Report on the assessment of genetic diversity, population structure of spotted seal based on mtDNA of in the Yellow Sea

First Institute of Oceanography, Ministry of Natural Resources

May 2020

Content

1 Background

The spotted seal, *Phoca largha*, which belongs to Carnivore, Phocidae, is widely distributed in the north and west of the North Pacific Ocean. It is mainly distributed in the Bohai Sea and the Yellow Sea in China, occasionally in the South China Sea. The spotted seal is the only pinniped animal that can reproduce in the sea area of China, and one of second-class protected endangered animals in China.

Most of the researches on spotted seals mainly focus on its conventional field investigation of ecology and general biology, including the living habits, population distribution, and quantity distribution and so on. However, the studies on population diversity and structure are very limited. Several studies have been reported the genetic diversity of spotted seal in Liaodong Bay. However, only Mizuno et al. evaluated the population genetic structure of the spotted seal along the coast of Hokkaido based on mitochondrial DNA. The results showed a high level of diversity but no genetic structure, and did not deny the possibility that seals in the Okhotsk breeding concentration mainly stayed in the fall Okhotsk and also inhabited in the winter Sea of Japan. Wang (2003) sequenced 15 spotted seal individuals from Liaodong Gulf. The results implied population from Laodong Gulf may undergo a genetic bottleneck effect, and it was thought to be an independent population which had not genetic exchanges with those along Sea of Japan and Sea of Okhotsk.

The direct sequencing of DNA sequence can accurately reflect the variation of bases among individuals, which is one of the most widely used genetic analysis methods. Mitochondrial DNA (mtDNA, mitogenome) is the only hereditary material outside the zooblast nucleus. It is usually a circular, double stranded molecule of 16-19 kb in length that comprises a set of 37 genes for 22 tRNAs, 2 rRNAs, and 13 proteins (Miya and Nishida, 2015). The mtDNA has many obvious characteristics such as simple structure, fast evolution, low ratio of gene rearrangement and maternal inheritance. At present, it has become a powerful tool to study the origin and phylogenetic inferences of species, genetic differentiation between interspecific and/or intraspecific populations. Mitochondrial DNA control region has the advantages of rapid base substitution rate, simple structure, strict adherence to maternal inheritance, and almost no recombination. It

is widely used in the research fields of population genetics and intraspecific micro level phylogenetic relationship. Although there have been some reports about its genetic research, but in recent years, there are few studies. As an important protected animal, it is imperative to carry out the research on its genetic background. In the present study we collect 66 individuals from Liaodong Gulf and fragment of mitochondrial DNA was employed to analyze the population diversity status.

In this report, the complete mtDNA sequence of the spotted seals collected from the Liaodong Gulf, China was determined by long and accurate polymerase chain reaction (LA-PCR) and primer walking technology. Comparative analysis on gene length, structure, organization and polymorphic sites among Liaodong (CPL), Korean (KPL) and Alaskan (APL) individuals are carried out as well. Meanwhile, we amplified the mt DNA control region of 66 spotted seal individuals from Liaodong Bay, and compared the genetic diversity and population structure with population from Russia.

2 Materials and methods

2.1 Study methods

The original purpose of this research is to evaluate the genetic diversity of spotted seal and compare the genetic differentiation with population in others region. In previous, we planned to use environmental-DNA to research the population diversity and genetic structure, because we could not get muscle or blood sample of spotted seal. Environmental-DNA is the DNA extracted directly from environmental samples (e.g. sediment, feces, water, soil or air) without isolation of a particular organism, it is widely applied in genetic diversity research of marine mammals. However, due to the degradation of genomic DNA in water, its accurate degree is lower than extracting DNA from muscle or flood. In 2019, we rescued the 66 individuals in China, and got the blood samples, and 66 individuals are enough to evaluate the genetic diversity and population structure of spotted seal. Therefore, we extracted DNA from blood, and combined two DNA markers (mtDNA control region of 66 individual and mtDNA genome of one

individual) to estimate the genetic diversity and genetic structure compared with others regions population (public sequence from GeneBank).

2.2 Sample and sequencing

In total, 66 individuals from Liaodong Gulf were used in the present study (Fig. 1, Table 1). All individuals were identified on the basis of morphology, and the blood was taken from each individual and preserved in -20℃ for DNA extraction. Genomic DNA was isolated from muscle by proteinase K digestion, followed by a standard phenol–chloroform method.

The segment of the mtDNA control region was amplified using all the samples. A total of 17 individuals from Liaodong Gulf at 2005 and 28 individuals from Russia coastal waters at 2003 were downloaded from GenBank were used in the comparative analysis of genetic diversity and population structure. One individual was used to sequencing the whole mtDNA genome with long PCR amplification.

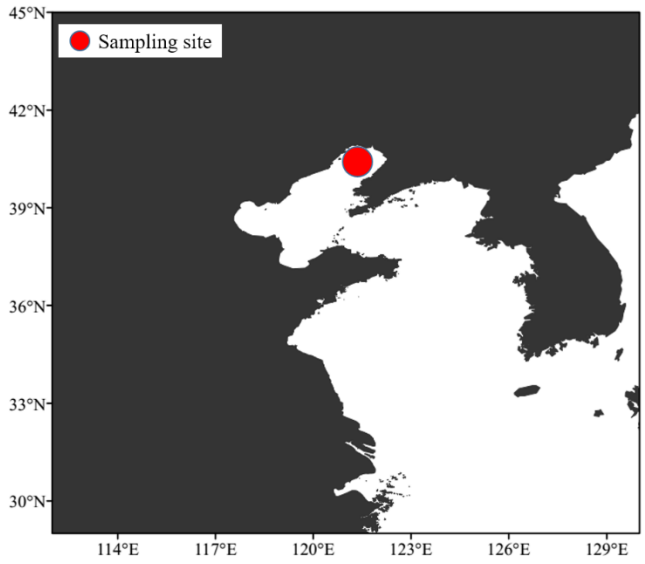

Fig.1 Sampling site in the present study

Table 1 Sampling information and genetic diversity indices for three populations

Localities	Abbreviation	Time	N	NH	*h*	π
Liaodong Gulf	LD2019	2019	66	22	0.8308±0.0398	0.0096±0.0054
Liaodong Gulf	LD2005	2005	17	13	0.9632±0.0328	0.0055±0.0035
Russian Coastal waters	RU	2003	28	14	0.9153±0.0345	0.0120±0.0067

N, sample number; NH, haplotype numbers; h, haplotype diversity; π, nucleotide diversity

2.3 Mitochondrial genome gene prediction and phylogenetic analysis

PCR products are directly sequenced after purification. Sequence assembly, annotation, analysis and nucleotide composition calculations are conducted with DNAstar (DNAStar, www.dnastar.com) and MEGA 6 (Tamura *et al*., 2007). The sequences on the minority strand (also namely light strand, L strand) are reversely complemented in EditSeq, and the neighboring sequences are aligned by ClustalW version 1.6 (Thompson *et al*., 1994) as implemented in MEGA 6 to find the overlapping regions. With the help of MEGA 6 and EditSeq, the sequence assembly and annotation are conducted in the Staden sequence analysis package (Staden *et al*., 2000). The locations of protein-coding genes and rRNA genes are identified by comparison with those of other *Phoca* species, while tRNA genes are identified using the tRNAscan-SE server (Lowe and Eddy, 1997). Potential secondary structure folds in the control region are predicted using Mfold v. 3.2 (Zuker, 2003).

ClustalW version 1.6 and MEGA 6 were used to compare the mtDNA sequence of *Phoca largha* from Liaodong, Korean and Alaskan. In this study, we mainly investigate the variation of nucleotide and amino acid sequences of tRNAs, rRNAs, protein-coding

genes and non-coding regions. Seventeen complete mitogenomes belonging to 15 Phocidae species are included in analyses (Table 2). 13 protein-coding genes are used to construct the phylogenetic relationships in all species. Each gene is individually aligned using ClustalW. All stop codons are excluded from the analysis. The possible bias of substitution saturation at each codon position of protein-coding genes is investigated using DAMBE v.4.5.57 (Xia and Xie, 2001), and the results suggested that the third codons position are saturated both for transitions and transversions in the plot against with pairwise sequence divergence. MEGA6 were used to construct the neighbor-joining tree.

Table 2 List of sequence used in the phylogenetic analysis

Species	GenBank accession number
Phoca largha	FJ895151, NC_008430
Phoca vitulina	NC_001325
Pusa hispida	NC_008433
Pusa sibirica	NC_008432
Pusa caspica	NC_008431
Halichoerus grypus	NC_001602
Phoca groenlandica	NC_008429
Terapon jarbua	NC_027281
Phoca fasciata	NC_008428
Cystophora cristata	NC_008427
Erignathus barbatus	NC_008426
Monachus schauinslandi	NC_008421
Lobodon carcinophaga	NC_008423
Leptonychotes weddellii	NC_008424
Mirounga leonina	NC_008422
Mirounga leonina	NC_008422
Lipotes vexillifer	NC_007629
Dugong dugon	NC_003314

2.4 mtDNA control region data analysis

All sequences were edited and aligned using DNASTAR software (DNASTAR Inc., Madison, USA). Some sequences from GenBank were downloaded to compare the genetic diversity. Genetic diversity indices such as number of haplotypes, polymorphic sites, transitions, transversions, and indels were obtained using the program ARLEQUIN version 3.5 (Excoffier and Lischer 2010). The haplotype diversity (*h*), nucleotide diversity (π) and the mean number of pairwise differences (*k*) were calculated using the program ARLEQUIN version 3.5 (Excoffier and Lischer 2010). A neighbour-joining tree of the haplotypes was constructed using MEGA5.0 and evaluated with 1000 bootstrap replicates (Tamura et al. 2011).

Pairwise genetic divergences between different populations were tested by the fixation index F_{st} (Excoffier et al. 1992). The significance of the F_{st} was tested by 10,000 permutations for each pairwise comparison in ARLEQUIN. When multiple comparisons were performed, *P* values were adjusted using the sequential Bonferroni procedure (Rice 1989). To further examine hierarchical population structure as well as the geographical pattern of population subdivision, we used analysis of molecular variance (AMOVA) (Excoffier et al. 1992). The D test of Tajima (1989) and *F*s test of Fu (1997) were used to test for neutrality. Significant negative D and *F*s statistics can be interpreted as signatures of population expansion. The concordance of the observed with the expected distribution in the sudden-expansion model was tested by a least-squares approach (Rogers & Harpending, 1992). The value of τ was transformed to estimates of real time since expansion with the equation $\tau = 2\times\mu t$ where μ is the mutation rate for the whole sequence under study and t is the time since expansion. Both mismatch analysis and neutrality tests were performed in ARLEQUIN.

3 Results

3.1 Genome structure, organization, and composition

The complete mitogenome of *Phoca largha* from Liaodong Gulf is circular, 16754 bp in length. It has 13 protein coding genes, 22 tRNA genes, and 2 rRNA genes, as is also the case in other marine animals. The 13 protein-coding genes included: 1 cytochrome b gene (CytB), 2 atpase subunits (ATPase6, ATPase8), 3 cytochrome oxidase subunit gene (CO I, CO II, CO III) and 7 NADH dehydrogenase subunit gene (ND1-6, ND4L). The two rRNA genes were 12S rRNA and 16S rRNA, respectively. The ND6 gene is located on the minority strand (L strand), and the remaining 12 protein-coding genes are all expressed by majority strand (H strand). The ND6, tRNAAla, tRNAAsn, tRNACys, tRNATyr, tRNASer(UCN), tRNAGlu and tRNAPro genes were 1ocated on the light strand (L strand), and the remainders lied on the heavy strand (H strand).

The mitochondrial genomes of the seal from Liaodong Gulf were compared with those from Western Coast of South Korean (GenBank: FJ895151) and Alaska (GenBank: NC_008430), American. The results showed that the length of three mitochondrial genomes was slightly different in length, with 16754, 16728 and 16701 bp, respectively. The main variation occurred in the control region (O_H), and the length of O_H was 1316, 1290 and 1264 bp, respectively. Thirty seven genes of the three mitochondrial genomes were compared and 5 nucleotide variation were found in rRNAs. As for tRNA, one base deletion was found in the $tRNA^{Leu}$ sequence of Alaskan. The $tRNA^{Phe}$, $tRNA^{Ala}$, $tRNA^{Tyr}$, $tRNA^{Thr}$ have one nucleotide mutation as well. However, the remaining 17 tRNAs were identical. Of 13 protein-coding genes, 8 genes (ATPase6、CO I、CO II、CytB、ND1、ND3、ND4、 ND4L) were relatively conservative, and their amino acid sequences were completely consistent among each other. The remaining 5 genes showed some amino acid sequence variation (Table 3). Among them, the nucleotide mutation of ND5 is the largest. There are 18 mutation sites in ND5 (1821 bp, in total). The mutation rate is 0.99% and defining 4 amino acid sequence mutations.

Table 3 Comparative analysis of three protein-coding gene mutation of spotted seals

Gene	Total variation		CPL/KPL		KPL/APL		APL/CPL	
	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid
ND1	2	0	0	0	2	0	2	0
ND2	6	2	2	0	5	2	5	2
CO I	9	0	2	0	8	0	8	0
CO II	2	0	0	0	2	0	2	0
ATP8	1	1	0	0	1	1	1	1
ATP6	4	0	0	0	4	0	4	0
CO III	5	2	2	0	4	2	4	2
ND3	1	0	0	0	1	0	1	0
ND4L	2	0	0	0	2	0	2	0
ND4	8	0	1	0	8	0	7	0
ND5	18	4	1	1	18	4	17	3
ND6	2	1	0	0	2	1	2	1
CytB	1	0	0	0	1	0	1	0
Total	61	10	8	1	58	10	56	9

3.2 Phylogenetic relationships inferred from mitochondrial genome

Based on 13 amino acid sequence of 17 mtDNA belonging to the family Phocidae, a NJ trees was constructed (Fig. 1). *Lipotes vexillifer* (NC_007629) and *Dugong dugon* (NC_003314) were selected to root the tree. The results show that the individual from Liaodong Gulf first clustered with that from South Korea, and then with the specimen from Alaska. The three spotted seal population then clustered with the other seals belonging to the subfamily Phocinae. The five species sequences of Monachinae form an independent clade well supported by high bootstrap values. The results of the phylogenetic tree based on the nucleotide and amino acid sequences were basically the same, and showed that group of Liaodong Gulf is more closely related to that of Korea. However, the spotted seal individuals from different location have a certain level of genetic differences.

3.3 Population genetic of mtDNA control region

A 430-bp segment of the 5' end of the control region was sequenced in 66 specimens, and used in the present study. Sixty-nine variable sites were checked within the 430-bp mitochondrial DNA control region fragments, and variations consisted predominantly of transition substitutions (46 transitions and 23 tranversions) (Fig.2). Three insertions/deletions were detected. These polymorphic sites defined 22 haplotypes among 66 individuals. Among 22 haplotypes, 7 of them were shared by different individuals, and the other 15 haplotypes were unique to one individuals. Hap 2 was shared by the most individuals (Table 1). The nucleotide base composition of the CR showed that the content of A+T (54.84%) was slightly higher than that of C+G (45.16%), and the content of G (16.48%) was the lowest showing strong base composition bias. The spotted seal population in the present study showed high haplotype diversity (0.8308±0.0398) and lower nucleotide diversity (0.0096±0.0054).

```
[                     11111 1111111111 1111112222 2222222222 2222223333 33344]
[            12333355 6789901123 3455667778 8899990000 1122233345 5668891122 33500]
[          4978025723 9900480581 6818891890 7827893567 2301624833 9162994605 79648]
#CHDL_66seq TAAGTCCTAG ATTCCTGTGC AACGAGAATC ACGTAGCACG GTGCTGAAAG GGAGGTACGG TTGGC
#CHDL_1seq  .......... .......... .......... ......T... .......... .......... .....
#CHDL_2seq  .......... .......... .......... .......... .......... .......... .....
#CHDL_3seq  .......... .C.T...... .......... .......... .......... .......... .....
#CHDL_4seq  .......... .......... .......... .......... .......... .......... .....
#CHDL_5seq  .......... .CCT...... .......... .......... .......... .......... .....
#CHDL_6seq  .......... .......... .......... ......T... .......... .......... .....
#CHDL_7seq  .......... .C........ .......... .......... .......... .......... .....
#CHDL_8seq  .......... .......... .......... ......T... .......... .......... .....
#CHDL_9seq  .......... .C........ .......... .......... .......... .......... .....
#CHDL_10seq .......... .......... .......... .......... .......G.. .......... .....
#CHDL_11seq .......... .......... .......... ...C..T... .......... .......... .....
#CHDL_12seq .......... .C........ .......... .......... .......... .......... .....
#CHDL_13seq .......... .C.T...... .....A.... .......... .......... .......... ...A.
#CHDL_14seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_15seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_16seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_17seq .......... .....C.... .....A.... ......T... .........T ........C. ...A.
#CHDL_18seq .......... .......... .....A.... ......T... .........A .A......C. ...A.
#CHDL_19seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_20seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_21seq .......... .......... .......... .......... .......... .......... .....
#CHDL_22seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_23seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_24seq .......... ...T...... .......... .......... .......... .......... .....
#CHDL_25seq .......... .C........ .....A.... .........A .......... .......... ...A.
#CHDL_26seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_27seq .......... .......... .......... .......... .......G.. .......... .....
#CHDL_28seq .......... .......... G......... .......... .......... .......... .....
#CHDL_29seq .......... .......... .......... .......... .......... .......... ..A..
#CHDL_30seq .......... .......... .......... .......... .......... .......... .....
#CHDL_31seq .......... .......... G......... .......... .......... .......... .....
#CHDL_32seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_33seq .......... .....CC.CG .C.A.A.... .....A.... A.A.....A CA..A...CC ..AA.
#CHDL_34seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_35seq .......... .......... .......... .......... .......... .......... .....
#CHDL_36seq .......... .......... .......... .......... .......... .......... .....
#CHDL_37seq .......... .......... .......... .......... .......... .......... .....
#CHDL_38seq .......... .......... .....C.... .........A .....A.... .......... .....
#CHDL_39seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_40seq .......... .......... G......... .......... .......... .......... .....
#CHDL_41seq .......... ...T...... .......... .......... .......... .......... .....
#CHDL_42seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_43seq .......... .......... .......... .......... .......... .......... ..A..
#CHDL_44seq .......... ...T...... .......... .......... .......... .......... ...A.
#CHDL_45seq .......... .......... .......... .......... .......... .......... .....
#CHDL_46seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_47seq .......... .......... ....G..... ......T..A .......... .......... .....
#CHDL_48seq ......TC.. .......... .......G.. GTA..A.GTA ....C..... .......... .....
#CHDL_49seq .......... .......... .......... .......... .......... .......... ..A..
#CHDL_50seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_51seq .......... .......... .......... .......... .......... .......... .....
#CHDL_52seq .......... .......... .......... .......... .......G.. .......... .....
#CHDL_53seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_54seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_55seq .......... ...T...... .......... .......... .......... .......... .....
#CHDL_56seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_57seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_58seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_59seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_60seq .C.A.T.... G...T..C.. ......GGCT .T.CG.T..A ......TG.. ..G...GT.. ....T
#CHDL_61seq .......... ...T...... .......... .......... .C........ .......... .....
#CHDL_62seq .......... .......... G......... .......... .......... .......... .....
#CHDL_63seq C.T.A...TT .......... G.A..A.... .......... ...A....T. .T.AAA..C. GGAA.
#CHDL_64seq .......... .......... .......... ......T... .......... .......... .....
#CHDL_65seq .......... .......... .......... ......T... .......... .......... .....
```

Fig.2 Variable sites for 66 individuals

In order to evaluate the genetic status of the spotted seal collected in this study, we compared the results with the sequences from GenBank. A total of 111 individuals were used in the further analysis. Ten haplotypes were shared among three populations, and 24 haplotypes were shared by three populations (Table 5). Population LD2005 showed highest haplotype and lowest nucleotide diversity (Table 1).

Another 9 sequences of Korean coastal waters were also downloaded from GenBank to construct phylogenetic tree. The results showed that there was a branch which mainly consisted of the haplotype from Russian coastal waters (Fig.3). However, there were no obvious pedigree branches and geographic structure on the whole.

The result of pairwise population F_{st} showed that genetic differentiation among the three populations ranged from 0.038 (LD2005 vs RU) to 0.069 (LD2019 vs RU, $P<0.01$). To further examine hierarchical population structure as well as the geographical pattern of population subdivision, we used analysis of molecular variance (AMOVA). The results of AMOVA analysis for one group showed that the genetic variation among populations was 5.77% and within populations was 94.23%. Two groups were defined to do further AMOVA analysis (Table 6). Population LD2019 and LD2005 were pooled to form one group, and Population RU was pooled to form another group. The results showed that the genetic variation among groups was 3.39%, among populations within groups was 3.42% and within populations was 93.19% (Table 3). The results of population genetic structure supported weak but significant differentiation among populations.

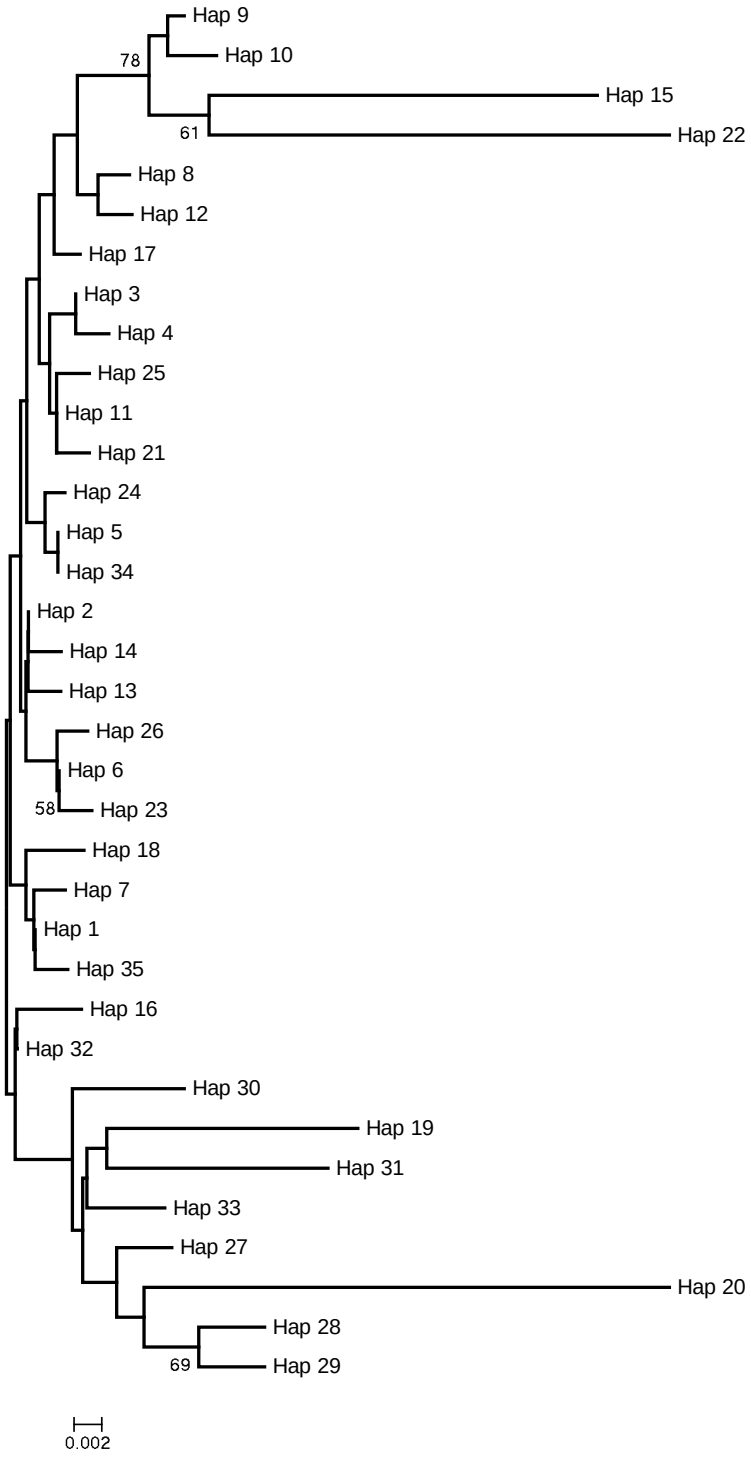

Fig.3 Neighbor-joining tree constructed using K-2P distances for 36 control region haplotypes of spotted seal. Bootstrap supports of >50% in 1000 replicates are shown.

Table 5 Haplotype distributions among three populations

Hap	LD2019	LD2005	RU	Hap	LD2019	LD2005	RU
Hap1	25	1	7	Hap18	1		
Hap2	10	3		Hap19	1		
Hap3	1	1	3	Hap20	1		
Hap4	1	1	2	Hap21	1		
Hap5	3	1		Hap22	1		
Hap6	3	1	3	Hap23		2	
Hap7	1			Hap24		2	1
Hap8	1			Hap25		1	
Hap9	1			Hap26		1	
Hap10	1			Hap27			3
Hap11	3	1	2	Hap28			1
Hap12	1			Hap29			1
Hap13	4	1		Hap30			1
Hap14	3	1		Hap31			1
Hap15	1			Hap32			1
Hap16	1			Hap33			1
Hap17	1			Hap34			1

Table 6 AMOVA of *Phoca largha* populations based on mtDNA control region sequences

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
One gene pool				
Among populations	2	11.951	0.12610Va	5.77
Within populations	108	22.563	2.06077 Vb	94.23
Total	110	234.514	2.18686	
Two gene pools				
Among groups	1	7.847	0.07492Va	3.39
Among populations within groups	1	4.104	0.07557Vb	3.42
Within populations	108	222.563	2.06077Vc	93.19
	110	234.514	2.21126	

The mismatch distribution of *Phoca larghais* appeared to be unimodal, and closely matched the expected distributions under the sudden-expansion model (Fig.4). Tajima's *D* and Fu's *Fs* statistics were significantly negative for these two clades (*D*=-2.24, *P*=0.00; *F*s=-16.60, *P*=0.00), which rejected the hypothesis of selective neutrality and were in accordance with the results of mismatch. Moreover, the significantly negative deviation of Fu's Fs value could be a sensitive genetic signature of a recent demographic expansion. The P_{SSD} and raggedness tests could not reject the expansion hypothesis. The mutational timescale $\tau=2\mu t$, which reflects the location of the mismatch distribution crest, can be used to estimate expansion time. The molecular clock for the control region seems to vary among major taxonomic groups of marine mammals, ranges from 2% to 10% per million years. Spotted seals with medium generation time and body size might have a medium molecular clock. In the present study, sequence divergence rate of 5%/Myr was

applied for the control region sequence of spotted seal. Based on this divergence rate, it was estimated that population expansion occurred approximately at 159, 000 years ago.

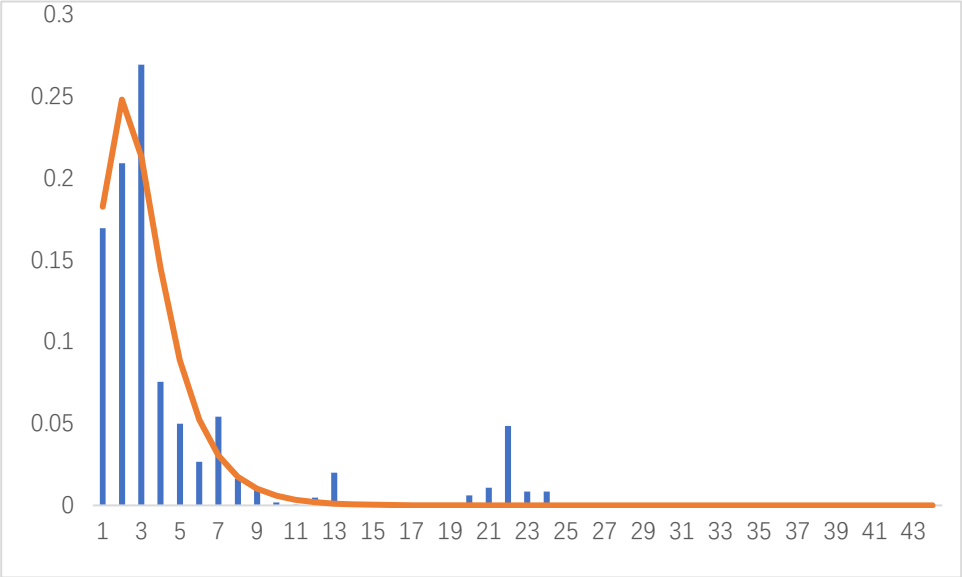

Fig.4 The observed pairwise difference (bars), and the expected mismatch distributions under the sudden expansion model (solid line) of control-region haplotypes for the spotted seal.

4 Conclusion

In this report, it can be seen that the sequence similarity between Liaodong and Korean is far greater than that between Alaskan and Liaodong, Alaskan and Korean, which indicates that the spotted seal in Liaodong Gulf of China and Korean in are probably a same population. The northern-western Pacific is the main distribution areas of spotted seals in the world (Rugh et al., 1997). The waters of Alaska, US are far away from the coasts of China and South Korea. From the satellite telemetry technique tracking spotted seals, it was found that the individuals in Liaodong Gulf can reach to the Baengnyeong Island of South Korea, which suggested that the spotted seals in China and

South Korea exchanged with each other (Han et al., 2013). Comparing to the Alaska group, the spotted seals distributed in Liaodong Gulf is closer to the South Korean group.

Compared with the samples in 2005, the genetic diversity of the spotted seal in 2019 is lower, and only 22 haplotypes were detected in 66 samples. The results of NJ tree analysis showed that the samples from China, Korea, and Russia coastal waters had no significant pedigree structure. However, the pairwise genetic divergences (F*st*) and AMOVA result show that there was small but significantly different between China population and Russia population, which suggested that the dispersal range of spotted was limited to the closed populations.

Duo to spotted seal is a long-distance migration species, distributed in Northwest Pacific including China, Korean, Japan, Russia and America. It has eight breeding ground along the coastline. It needs the cooperation of scientist from different countries to work together to research the population structure covering the whole distribution region. Two aspects of cooperative work can be carried out. On the one hand, different researchers can contribute their own samples, then we can study the genetic structure and diversity of the entire distribution region of spotted seals. Secondly, with the popular of sequencing technology, different researchers can carry out the sequencing work separately, and share the sequencing data with each other. Through this work we can obtain accurate population structure and phylogenetic relationship, which can promote to make effective protection and management measures for population resources.

References

Bahn S Y, Choi E H, Jang K H, Kim S K, Ryu S H, Hwang U W. Complete mitochondrial genome of the Korean spotted seal *Phoca largha* (Mammalia, Pinnipedia, Phocidae): genetic differences between *P. largha* and *Phoca vitulina*. *Mitochondrial DNA*, 2009, 20(5-6): 105–106.

Clary D O, Wolstenholme D R. Drosophila mitochondrial DNA: conserved sequences in the A+T–rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. *Journal of Molecular Evolution*, 1987, 25(2): 116–125.

Han J B, He C B, Wang Q, Ma Z Q, Xu X H. Sequence analysis of mitochondrial ND4, $tRNA^{Arg}$, ND4L and ND3 from spotted seal (*Phoca largha*) in Liaodong Gulf. *Fisheries Science*, 2006, 25(10): 500–504. (in Chinese with English abstract)

Han J B, He C B, Wang X M, Wang Q, Ma Z Q, Zhou Z C, Wang P L. Sequence analysis of mitochondrial $tRNA^{Thr}$, $tRNA^{Pro}$ and control region from spotted seals (*Phoca largha*) in Liaodong Gulf. *Fisheries Science*, 2007, 26(2): 74–78. (in Chinese with English abstract)

Han J B, Lu Z C, Tian J S, Ma Z Q, Wang Z H, Yang Y, Wang Q G, Song X R, Peng Z P. Release studies on spotted seals (*Phoca largha*) using satellite telemetry tracking technique. *Acta Theriologica Sinica*, 2013, 33(4): 300–307. (in Chinese with English abstract)

Lowe T M, Eddy S R. tRNAscan–SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 1997, 25(5): 955–964.

Miya M, Nishida M. The mitogenomic contributions to molecular phylogenetics and evolution of fishes: a 15–year retrospect. *Ichthyological Research*, 2015, 62(1): 29–71.

Mizuno A W, Onuma M, Takahashi M, Ohtaishi N. Population genetic structure of the spotted seal *Phoca largha* along the coast of Hokkaido, based on mitochondrial DNA sequences. *Zoological Science*, 2003, 20(6): 783–788.

Rugh D J, Shelden K E W, Withrow D E. Spotted seals, *Phoca largha*, in Alaska. *Marine Fisheries Review*, 1997, 59(1): 1–18.

Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning. New York: Cold Spring Harbor Laboratory Press, 1989.

Shaughnessy P D, Fay F H. A review of the taxonomy and nomenclature of North Pacific harbour seals. *Journal of Zoology*, 1977, 182(3): 385–419.

Staden R, Beal K F, Bonfield J K. The staden package, 1998. *Methods in Molecular Biology*, 2000, 132(6): 115–130.

Su Y. Conserved and heteroplasmy on mitochondrial DNA control region in animal. *Sichuan Journal of Zoology*, 2005, 24(4): 669–672. (in Chinese with English abstract)

Sun F Y. Studies on Population Genetic Characteristic of Spotted Seal (*Phoca largha*) in Liaodong Gulf. Dalian: Liaoning Normal University, 2009. (in Chinese with English abstract)

Tamura K, Dudley J, Nei M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 2007, 24(8): 1596–1599.

Thompson J D, Higgins D G, Gibson T J. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position–specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994. 22(1): 4673–4680.

Wang W, Zhu H, Hu H X, Tian Z H, Dong Y. Heteroplasmy in mtDNA control region and phylogenetics of five sturgeons. *Zoological Research*, 2009, 30(5): 487–496. (in Chinese with English abstract)

Won C, Yoo B H. Abundance, seasonal haul-out patterns and conservation of spotted seals *Phoca largha* along the coast of Bak-ryoung Island, South Korea. *Oryx*, 2004, 38(1): 109–112.

Xia X, Xie Z. DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity*, 2001, 92(4): 371–373.

Xing J J. The types of molecular markers and the research and applications of molecular markers technology on the aquatic creature. *Chinese Journal of Fisheries*, 2002, 16(1):61–70. (in Chinese with English abstract)

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 2003, 31(13): 3406–3415.