International Journal of Biological Macromolecules 121 (2019) 223–232

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: http://www.elsevier.com/locate/ijbiomac

The complete mitochondrial genome of *Budorcas taxicolor tibetana* (Artiodactyla: Bovidae) and comparison with other Caprinae species: Insight into the phylogeny of the genus *Budorcas*

Ming Zhou, Jiaojiao Yu, Biao Li, Bo Ouyang, Jiandong Yang *

College of Animal Sciences and Technology, Sichuan Agricultural University, Chengdu 611130, China

ARTICLE INFO

Article history:
Received 21 May 2018
Received in revised form 3 October 2018
Accepted 5 October 2018
Available online 06 October 2018

Keywords:
Budorcas taxicolor tibetana
Mitogenome
Molecular phylogeny

ABSTRACT

Takin (*Budorcas taxicolor*) is one of the most endangered species. However, the taxonomy of takin is still in dispute. Here, the complete mitochondrial genome of *Budorcas taxicolor tibetana* was reported for the first time, which featured a typical circular molecule of 16,665 bp in length, including 13 PCGs, 22 tRNAs, two rRNAs, and one control region. A + T content was higher than G + C content. All of the genes were encoded on the heavy strand, except for eight tRNAs and *ND6* gene. The O_L region was 49 bp in length and highly conserved in the synthesis and stem-loop regions, and all of the substitutions and indels were found only in the loop structure. Three types of tandem repeat units, six pairs of hairpin loop structure (TACAT, ATGTA) and six CSBs were predicted in the control region. Our results clearly revealed the systematic status of *Budorcas* species, and the phylogenetic analyses indicated that *Budorcas* was closer to *Capra* and *Pseudois*, rather than to *Ovis*, which should be merged into the subfamily of Caprinae.

1. Introduction

In the last decade, rapid advances in molecular biology technology have promoted the accuracy of phylogenetic studies. Specifically, the mitochondrial DNA (mtDNA) exhibits maternal inheritance, low rates of recombination and has no intron, and this has resulted in its wide application in the field of molecular ecology, population genetics, taxonomy, phylogeography, and phylogenetic studies [1–5]. Previous studies have suggested that long DNA or multigene concatenated sequences provide richer phylogenetic information than short or single genes, which might reveal phylogenetic relationships and systematic-taxonomic status more accurately [6–8]. In recent years, researchers tend to apply all protein-coding genes or complete genome of mitochondria as molecular marker to phylogenetic analyses [9–11].

Budorcas taxicolor (Artiodactyla: Bovidae), also known as takin, lives in high elevation regions from Eastern Himalayas to South-Central China. This species is one of the most endangered species, classified as vulnerable (VU) under the International Union for Conservation of Nature Criteria (IUCN) [12]. The four reported subspecies are *Budorcas taxicolor bedfordi*, *B. taxicolor taxicolor*, *B. taxicolor tibetana* and *B. taxicolor whitei* [13]. However, the taxonomy of takin is still in dispute, particularly at the level of subfamily. The main reason is that the takin is of a larger size similar to the Bovinae species, but the external shape resembles that of the Caprinae species. In view of the external morphological characteristics, previous studies have attributed it to the Rupicaprinae, Antelepinae, Caprinae, but some researchers have independently classified it with the Musk Ox (*Ovibos moschatus*) as the Ovibovinae species [14–16]. An additional reason is that the comparative morphology of Artiodactyla species is not always accurate for classification because of morphological and behavioral similarities between takin and other Artiodactyla species, particularly Musk Ox [17,18]. In recent years, researches on the takin have also expanded to the molecular level. However, these studies have only been based on the usage of a single gene fragment (such as *Cytb* or *COI*) or partial control region sequence [15,18–20].

In this study, to further confirm the systematic relationship and taxonomic status, and to understand the mitochondrial genome structure of *B. t. tibetana*, we for the first time reported the complete mitochondrial genome sequence of *B. t. tibetana*, and compared its mitochondrial genome features and architecture with the previously reported Caprinae mitochondrial genomes. Furthermore, to reconstruct a more reliable and robust phylogeny, we applied the Maximum Likelihood (ML) and Bayesian Inference (BI) approaches using the concatenated

* Corresponding author.
E-mail address: yangjd@sicau.edu.cn (J. Yang).

https://doi.org/10.1016/j.ijbiomac.2018.10.020
0141-8130/

nucleotide sequences of the 13 protein coding genes (PCGs) and two ribosomal RNA genes (rRNAs) of 44 Bovidae species (34 Caprinae species, 10 Bovinae species).

2. Material and methods

2.1. Ethics statement

All sample collection and research processes conformed to the ethical guidelines, religious beliefs, and laws and regulations of the People's Republic of China (PRC). All animal experiments were conducted in accordance with the protocols in the Guide for the Care and Use of Laboratory Animals of the Ethics Committee for Laboratory Animal Research in China, and were approved by the Committee of Laboratory Animal Experimentation at the Sichuan Agricultural University.

2.2. Sample collection and DNA extraction

The takin muscle was stored in the Wildlife Conservation Laboratory of Sichuan Agricultural University in 95% ethanol. Total genomic DNA was extracted according to traditional phenol-chloroform protocols [21]. The quality and concentration of DNA was assessed by NanoDrop 2000 Ultra-Vis spectrophotometer (Thermo Scientific, USA), and then transferred to −20 °C for long-term storage.

2.3. Primer design, PCR amplification and DNA sequencing

We designed 18 specific primer pairs using the mitochondrial genome of *B. t. bedfordi* [20] and *B. t.* sp. [22] and synthesized them at TSINGKE Biology Co., Ltd. (Chengdu, China) (Table 1). The fragments were amplified using a 25 μL reaction cocktail that contained 1.5 μL template genomic DNA (120 ng/μL), 0.5 μL of each primer (10 μmol/L) and 22.5 μL TSINGKE Multiplex PCR Master Mix (green) (TSINGKE, Beijing, China). The PCR profile was 94 °C/5 min, followed by 35 cycles at 94 °C - 30 s, 55 °C - 30 s, 72 °C - 35 s, then 72 °C/8 min. PCR products were electrophoresed on a 1.2% agarose gel and then directly sequenced using the ABI 3730 DNA sequencer (TSINGKE Biology (Chengdu) Co., Ltd., China).

2.4. Mitochondrial genome analysis and gene identification

To ensure the accuracy of the sequences, we edited and assembled the DNA sequences manually using SeqMan, aligned by MEGA v6.0 [23], then used the BLAST searches on NCBI web site (http://blast.ncbi.nlm.nih.gov/Blast) to confirm the PCGs and rRNA genes. Locations of PCGs and rRNA genes were annotated according to the mitochondrial genomes of *B. t. bedfordi* (KU361169) and *B. t.* sp. (FJ207524). ORF Finder was used to find the initiation and termination codons of the PCGs. The locations and secondary structures of transfer RNA genes (tRNAs) were predicted by tRNAscan-SE server (http://lowelab.ucsc.edu/tRNAscan-SE/) [24] under the default search mode, using the vertebrate mitochondrial genetic code. Non-confirmed tRNAs were compared with the nucleotide sequences of *B. t.* sp. and *B. t. bedfordi* [25], then annotated by proposed secondary structures and identification of the anticodons. Intergenic spacers and overlapping regions were also counted manually.

Composition skews were calculated according to the following formulas: AT skew = [A − T] / [A + T] and GC skew = [G − C] / [G + C] [26]. The basic nucleotide compositions and codon usage statistics were estimated using MEGA v6.0. The control region (D-loop) and its conserved motifs were identified by comparing with the reference sequences. Repeated regions in the D-loop were detected by Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html) [27]. In most vertebrates, the stem-loop structure of the putative origin of light-strand replication (O_L) begins with 5′ - CTTCT - 3′. Therefore, we identified the conserved element located in the O_L sequence, compared it with the homologous sequences of other takins, and then used the Mfold v.3.2 program (http://mfold.bioinfo.rpi.edu/) under the default settings to find the O_L sequence and derive its secondary structure [28]. The variable sites of O_L sequences were identified using MEGA v6.0. Finally, the mitochondrial genome organization of *B. t. tibetana* was visualized by GenomeVx (http://wolfe.ucd.ie/genomevx/) and embellished manually [29].

2.5. Phylogenetic analysis

We downloaded 44 mitochondrial genome sequences from GenBank database, including 33 Caprinae, 10 Bovinae and one outgroup species (*Sus scrofa*, NC_012095) to investigate the phylogenetic position of the genus *Budorcas* (Table S1). Firstly, initiation and termination codons were manually deleted from all PCGs. Then, the concatenated sequences of the PCGs and two rRNAs were aligned using ClustalW with default parameters [30], ambiguous alignment regions were trimmed by the Gblocks Server (http://molevol.cmima.csic.es/castresana/Gblocks_seever.html) [31], and the type of sequence was set to codons for PCGs with all options for a less stringent selection. PartitionFinder 2 [32] was used to select partitioning schemes and substitution models for matrix. The search models for DNA sequences were set to “mrbayes”

Table 1
Primers used in amplification of *B. t. tibetana* mitogenome.

Primer	Sequence (5′–3′)		Annealing temperature
F1	ATCCTTGCTAATACAGTCT	CATAGGGTCTTCTCGTCTT	47.3 °C
F2	GCAGTTTTGGTTGGGGTGAC	TAATGGGAATGATGGCGAGT	53.6 °C
F3	GAATCCCAGAAGTAACAC	TCAGTACAGATCACACGA	50.5 °C
F4	CCTCCTCTCACTTCCTGT	GGTGTGGGTTAGTTTTGT	55.2 °C
F5	AACCCATCTCTCACAGTAA	GTGTCATTATGTGTTGTCG	48.4 °C
F6	ACCAATGATGACGAGATGT	TTGATAGCCAGGTTAGGGG	48.4 °C
F7	CTCATTATTATTACCCGAT	GGTCAGTCAGTGTCAGTTT	55.2 °C
F8	AAATAAGAGAGGAAGGAA	TACAGGGGTAATGAAAGA	52 °C
F9	ATAGGCGAAGGTTTTGAAG	TCGCTTCTTCCTTGAGTCT	50.5 °C
F10	ATAGCCTCACCATCAACAC	TGCTCCTCTTAGTCCTGC	57 °C
F11	TCCCATCCCTCACACTAACT	ATGAATCCTGCGAAAAGAC	50.5 °C
F12	TTTTCTACCCCTAACGCTA	TTGAAGGCTCTTGGTCTAT	50.5 °C
F13	TAGTTACAGGCTTCCGCA	GGCTTTGGGTAGTCAGAG	49.9 °C
F14	ACGCAGGAGCCTCAGTAGA	TAGCGGATGTAAAGTAAGC	49.4 °C
F15	TTATCTTCTTCACACTCC	ATGCTATTGTTGTATCGG	48.2 °C
F16	TTAGTCCTTTCCATCTTGA	GAGGTTCTTAGTGTTTTTG	49.4 °C
F17	ATGGGCACTTCAGTCAATG	GTCGTAAACCCTATTGTCG	49.9 °C
F18	TGGGTCTGTTTACATACTA	TGGTTCAACTAAGCACTCT	48.2 °C