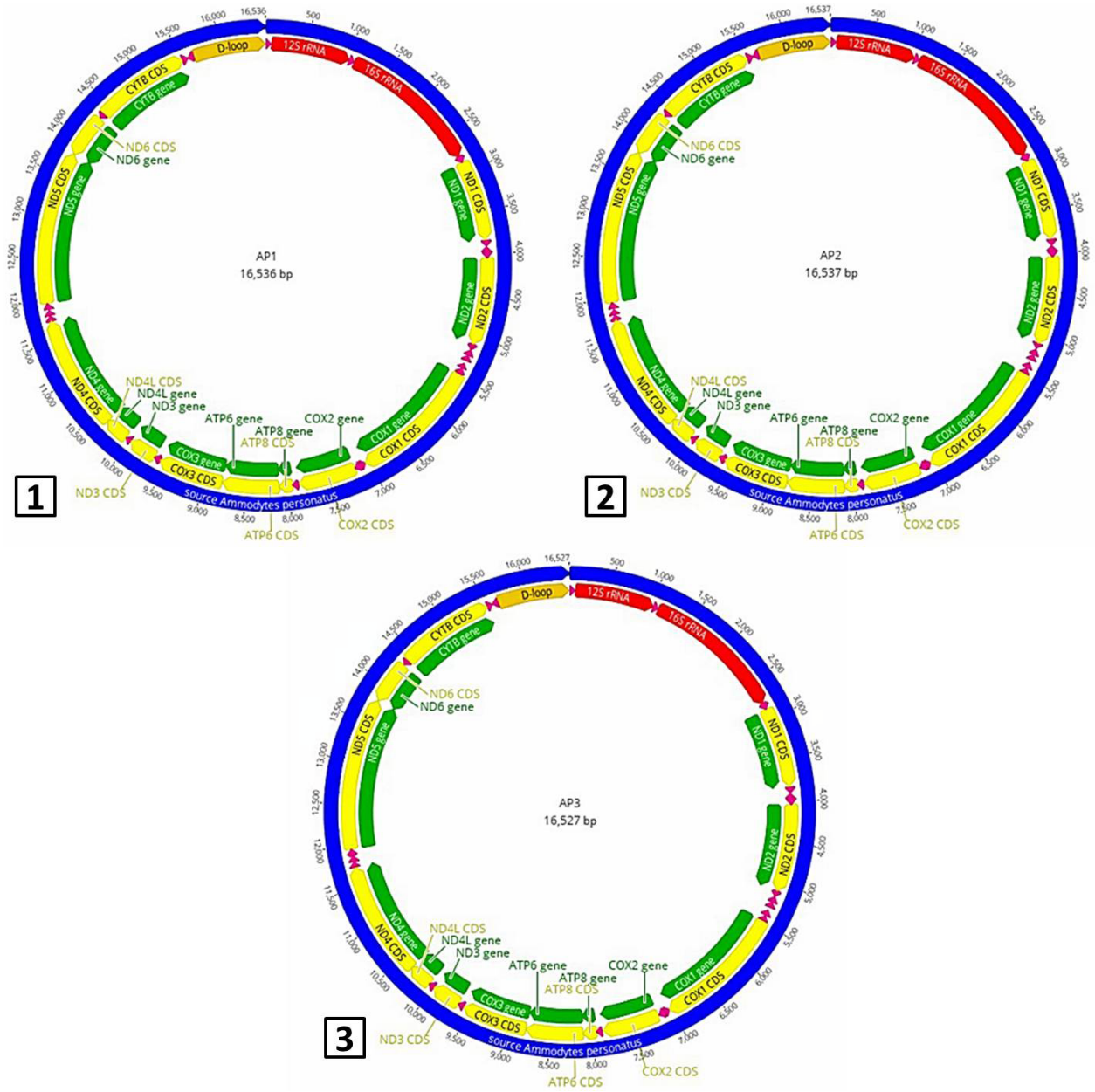

Fig. 1. Schematic map of the mitochondrial genome of *A. personatus* in this study.

Table II.- Results of the complete mitochondrial genome annotation for *Ammodytes personatus* (AP1).

Gene	Size (bp)	Position		IGS (bp)	Anticodon	Codons		Coding strand
		Start	Stop			Start	Stop	
tRNA-Phe	68	1	68	0	GAA			H
12S rRNA	945	69	1,013	0				H
tRNA-Val	74	1,014	1,087	0	TAC			H
16S rRNA	1,698	1088	2,785	0				H
tRNA-Leu	74	2,786	2,859	0	TAA			H
ND1	975	2,860	3,834	0		ATG	TAG	H
tRNA-Ile	71	3,839	3,909	4	GAT			H
tRNA-Gln	71	3,909	3,979	−1	TTG			L
tRNA-Met	69	3,979	4,047	−1	CAT			H
ND2	1,046	4,048	5,093	0		ATG	TA	H
tRNA-Trp	71	5,094	5,164	0	TCA			H
tRNA-Ala	69	5,166	5,234	1	TGG			L
tRNA-Asn	73	5,236	5,308	1	GTT			L
tRNA-Cys	67	5,343	5,409	34	GCA			L
tRNA-Tyr	71	5,410	5,480	0	GTA			L
COX1	1560	5,482	7,041	1		GTG	AGG	H
tRNA-Ser	71	7,033	7,103	−9	TGA			L
tRNA-Asp	73	7,107	7,179	3	GTC			H
COX2	690	7,188	7,877	8		ATG	TAA	H
tRNA-Lys	74	7,879	7,952	1	TTT			H
ATP8	168	7,954	8,121	1		ATG	TAA	H
ATP6	683	8,112	8,794	−10		ATG	TA	H
COX3	785	8,795	9,579	0		ATG		H
tRNA-Gly	71	9,580	9,650	0	TCC			H
ND3	349	9,651	9,999	0		ATG	T(aa)	H
tRNA-Arg	69	10,000	10,068	0	TCG			H
ND4L	297	10,069	10,365	0		ATG	TAA	H
ND4	1381	10,359	11,739	−7		ATG	T(aa)	H
tRNA-His	69	11,740	11,808	0	GTG			H
tRNA-Ser	68	11,809	11,876	0	GCT			H
tRNA-Leu	73	11,882	11,954	5	TAG			H
ND5	1839	11,955	13,793	0		ATG	TAA	H
ND6	522	13,790	14,311	−4		ATG	TAG	L
tRNA-Glu	69	14,312	14,380	0	TTC			L
CYTB	1141	14,386	15,526	5		ATG	T(aa)	H
tRNA-Thr	72	15,527	15,598	0	TGT			H
tRNA-Pro	70	15,598	15,667	−1	TGG			L
D-loop	870	15,668	16,537	0				-

Table III.- Results of the complete mitochondrial genome annotation for *Ammodytes personatus* (AP3).

Gene	Size (bp)	Position		IGS (bp)	Anticodon	Codons		Coding strand
		Start	Stop			Start	Stop	
tRNA-Phe	68	1	68	0	GAA			H
12S rRNA	945	69	1013	0				H
tRNA-Val	74	1014	1087	0	TAC			H
16S rRNA	1696	1088	2783	0				H
tRNA-Leu	74	2784	2857	0	TAA			H
ND1	975	2858	3832	0		ATG	TAG	H
tRNA-Ile	71	3837	3907	4	GAT			H
tRNA-Gln	71	3907	3977	−1	TTG			L
tRNA-Met	69	3977	4045	−1	CAT			H
ND2	1046	4046	5091	0		ATG	TA	H
tRNA-Trp	71	5092	5162	0	TCA			H
tRNA-Ala	69	5164	5232	1	TGG			L
tRNA-Asn	73	5234	5306	1	GTT			L
tRNA-Cys	67	5341	5407	34	GCA			L
tRNA-Tyr	71	5408	5478	0	GTA			L
COX1	1560	5480	7039	1		GTG	AGG	H
tRNA-Ser	71	7031	7101	−9	TGA			L
tRNA-Asp	73	7105	7177	3	GTC			H
COX2	690	7186	7875	8		ATG	TAA	H
tRNA-Lys	74	7877	7950	1	TTT			H
ATP8	168	7952	8119	1		ATG	TAA	H
ATP6	683	8110	8792	−10		ATG	TA	H
COX3	785	8793	9577	0		ATG		H
tRNA-Gly	71	9578	9648	0	TCC			H
ND3	349	9649	9997	0		ATG	T(aa)	H
tRNA-Arg	69	9998	10066	0	TCG			H
ND4L	297	10067	10363	0		ATG	TAA	H
ND4	1381	10357	11737	−7		ATG	T(aa)	H
tRNA-His	69	11738	11806	0	GTG			H
tRNA-Ser	68	11807	11874	0	GCT			H
tRNA-Leu	73	11880	11952	5	TAG			H
ND5	1839	11953	13791	0		ATG	TAA	H
ND6	522	13788	14309	−4		ATG	TAG	L
tRNA-Glu	69	14310	14378	0	TTC			L
CYTB	1141	14384	15524	5		ATG	T(aa)	H
tRNA-Thr	72	15525	15596	0	TGT			H
tRNA-Pro	70	15596	15665	−1	TGG			L
D-loop	862	15666	16527	0				-

Table IV.- Comparison of the base content in different genes or region of *Ammodytes personatus* mitochondrial genomes.

Gene/region	Base composition (%)				
	T	C	A	G	G+C
Protein coding					
1st	22	28.4	25.8	23.6	52.0
2nd	37	29.7	19.9	13.1	42.8
3rd	25	36.1	25.1	13.6	49.7
Total	28.3	31.4	23.6	16.8	48.2
tRNA	24.8	24.7	30.3	20.3	45.0
rRNA	24.8	24.7	30.3	20.3	45.0
Control region	24.8	24.7	30.3	20.3	45.0
Overall	26.8	29.4	25.8	17.9	47.3

Taking AP1 as an example, the total length of the 13 PCG genes is 11,436 bp, and 3,803 codons are found aside from the terminating codon. Using MEGA6.0 to analyse base use frequency, the bias of the first loci of codons is the smallest and the second loci bias is the largest. An obvious base bias is found. The T and G contents are 37% and 13.1%, respectively (Table IV). The above base shift is also common in other vertebrates. Analysis of the average use frequency of codon and the average use frequency of relative synonymous codon indicate that 32 preference codons (RSCU ≥ 1) are found in the 13 PCG genes of *A. personatus*, of which the codon with the third locus as U or C has a high frequency of use in the encoding of the same amino acid. The preference of codon third locus to U and C is positively related to the third locus of protein coding codon to T and C bias.

The difference in protein coding genes between the two lineages is very significant. The difference between AP1 and AP3 is 932 bases, accounting for 8.15% of the total length of protein coding genes. The difference between AP2 and AP3 is 937 bases, accounting for 8.19% of the total length of protein coding genes. The differences in protein sequences between the two lineages are shown in Table V.

rRNA genes

This study found that like that of other vertebrates, the *A. personatus* mitochondrial genome contains two rRNAs: 12S rRNA (between tRNA-Phe and tRNA-Val) and 16S rRNA (between tRNA-Val and tRNA-Leu) with lengths of 945 bp and 1,698 bp (AP1, AP2) and 1,696 bp (AP3), respectively. The rRNA second-structure is generally conservative and contains many stem ring structures. The rRNA gene is very conservative in evolution, and it has the slowest evolution in the mitochondrial genome. At the same time, 12S rRNA gene is more conservative than 16S rRNA gene. The difference of rRNA gene between the two lineages was not significant, showing a high conservatism.

Table V.- Amino acid variation information between two lineages.

Gene	Variation position	Amino acids	
		AP1	AP3
ND1	8	I	V
	18	I	V
ND2	142	I	V
	158	A	T
	310	V	I
	333	C	G
	338	T	A
	342	M	V
ATP6	17	I	V
	21	V	A
	144	I	V
	188	T	A
COX3	159	A	T
ND3	17	V	I
	90	I	V
ND4L	55	S	N
ND4	46	A	T
	177	I	L
ND5	134	I	V
	212	S	G
	479	I	V
	485	V	I
	498	V	I
	540	I	V
	587	V	I
CYTB	353	I	V
	371	L	M

tRNA genes

The location and secondary structure prediction of 22 tRNA genes in AP1 mitochondrial DNA revealed that 22 tRNA genes are distributed between 13 protein-coding genes with size ranging from 66 bp to 74 bp and a total length of 1,556 bp. Through comparative analysis, the three pairs of adjacent tRNA (tRNA-Ile and tRNA-Gln, tRNA-Gln and tRNA-Met, and tRNA-Thr and tRNA-Pro) have 1 overlapping nucleotide in the *A. personatus* tRNA gene, which is also observed in other fish, with the difference in the number of folded bases. Fourteen of the 22 tRNA genes are encoded by H- chains, and the other 8 are encoded by L- chains. The G+C base content of all tRNA genes was higher than that of the control region, but lower than that of PCG and rRNA genes. Except for tRNA-Ser (AGN),

the second-structure of other tRNAs are of typical clover structure. The difference of tRNA gene between the two lineages was not significant, showing a high conservatism.

Control region

Although the A+T-rich control region does not encode proteins, it plays an important role in regulating mtDNA replication and transcription. In this study, a control region consisting of a sequence of 870 bp was identified between the tRNA-Pro and the tRNA-Phe genes. The region contains a sequence of initiation of H- chain replication (Fig. 1). Compared with other Perciformes fishes, the control area was relatively conservative except for the insertion/deletion in some loci, and no significant difference in the sequence length and A+T content in different species was found (Supplementary Table I).

Fig. 2. Schematic map characterising the control region of *Ammodytes personatus*.

Similar to many vertebrates, the control areas in *A. personatus* can be divided into three domains (Fig. 2): (i) the termination region is also known as the Hypervariable Domain, which has the largest variability among the three regions. The terminating sequence area of *A. personatus* has two sequences, namely, TAS-1 and TAS-2, which control the termination of mitochondrial DNA replication. The sequence can form a thermally stable stem ring structure containing 32 bases from two pairs of core sequences TACAT and their complementary ATGTA sequence. (ii) The central conservative region contains three conservative sequences, namely CSB-D, CSB-E and CSB-F. The key sequence of CSB-F is AGAACCGACCA, which is considered to be a marker for distinguishing between the terminating region and the central conservative region (Zhao *et al*., 2006). The existence of GTGGG-box in CSB-E is widely recognised by researchers (Lee *et al*., 1995). In this study, the key sequence of the GTGGG-box in *A. personatus* is AGGGACA ACTATTGTGGGGG. CSB-D is highly variable in Perciformes fishes. The key sequence identified in *A. personatus* is TATTCCTGGCATTTG. (iii) In the conservative sequence, the three conserved sequences (CSB1, CSB2 and CSB3) were identified in *A. personatus*. The conservatism of CSB-1 in conservative sequences is relatively low (Broughton *et al*., 2001). Similar to other teleost fishes, *A. personatus* followed a TGCCCCC conservative sequence after the sequence. The key sequence of *A. personatus* CSB-1 identified in this study is CTTGGCATCTCACA-GTGCATGCTAACCTATAGGAA. By contrast, the conservatism of CSB-2 and CSB-3 sequences is higher than that of CSB-1. The two conservative sequences in *A. personatus* mitochondria are AAACCCCCCTACCCCCCTAAA and TGAAAACCCCCCGGAAACA.

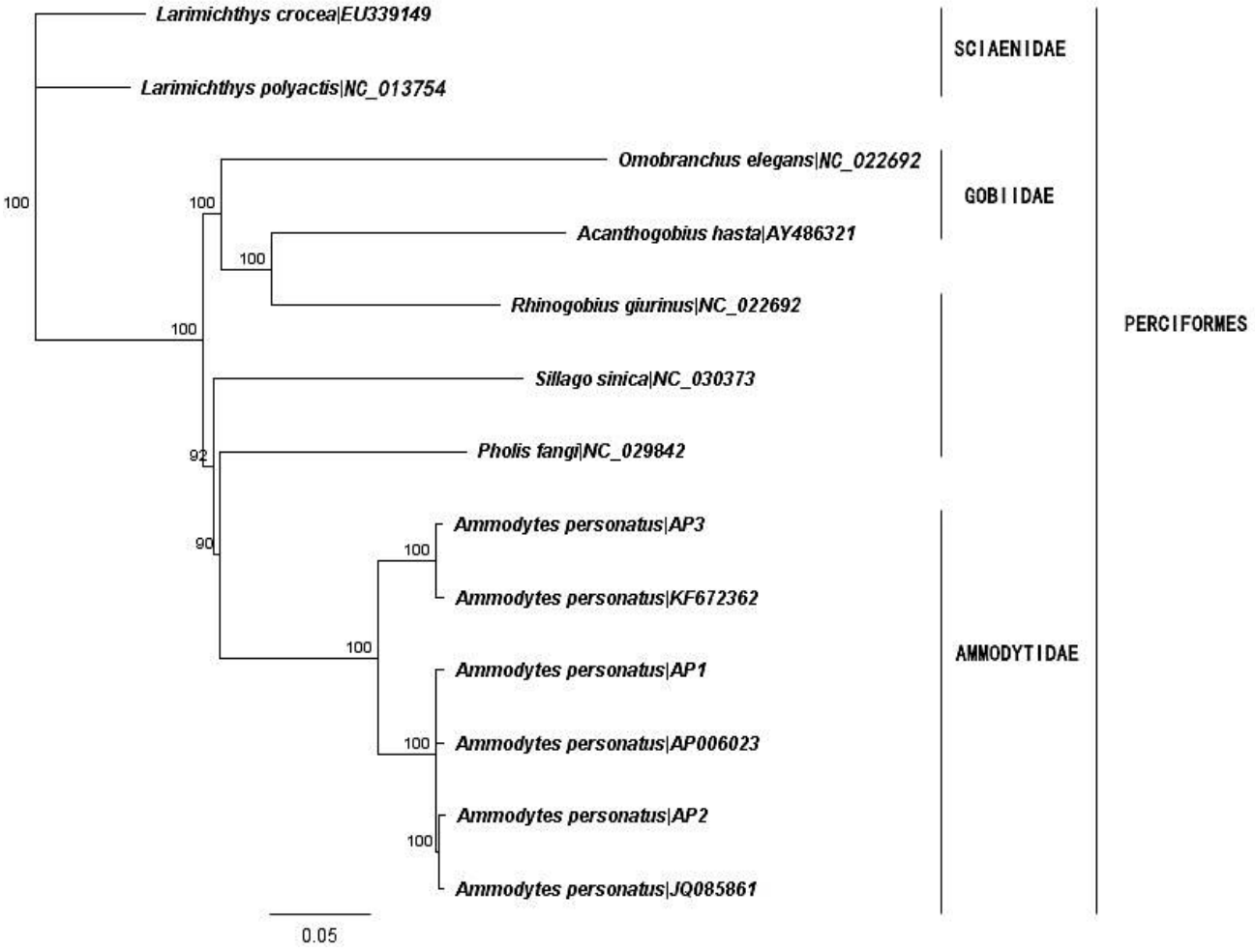

Fig. 3. Phylogenetic tree based on the NJ analysis of the whole Perciformes mitochondrial genome.

Non-coding regions outside the control area

The non-coding regions of the mitochondrial genome of vertebrates are some small fragments that regulate the replication and transcription of mtDNA, and are mainly distributed between the L-chain replication initiation region and the tRNA gene. According to the structural composition and location of the non-coding region, the region can also be divided into a gene interval sequence region and a gene overlap area. In the AP1 mitochondrial region, seven overlapping regions of gene sequences were identified with a total length of 33 bp. The largest base overlap of 10 bp is located between ATP6 and ATP8 genes, whereas the minimum base overlap is 1 bp. A total of 11 spacer sequences in the non-coding region with a length ranging from 1 bp to 34 bp were found. Among the tRNA-Asn and tRNA-Cys genes, the largest spacer sequences were found (34 bp), and the sequence of L-chain replication was identified in the region. The key sequence in the Perciformes fishes is 5′-GCCTA-3′, which forms a conservative stem ring structure from 10 bases. In the mitochondrial genome of vertebrates, the stem ring structure mainly regulates the replication of L-chains (Shadel *et al*., 1997).

Phylogenetic analysis

The whole mitochondrial genome sequence phylogenetic tree of Perciformes fishes was constructed using the NJ method (Fig. 3). Results showed that the individual AP1 and AP2 were clustered together with individuals reported by Gao *et al*. (2013) and Satoh *et al*. (2016). The individual AP3 was clustered together with individuals reported by Li *et al*. (2015). The phylogenetic tree of other Perciformes fishes is the same as that of traditional classification.

COX1 can be used as a general label to classify species effectively. In this study, COX1 was used to construct the phylogenetic tree using the NJ method (Fig. 4). In view of the special status of the sand lance, its systematic relationship with other Perciformes fishes must be confirmed through further research.

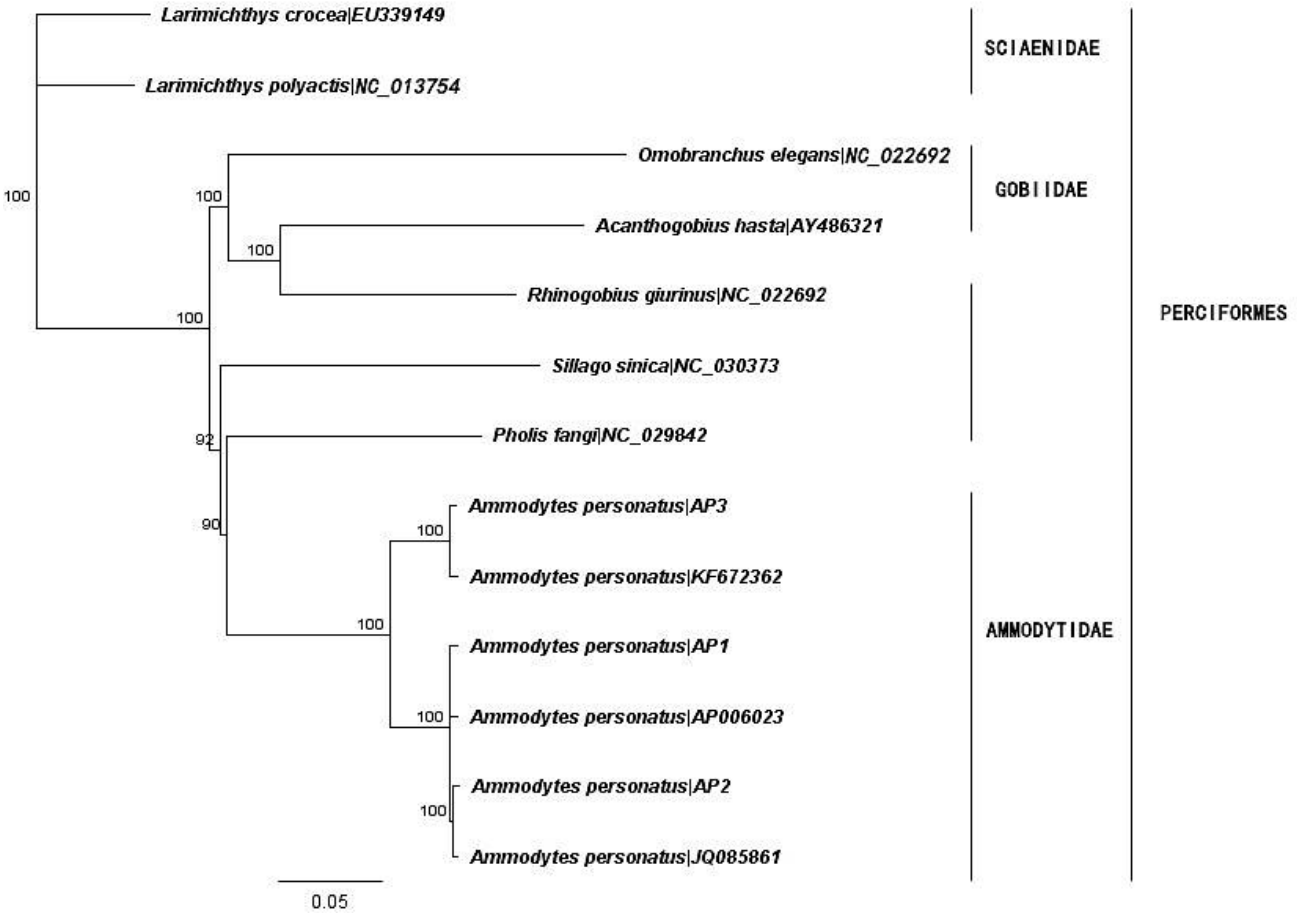

Fig. 4. Phylogenetic tree based on the NJ analysis of the COX1 genes.

Two Perciformes species, namely, *L. crocea* and *L. polyactis* were used as outgroups. Compared with the evolutionary trees constructed using the two regions (whole mitochondrial genome and COX1), results showed that AP1 and AP2 were clustered together with individuals as reported by Gao *et al*. (2013) and Satoh *et al*. (2016); AP3 was clustered together with the individual reported by Li *et al*. (2015). They get together with the *Pholis fangi*. *P. fangi* is a small Zoarcidae fish living near coastal rocks, the body is slender and flat. In traditional taxonomy, Zoarcoidei is closely related to the Ammodytoidei. This classification is consistent with the results of this study. The difference between the two methods is that the Gobiidae species was properly clustered into a branch in the whole mitochondrial genome NJ tree, which is consistent with classical taxonomy.

DISCUSSION

This study found that the complete *A. personatus* mitochondrial genome has similar gene composition and obvious AT bias to those of other typical vertebrates. The mtDNA characteristics of high A+T content and low G+C content are also present in other reported Perciformes fishes, but A+T or G+C content varies from species to species. At present, the order of mitochondrial genes in most fish is consistent with that of other vertebrates. Most L-chain in fishes encode the 8 tRNA and ND6 genes, while the remaining 28 genes are encoded by H- chain, which is consistent with the gene-coding position of *A. personatus* in this study.

In the 13 protein coding genes in *A. personatus*, the initiation codon is consistent with the results of most teleost fish and has two relatively stable initiation codons (GTG, ATG). Relative to the stability of the initiation codon, termination codon is more variable among different species. In *Sebastiscus marmoratus*, ND1, ND2, COX3, ND4, COX2, ND4L, ND5, ATP6 and ATP8 genes used TAA as the terminating codon; COX1 used AGA as the terminating codon; and ND3, ND6 and CYTB used incomplete TA or T as the terminating codon. In *Abbottina rivularis*, COX1, ND4L, ND5, ATP6 and ATP8 genes used TAA as the terminating codon; ND1, ND2, ND6 used TAG as the terminating codon; and COX2, COX3, ND3, ND4 and CYTB used incomplete TA or T as the terminating codon. In *Pseudorasbora parva*, ATP6, COX1, ND2, ND4L, ND5 and ND6 genes used TAA as the terminating codon; ND1, ND2, ND6 used TAG as the terminating codon; ATP8 and ND1 genes used TAG as the terminating codon; and CYTB, ND4, ND3, COX2 and COX3 genes used T as the terminating codon. In this study, *A. personatus* ND4L, COX2, ND5 and ATP8 genes used TAA as the terminating codon; TAG is the terminating codon of ND1; AGG is the terminating codon of COX1; ND2, ND6, ND3, COX3, ATP6, ND4 and CYTB genes used incomplete T and TA as terminating codons. The above results suggest that the protein-coding genes of different fish use different terminating codons. Most of the protein-coding genes used TAA or TAG as complete terminating codon, T and TA as incomplete codon, and some used AGA or AGG as terminating codon. Other special codons were not found. The phenomenon of incomplete termination of codon in the mitochondrial protein coding gene of vertebrates is common. Ojala *et al*. (1981) and other researchers speculate that incomplete codons are caused by the addition of A residues at the post-transcriptional mRNA 3′ end.

The control areas of fish include the termination sequence area (TAS), the central conservative region (CD), the conservative sequence area (CSB) and the H-chain replication initiation region (Guo *et al*., 2004). In this study, only TAS, CD and CSB regions were found in the non-coding region of mtDNA. In the TAS region, we found two pairs of core sequence TACAT and ATGTA that complement each other to form a thermal stable hairpin containing 32 bases and participate in the termination regulation of mitochondrial genome replication. As a key sequence of CSB-F, the AGAACCGACCA gene sequence is highly conserved in Perciformes (Zhao *et al*., 2006). The AGAACCGACCA gene sequence is considered to be a marker for distinguishing between terminating regions and central conserved areas. This study found that the key sequence in the initiation region of the mitochondrial L-chain is 5′-GCCTA-3′, which plays an important role in the synthesis of RNA and DNA. Its position in different Perciformes species is relatively conservative, but the nucleotide sequence is different.

In traditional taxonomy, the *A. personatus* taxonomy is special, The more obvious morphological difference between the two lineages is mainly reflected in the gill rake. The number of gill rakes of individuals in Aomori is slightly higher than that of Qingdao individuals (Ji *et al*. 2006). Pietsch and Zabetian (1990) believed that Ammodytids should be the sister group of Trachinidae and Uranoscopidae, and gave some morphological evidence. However, great morphological differences exist between Ammodytoidei and Trachimoidei species. Many taxonomists still regard it as a suborder. At present, only the complete mitochondrial genome sequences of *A. personatus* were determined, whereas only those of COX1 were measured in the rest of the Ammodytoidei species. In addition, Trachimoidei, which are closely related to Ammodytoidei in traditional taxonomy, belong to less studied species. Han *et al*. (2012) reported two distinct *A. personatus* populations in the coastal areas of Japan

and China. In this study, AP3 individuals from Aomori, Japan and AP and AP2 individuals from Qingdao were significantly differentiated into two branches, which is consistent with the results of this study.

In the study of fish mitochondria, related genes or structural regions are widely used in the field of phylogeny, species classification and population identification as molecular genetic markers. For example, using mitochondrial 16S rRNA and ND2 gene markers, Near *et al*. (2003) successfully analysed the evolutionary relationship of Antarctic icefishes. Tinti *et al*. (1999) have amplified the mtDNA control region sequence of three species of Pleuronectiformes in the Adriatic Sea, and by combining the control region sequences of other six species of Pleuronectidae, the molecular system relationship was studied with results consistent with morphological classification. Morin *et al*. (2010) studied the geographical population and systematic classification of *Orcinus orca* by using the complete sequence and control region sequence of mitochondrial genome. They found that the complete sequence of mitochondrial genome was more accurate than the control area in species identification, and the new geographic population was divided by the complete sequence of mitochondrial genome. Gilles *et al*. (2001) studied the phylogenetic relationships among the subfamilies of European cyprinid fishes using mtDNA control region complete sequence, and the molecular phylogenetic tree obtained from the control region sequence was more accurate and reliable than the result of Cytb and 16S rRNA sequences. Sun *et al*. (2002) used mtDNA control region sequence to analyse the genetic diversity in the mitochondrial control region of Chinese sucker (*Myxocyprinus asiaticus*). Results showed that the variation of mitochondrial control area of Chinese sucker was much greater than that of *Moxostoma robustum*.

In recent years, the mtDNA COX1 gene is widely used in the field of germplasm identification, species classification and phylogenetic studies because of its moderate length and evolution rate and rich phylogenetic information. For instance, Hebert (2003) used COX1 as a DNA barcode, and studied the phylogenetic relationship of 200 closely related lepidopteran insects. Zhang (2011) used COX1 gene to effectively divide and analyse the evolutionary and classification relationship of marine fishes. In this study, the phylogenetic relationship and species identification of *A. personatus* and other Perciformes fishes were carried out using complete mtDNA and COX1 sequences, respectively. Results showed that the above two markers could effectively analyse the phylogenetic and taxonomic relationships of Perciformes fishes. However, in view of the complexity of Perciformes fish, learning from successful experiences of mitochondrial phylogenetic analysis of other fishes is necessary to further develop the phylogenetic relationship and taxonomic identification of Perciformes fishes, especially those of the Ammodytoidei fishes.

ACKNOWLEDGEMENTS

We sincerely thank the reviewer for the constructive criticisms and valuable comments. The research was funded by the National Natural Science Foundation of China (31472281).

Supplementary material

There is supplementary material associated with this article. Access the material online at: http://dx.doi.org/10.17582/journal.pjz/2019.51.5.1869.1880

Statement of conflict of interest

The authors declare no conflicts of interest.

REFERENCES

Broughton, R.E., Milam, J.E. and Roe, B.A., 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genom. Res*., **11**: 1958-1967. https://doi.org/10.1101/gr.156801

Farjallah, S., Amor, N., Merella, P. and Said, K., 2012. Pattern of genetic diversity of North African green frog *Pelophylax saharicus* (amphibia) in Tunisia. *Pakistan J. Zool*., **44**: 901-907.

Gao, T., Li, N., Zhang, Y. and Shi, P., 2013. The complete mitochondrial genome of Japanese sandeel *Ammodytes personatus* (Perciformes, Ammodytidae): Rare structure in control region compared. *Mitochondrial DNA*, **24**: 320-322. https://doi.org/10.3109/19401736.2012.760074

Gilles, A., Lecointre, G., Miquelis, A., Loerstcher, M., Chappaz, R. and Brun, G., 2001. Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. *Mol. Phylogen. Evolut*., **19**: 22-33. https://doi.org/10.1006/mpev.2000.0916

Guo, X.H., Liu, S.J., Liu, Q. and Liu, Y., 2004. New progresses on mitochondrial DNA in fish. *J. Genet. Genom*., **31**: 983-1000.

Han, Z., Yanagimoto, T., Zhang, Y. and Gao, T., 2012. Phylogeography study of*Ammodytes personatus* in Northwestern Pacific: Pleistocene isolation, temperature and current conducted secondary contact. *PLoS One*, **7**: e37425. https://doi.org/10.1371/journal.pone.0037425

Hebert, P.D.N., 2003. Biological identification through DNA barcodes. *Proc. R. Soc. B: Biol. Sci*., **270**: 313-321. https://doi.org/10.1098/rspb.2002.2218

Ingman, M. and Gyllensten, U., 2006. *Vertebrate mitochondrial DNA: Reviews in cell biology and molecular medicine*. Wiley-VCH Verlag GmbH and Co, KGaA.

Ji, Y.P., Gao, T.X., Chen, Y.C. and Yanagimoto, T., 2006. A comparative study on the morphology and genetics of sandeel populations. *Period. Ocean Univ. China*, **36**:77-83.

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

Li, N., Song, N. and Gao, T., 2015. The complete mitochondrial genome of Japanese *Ammodytes personatus* (Perciformes, Ammodytidae). *Mitochondrial DNA*, **26**: 781. https://doi.org/10.3109/19401736.2013.855750

Liu, Z., Song, N., Yanagimoto, T., Han, Z., Shui, B. and Gao, T., 2018. Complete mitochondrial genome of three species (Perciformes, Amblyopinae): Genome description and phylogenetic relationships. *Pakistan J. Zool.*, **50**: 1999-2398

Morin, P.A., Archer, F.I., Foote, A.D., Vilstrup, J., Allen, E.E., Wade, P., Durban, J., Parsons, K., Pitman, R., Li, L., Bouffard, P., Nielsen, S.C.A., Rasmussen, M., Willerslev, E., Gilbert, M.T.P. and Harkins, T., 2010. Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genom. Res*., **20**: 908. https://doi.org/10.1101/gr.102954.109

Moritz, R.F.A., Cornuet, J.M., Kryger, P., Garnery, L. and Hepburn, H.R., 1994. Mitochondrial DNA variability in South African honeybees (*Apis mellifera*). *Apidologie*, **25**: 169-178. https://doi.org/10.1051/apido:19940205

Near, T.J., Pesavento, J.J. and Cheng, C.H., 2003. Mitochondrial DNA, morphology, and the phylogenetic relationships of Antarctic icefishes (*Notothenioidei channichthyidae*). *Mol. Phylogen. Evolut*., **28**: 87-98. https://doi.org/10.1016/S1055-7903(03)00029-0

Ojala, D., Montoya, J. and Attardi, G., 1981. Trna punctuation model of RNA processing in human mitochondrial. *Nature*, **290**: 470-474. https://doi.org/10.1038/290470a0

Okamoto, H., Yamazaki, F. and Mishima, S., 1988. Genetic divergence among sand-lance *Ammodytes personatus* populations in Japan. *Nsugaf*, **54**: 1297-1304. https://doi.org/10.2331/suisan.54.1297

Pan, Q., Lei, F., Yang, S., Yin, Z., Huang, Y., Tai, F. and Kristin, A., 2006. Phylogeny of some Turdinae birds based on mitochondrial cytochrome b gene sequences. *Acta Zool. Sin*., **52**: 87-98.

Pietsch, T.W. and Zabetian, C.P., 1990. Osteology and interrelationships of the sand lances (teleostei: Ammodytidae). *Copeia,* **1**: 78.

Satoh, T.P., Miya, M., Mabuchi, K. and Nishida, M., 2016. Structure and variation of the mitochondrial genome of fishes. *BMC Genom*., **17**: 719. https://doi.org/10.1186/s12864-016-3054-y

Shadel, G.S. and Clayton, D.A., 1997. Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem*., **66**: 409-435. https://doi.org/10.1146/annurev.biochem.66.1.409

Sun, Y.H., Wang, W., Liu, S.Y., He, S.P., Shao, X.L. and Xie, Z.X., Deng, F.J., Liu, Y., Tong, J.G. and Wu, Q.J., 2002. Genetic diversity analysis of mitochondrial D-loop region of Chinese sucker (*Myxocyprinus asiaticus*). *J. Genet. Genom*., **29**: 787-790.

Tinti, F., Colombari, A., Vallisneri, M., Piccinetti, C. and Stagni, A.M., 1999. Comparative analysis of a mitochondrial DNA control region fragment amplified from three Adriatic flatfish species and molecular phylogenesis of pleuronectiformes. *Mar. Biotechnol*., **1**: 20-24. https://doi.org/10.1007/PL00011747

Yamada, H., Tsumoto, K. and Kuno, M., 1998. Cannibalistic mortality of larval sand eel *Ammodytes personatus* by adults in Ise Bay central Japan. *Nsugaf*, **64**: 807-814. https://doi.org/10.2331/suisan.64.807

Yamashita, Y. and Aoyama, T., 1986. Starvation resistance of larvae of the Japanese sand eel *Ammodytes personatus*. *Nsugaf*, **52**: 635-639. https://doi.org/10.2331/suisan.52.635

Zhang, J., 2011. Species identification of marine fishes in china, with DNA barcoding. *Evidence-Based Complem. Altern. Med*., **2011**: 978253.

Zhao, J.L., Wang, W.W., Li, S.F. and Cai, W.Q., 2006. Structure of the mitochondrial DNA control region of the Sinipercine fishes and their phylogenetic relationship. *Acta Genet. Sin*., **33**: 793-799. https://doi.org/10.1016/S0379-4172(06)60112-1