GENETIC VARIATION AND MOLECULAR EVOLUTION OF ENDANGERED KOZLOV’S PIKA (OCHOTONA KOSLOWI BÜCHNER) BASED ON MITOCHONDRIAL CYTOCHROME B GENE

ABSTRACT: Kozlov’s pika is a rare and endangered lagomorph species with a limited distribution in the southern Kunlun Mountains in western China. Because of its endangered status, Kozlov’s pika is considered a priority species for research and conservation action. Genetic variation and molecular evolution of the Kozlov’s pika were studied based on a total of 14 individuals from four locations along the eastern boundary between Xinjiang and Tibet province (35.20–36.48°N, 86.08–83.04°E) on extremely high elevation (usually over 4800 m a.s.l.). The density of local populations was about 3–4 per ha, living in a typical alpine desert grassland habitat. The complete mitochondrial cytochrome b (cytb) gene was amplified and sequenced. Based on the cytb gene sequences the genetic variation and molecular evolution were analyzed. Unexpected high haplotype diversity (0.956 ± 0.045) but low nucleotide diversity (0.00537 ± 0.00126) was found, indicating past demographic expansion. Significant partitioning of variance (P < 0.01) among populations (46.7%), and within populations (53.3%), indicating low level of genetic differentiations among local populations. Our results gave an optimistic survival status of Kozlov’s pika at the genetic level. Bayes Empirical Bayes analysis with model M2a and M8 detected three positively selected amino acid sites at the significance level of 0.05. The mutant types with either or both of the mutations aspartic acid to asparagine and glutamic acid to lysine had higher isoelectric point values. We suggested these mutant types might have biological significance to help individuals to adapt to the extremely high elevation habitats.

KEY WORDS: Kozlov’s pika, genetic diversity, genetic structure, molecular evolution

1. INTRODUCTION

Kozlov’s pika is a rare and endangered lagomorph species with a limited distribution in the southern Kunlun Mountain in western China (IUCN 2007). After initial description in 1884 it was not recorded by scientists for 100 years until it was rediscovered in 1984 (Zheng 1986). Kozlov’s pika is considered a priority species for research and conservation action because of its endangered status (Li et al. 2006). However, with the exception of the brief description of basic biology by Li et al. (2006), little is known about the survival status and ecology of Kozlov’s pika.

Molecular approaches have been developed to evaluate the survival status of a species at the genetic level. The mitochondrial cytochrome b (cytb) gene is one of the most commonly used gene markers. Genetic diversity and genetic structure are important indices that reflect a species’ genetic status. There is now a compelling evidence that loss
of genetic diversity can increase the extinction risk for a species (Frankham 2005). In addition, high level genetic structure between rudimental populations can cause further deficiency of total diversity due to the 'Wahlund effect' (Hedrick 2000). Therefore studies of pika’s genetic diversity and genetic structure are crucial for the conservation of this endangered species.

Moreover, Kozlov’s pika has the extremely high elevation habitats in Qinghai-Tibetan Plateau. Recent study on the plateau pika (Ochotona curzoniae Hodgson), which also lives at high elevation habitats, detected positively selected sites in the nuclear ob gene (Zhao et al. 2008). They have been interpreted as adaptations to extremely high elevation. The mitochondrion is the ‘energy factory’ of eukaryote cell, and the product of cyt b gene constitutes the cytochrome bc1 complex which plays an important role in mitochondrial respiratory chain (Iwata et al. 1998). We suggested that the study on molecular evolution of cyt b gene in Kozlov’s pika may help in understanding evolution of adaptations to high elevation habitats.

The aim of this paper is to provide estimates of genetic variation in the pika and to discuss selection of molecular adaptations to life at high elevation in mammals.

2. MATERIALS AND METHODS

In October and November 2007, we investigated the distribution of the Kozlov’s pika in the Kunlun Mountains along the eastern boundary between Xinjiang and Tibet province during an expedition organized by the Chinese Academy of Sciences. The study area ranged from 35.20–36.48°N and 86.08–83.04°E and the average elevation was over 4800 m a.s.l. (Fig.1). The pika lived in a typical alpine desert grassland habitat, with a small number of plant species including Carex moorcroftii, Ceratoides compacta, as well as some Leguminosae and Cruciferae species. The density of local populations was about 3–4 per ha. The body weight of adults was about 300 g and each had 7–8 shallow burrows with a depth of only 30–40 cm. With the permission of local government, we sampled some individuals to analyze the genetic variation and molecular evolution of this species based on the mitochondrial cyt b gene.

Pikas were trapped on 1–2 ha plots and each plot was sampled in the day time with

![Fig. 1. Study region in the west of China. P1~P4, the four local populations in this study; M1, Arjin Mountains; M2, Kunlun Mountains; M3, Kekexili Mountains.](image-url)
15 traps randomly placed at burrows. Pikas were decollated immediately after trapping and muscle tissue samples for DNA extraction were collected and preserved in 95% alcohol. The longitude, latitude and elevation at the centre of each sampling plot were measured using a GPS unit (Garmin Etrex, Taiwan). Due to the limited time available and the low activity of pikas only a small number of individuals were trapped (a total of 13 pikas were trapped from four populations (P1-P4) (Fig.1).

Total genomic DNA was extracted from approximately 0.3 g muscle following a modification of the Sambrook and Russell method (Sambrook and Russell 2001). Cytb gene was amplified using the primers: L 1 4 7 2 4 (5’- CGA AGC TTG ATAT- GAAACATCGTTG-3’) and H15915 (5’- CGGAATTCCATTTTG- GTTTACAAGAC-3’) with the reference of Kocher et al. (1989). Thermocycling was conducted in a Biometra T-Gradient Thermoblock PCR machine (Biometra, Gottingen, Germany). The PCR products were purified using a CASpure PCR Purification Kit (Casarray, Shanghai, China), and directly sequenced using the same primers used for amplification mentioned above. Sequencing reactions were conducted in a Biometra thermocycler using a DYEnamic Dye Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech Inc., USA.) following the manufacturer’s protocol. Sequencing products were separated and analyzed on a MegaBACE 1000 DNA Analysis System (Amersham Pharmacia Biotech Inc., USA.). The sequences were checked by eye and aligned using CLUSTAL W (Thompson et al. 1997) and refined manually. Due to the small sample size from the Kunlun Mountains, the sequence from a single individual collected in 2006 near Jinyuhu Lake was also used (P4, see Fig 1). In total, 14 sequences from four populations were used in this study.

Genetic diversity, both haplotype diversity (Hd) and nucleotide Diversity (Pi) were calculated with Arlequin 3.0 software based on Kimura 2P algorithm (Excoffier et al. 2005). There was a high haplotype diversity but a low nucleotide diversity level (see Chapter 3 – Results and Discussion), indicating a past population expansion. In order to test the hypothesis on past population growth, we used Fu’s Fs neutral test (Arlequin 3.0. software), which is one of the most powerful tests for detecting population growth (Ramos-Onsins and Rozas 2002).

Positive selection sites were detected using codeml program in PAML (Phylogenetic Analysis by Maximum Likelihood) program package version 4.1 (Yang 2007). The tree was an unrooted Neighbor Joining tree generated using MEGA 4 (Molecular Evolutionary Genetic Analysis; Kumar et al. 2008) based on the haplotypes. Following the suggestions in User Guide of PAML program package, the M2a (selection) and M8 (beta&ω) models with the Bayes Empirical Bayes algorithm were used. The methods use nonsynonymous/synonymous substitution rate ratio (dN/dS, or synonymically ω) as a measure of selective pressure at the protein level, with ω > 1 indicating positive selection.

3. RESULTS AND DISCUSSION

3.1. Genetic variation

A total of 24 variable sites and 11 haplotypes were identified (Table 1). Unexpected high haplotype diversity (Hd ± SD = 0.956 ± 0.045) was found, while the nucleotide diversity is low (Pi ± SD = 0.00537 ± 0.00126). The Fu’s Fs neutral test showed a nearly significant (P = 0.059) negative Fs value (–3.055), indicating that there was at least once past population expansion events (Excoffier et al. 2005). Standard AMOVA analysis revealed significant partitioning of variance (P < 0.01) among populations (46.7%), and within populations (53.3%), indicating a low level of genetic differentiations among local populations.

It is generally thought that a rare species will exhibit low genetic diversity, high interpopulation genetic differentiations and decreasing population size (Jacquemyn et al. 2007). Our results to some degree contradict this common view, although we must be cau-5
tious given the small sample sizes obtained during our expedition. We found evidence of high haplotype diversity and low nucleotide diversity which may represent a signal of rapid demographic expansion from a small effective population size (Avise 2000). The negative value in Fu’s Fs test also supports this hypothesis. Moreover, field surveys in 1999 (Li et al. 2006) showed that density of Kozlov’s pika in suitable habitats can be high, to some degree supporting our view. Based on our survey, the distribution of Kozlov’s pika is continuous in the study region and there is no obvious fragmentation between the four local populations. We suggest that the low level inter-population genetic differentiations have resulted from the habitat connectivity among local populations.

Our results gave an optimistic prognosis for the survival of Kozlov’s pika. However, this does not imply that there are no current conservation concerns. The distribution of Kozlov’s pika is so limited that previous expeditions to the Kunlun Mountains in 1989 and 1992 were unable to find the species (Li et al. 2006). Moreover, the distribution of Kozlov’s pika appears to be at risk of invasion by the plateau pika which is the dominant and keystone species on the Qinghai-Tibetan Plateau (Smith and Foggin 1999). Li et al. (2006) suggest that the plateau pika competes with Kozlov’s pika and plateau pika individuals were also found in several Kozlov’s pika populations during our own survey. However, whether and how Kozlov’s pika is affected by plateau pika remains unknown and requires further investigation.

Table 1. The geographic records and haplotype constitution in local populations (P1-P4) (see Fig. 1) and the isoelectric point as well as variable site information of each haplotype (H1 – H11). The dot in each nucleotide position of relevant haplotype means the same as in H1.

<table>
<thead>
<tr>
<th>Population, longitude (E °), latitude (N °), elevation (m)</th>
<th>Haplotype (frequency)</th>
<th>Isoelectric point</th>
<th>Variable nucleotide position</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1, 86.0761, 36.4954, 4884</td>
<td>H1</td>
<td>7.94</td>
<td>ACGCGCAAGGCCCCTTCTCCGA</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>7.94</td>
<td>. . A. . . . . . . . . . . . TC . CT . . . .</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>7.93</td>
<td>. . A. . . . . . . . . . . GG . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>7.94</td>
<td>. . A. . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H5</td>
<td>8.54</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H6</td>
<td>8.76</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td>P2, 86.1084, 35.9393, 4996</td>
<td>H2(2)</td>
<td>7.94</td>
<td>. . A. . . . . . . . . . . TC . CT . . . .</td>
</tr>
<tr>
<td></td>
<td>H7</td>
<td>7.94</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H8</td>
<td>7.94</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td>P3, 83.0442, 35.2043, 4907</td>
<td>H7</td>
<td>7.94</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H9</td>
<td>8.26</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td></td>
<td>H10</td>
<td>8.26</td>
<td>. . . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td>P4, 89.2520, 36.4531, 4719</td>
<td>H11</td>
<td>8.76</td>
<td>GCT. TAATG. AA . . T . . C. TAT</td>
</tr>
</tbody>
</table>

Gonghua Lin et al.
3.2. Molecular evolution

The Bayes Empirical Bayes analysis with model M2a and M8 detected 3 positively selected amino acid sites at the significance level of 0.05 (Table 2). One (111 E&K) locates in transmembrane region (next to the mitochondria matrix side) and two (58 D&N, 171 D&N) locate in the mitochondria matrix side, according to Irwin et al. (1991) (The capital letters D, E, K and N in the context represents aspartic acid, glutamic acid, lysine and asparagine, respectively).

Previous study (Irwin et al. 1991) on evolution of the cytb gene in mammals found that these three sites were very conservative, i.e. most of the sequences used in their study had D, E and D in No. 58, 111 and 171 Aa site, respectively. Interestingly however, these sites in some of our haplotypes (mutant types) mutated to other Aa types with significantly higher isoelectric point values (D to N, 2.77 to 5.41; E to K, 3.22 to 9.74) than the original products detected prior to mutation (Haig and Hurst 1991). Using the program EditSeq in DNAStar program package (version 6.0; Dr. Steve ShearDown, Madison, USA), we calculated the rough isoelectric point values of the products of these haplotypes. The results indicated that either of the mutations or both mutations (mutation from D to N and E to K) in their Aa sequences might help individuals to adapt to the extremely high elevation habitats.

ACKNOWLEDGEMENTS: This study was supported by the Training Qualified People Plan “Hope of the West China” of The Chinese Academy of Sciences and the Ministry of Personnel of China (No. O95402121) and General Programs of the National Natural Science Foundation of China (No. 30970366). Simon J. Thirgood was funded by the International Exchange Programme of the Royal Society of Edinburgh.

4. REFERENCES


Table 2. Positively selected amino acid (Aa) sites (refer to conserved & mutated Aa type according to Irwin et al. (1991) detected by Bayes Empirical Bayes analysis). ω, the ratio of nonsynonymous/synonymous substitution rates; Pr (ω >1), probability of the case that ω >1; M2a, the selection model and M8, the beta&ω model according to Yang (2007); D, E, K and N represents aspartic acid, glutamic acid, lysine and asparagine, respectively; *, P >95%; **, P >99%.

<table>
<thead>
<tr>
<th>Aa site</th>
<th>Pr (ω &gt;1)</th>
<th>ω (mean ±SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M2a</td>
<td>M8</td>
</tr>
<tr>
<td>58 D&amp;K</td>
<td>0.995**</td>
<td>0.998**</td>
</tr>
<tr>
<td>111 E&amp;K</td>
<td>0.959*</td>
<td>0.974*</td>
</tr>
<tr>
<td>171 D&amp;N</td>
<td>0.969*</td>
<td>0.980*</td>
</tr>
</tbody>
</table>
Gonghua Lin et al.

Received after revision September 2009