Biochemical Systematics and Ecology 69 (2016) 226–235

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

journal homepage: www.elsevier.com/locate/biochemsyseco

Characterization of the complete mitochondrial genome of *Phodopus roborovskii* (Rodentia: Cricetidae) and systematic implications for Cricetinae phylogenetics

Li Ding, Wenjia Li, Jicheng Liao*

School of Life Sciences, Lanzhou University, 222 Tianshuinan Road, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form 10 October 2016
Accepted 15 October 2016

Keywords:
Phodopus roborovskii
Cricetinae
Complete mitochondrial genome
Molecular phylogeny

ABSTRACT

Phodopus roborovskii (subfamily Cricetinae) is widely distributed in the northern arid regions of China. This study reports its complete mitochondrial genome (mitogenome) for the first time. The complete sequence was 16,273 bp long, including 13 protein-coding genes, 2 ribosomal RNAs, 22 transfer RNAs, and 1 major noncoding region. The base composition and codon usage were described. The putative origin of replication for the light strand (O_L) of *P. roborovskii* was approximately 45 bp long and was highly conserved in the stem-loop and adjacent sequences, but the starting sequence of replication varied between genera among Rodentia. We analyzed the three domains of the D-loop region, and the results indicated that the central domain had higher G + C content and lower A + T content than two peripheral domains. Phylogenetic analyses indicated high resolution in four main divergent clades using mitogenomes data within Cricetidae. Within Cricetinae clade, *P. roborovskii* was at basal position which was in line with previous researches, and it shared a common ancestor with other extant hamsters. This work validated previous molecular and karyotype researches using mitogenomes data, and provided a set of useful data on phylogeny and molecular evolution in Cricetidae species.

1. Introduction

The subfamily Cricetinae comprises small-medium muroids with a wide distribution in the Holarctic region (Luo et al., 2000). These rodents are adaptable and abundant in a variety of habitats, including grasslands, meadows, scrubland, broad-leaved forest, mountains, hills and plains. Based on morphological characteristics and color variation in China, extant hamsters consist of nine recognized species across three genera (Wang and Zheng, 1973). However, current Cricetinae systematics is unclear: the actual number of species, as well as the phylogenetic relationships between genera, have been highly disputed (Corbet, 1978; Neumann et al., 2006). Recent developments with molecular markers have increased the accuracy of phylogenetic studies in investigating evolutionary histories between organisms. Among molecular markers, mitochondrial DNA (mtDNA) is the most efficient and sensitive; it has been widely applied to molecular ecology, population genetics and phylogeography (Zardoya and Meyer, 1996; Fan et al., 2012; Lu et al., 2015). Specifically, animal mtDNA exhibits maternal

* Corresponding author.
E-mail addresses: dingl14@lzu.edu.cn (L. Ding), 1132905461@qq.com (W. Li), liaojch@lzu.edu.cn (J. Liao).

http://dx.doi.org/10.1016/j.bse.2016.10.010
0305-1978/

inheritance, relatively low rates of recombination, and high substitution rates; these advantages result in its frequent application as a genetic marker for studying animal evolution, population genetics, and phylogenetic relationships (Hutchison et al., 1974; Harrison, 1989; Wang et al., 2013). However, multigenes markers and genome-level may be sufficient for resolving phylogenetic relationships because they provide amount of phylogenetic information (Boore, 2006; Tu et al., 2014). To resolve phylogenetic relationships in Cricetinae, we need sufficient molecular data obtained via mtDNA markers to validate previous taxonomic studies.

Phodopus roborovskii (desert hamster) belongs to *phodopus* with a wide distribution in the northern and northwestern arid regions of China. Its distribution ranges from northeastern China (western Liaoning and Jilin Provinces), through a large portion of northern China and Mongolia, to northwestern China and Kazakhstan (Ross, 1994). The smallest of all hamsters, *P. roborovskii*, can endure extreme desert conditions; they preferentially inhabit along sand dunes and avoid solid clay substrates and areas covered with dense vegetation (Flint, 1966; Newkirk et al., 1998). Interestingly, *Phodopus* are the only known hamsters that exhibit high social tolerance and biparental behavior (Wynne-Edwards and Lisk, 1987). In addition, it is the most primitive species within *Phodopus* compared with *Phodopus campbelli* and *Phodopus sungorus* in karyotype characters (Schmid et al., 1986). To date, the complete mitochondrial genome (mitogenome) sequence of *P. roborovskii* remains unavailable, thus hampering the development of well-supported phylogenetic inferences based on full mitogenomes analysis.

In the present study, we report on the first mitogenome sequence of *P. roborovskii* and compare it with sequences from other hamster species. A phylogenetic analysis was performed to clarify the evolutionary relationships within Cricetidae, combining the results of this study with previously published data in this family.

2. Materials and methods

2.1. Sampling and ethics statement

A voucher specimen of *P. roborovskii* was collected from Akesai county (N 38°53′57.13″, E 93°54′16.96″, 2848 m above sea level) in Gansu Province, China; it was identified based on external characteristics morphology (including of the skull), molecular data (Luo et al., 2000; Robins et al., 2007), and database searches (e.g., homologous gene identification using the basic local alignment search tool (BLAST) from NCBI) (Altschul et al., 1990). Muscle tissue of *P. roborovskii* was collected from the individual; then both tissue and voucher specimen was preserved in absolute ethanol and formalin solutions, respectively. The sample collections and the study did not conflict with the ethical guidelines, religious beliefs, or legal requirements of China. All animal experiments conformed to the guidelines of care and use of laboratory animals and were approved by the Committee of Laboratory Animal Experimentation at Lanzhou University.

2.2. DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from the muscle sample using the TIANamp Genomic DNA Kit (Tiangen, Beijing, China). Mitochondrial DNA sequence fragments were amplified with PCR; four primers were used for partially conserved sequences (Arèvalo et al., 1994; Hao et al., 2011), and the remaining long fragments were sequenced with primers designed specifically (Table S1) for long and accurate PCR, with each mtDNA sequence exhibiting a > 100-bp overlap with the next contig (Zhang et al., 2003b). The reaction volume was 50 μl: 10 μl 5 × PCR buffer (Mg^{2+} plus), 4 μl dNTP (2.5 mM each), 1.5 μl of each primer (10 μM each), 1 μl Primer STAR GXL DNA polymerase (1.25U/μl, Takara, Dalian, China), 2 μl total genomic DNA (100–200 ng) as template, and 30 μl ddH_2O. The PCR cycle was as follows: pre-denatureing for 3 min at 98 °C, followed by 35 cycles of denatureing for 10 s at 98 °C, annealing for 15 s at 50–60 °C, extension for 1 kb/min (kilobase) at 72 °C, and a final extension at 72 °C for 8 min. All PCR products were detected using 1% agarose gel electrophoresis and then subjected to Sanger sequencing (Genewiz Biotech Co., Ltd., Suzhou, China).

2.3. Mitogenome sequence analysis and gene identification

DNA sequences were edited and assembled manually in Lasergene version 5.0 (DNASTAR) and aligned with BioEdit version 7.0.5.2 (Hall, 1999). Contig assembly was performed with Seqman software (Lasergene subprogram). Nucleotide base compositions and codon usage of protein-coding genes (PCGs) were calculated using MEGA 6.0 (Tamura et al., 2013). The locations of PCGs, tRNAs, and rRNA genes were annotated via BLAST alignment to search for homologous sequences, which were then used as reference sequences. Specifically, conserved *C. griseus* and *M. auratus* mtDNA genes sequences were used in BioEdit and BLAST searches for different genes. Most tRNA genes were validated using tRNA scan-SE1.2.1 (Lowe and Eddy, 1997) with the default search mode, using the vertebrate mitochondrial genetic code source. We used the same method to compare the reference sequence with homologous DNA sequences for determining the control region using BioEdit (Sbisà et al., 1997). The circular mitochondrial genome map of *P. roborovskii* was drawn by the Mtv online tool (http://pacosy.informatik.uni-leipzig.de/mtviz/mtviz). The putative origin of L-strand replication (O_L) was located in the WANCY (amino acid abbreviation for a cluster of five transfer RNAs) cluster region between the ND2 and COX1 genes. To rapidly obtain the O_L sequence, we found the easy identified, conserved stem-loop structure of O_L (beginning with 5′-CTTCT-3′ in most vertebrates); other regions in the replication site were not suitable because the starting sequence of replication varies between genera and species (Zardoya et al., 1995; Zhang et al., 2003b; Jiang et al., 2012). Therefore, we identified the conserved motif

(5′-CTTCT-3′) to find the O_L sequence, and then using the mfold web server online (http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form) to draw the O_L secondary structure map of representative Rodentia species (Zuker, 2003).

2.4. Phylogenetic analysis

The complete *P. roborovskii* mitogenome and other available Cricetidae species mitogenomes downloaded from GenBank (Table 1) were included to estimate phylogenetic relationships. Multiple alignments and polymorphic site analysis of the mitogenome sequences were performed in BioEdit version 7.0.5.2 and DnaSP v5 (Librado Sanz and Rozas Liras, 2009), respectively. Phylogenetic analyses of the mitogenome were run in PAUP 4.0b10 (Swofford, 2003) and MrBayes 3.2.6 (Ronquist et al., 2012) using Bayesian inference (BI) (Yang and Rannala, 1997), maximum likelihood (ML) (Felsenstein, 1981), maximum parsimony (MP) (Fitch, 1971), and distance-based neighbor-joining (NJ) analyses (Saitou and Nei, 1987) methods. We employed the Akaike Information Criterion (AIC) (Akaike, 1974) in MrModeltest 2.3 (Nylander, 2004) to find the best-fit model of evolution. According to AIC results, the General Time Reversible (GTR) model was used for the BI, NJ and ML analyses. ML and NJ employed the same maximum likelihood model, whereas MP analyses were performed using a heuristic search and a tree-bisection-and-reconnection (TBR) branch swapping option. The robustness of inferences was assessed with bootstrapping (BT) (1000 random repetitions for MP and NJ, 100 for ML) (Felsenstein, 1985). In all analyses, only clades with a BT value equal or above 70% were considered as strong relationships (Hillis and Bull, 1993).

The AIC best-fit model for BI analysis (conducted in MrBayes 3.2.6) was computed in MrModeltest 2.3 with one million Markov Chain Monte Carlo (MCMC) simulations, run with default model parameters starting from a random tree, and sampled every 100 generations. The first 25% were discarded as a conservative burn-in and the remaining samples were used to generate a 50% majority rule consensus tree. Based on previous studies (Churakov et al., 2010), the jerboa *Dipus sagitta* (NC_027499.1) and *Jaculus jaculus* (NC_005314.1) were used as the most obvious outgroup. All tree files were drawn in Figtree v. 1.3.1 program.

3. Results and discussion

3.1. Mitochondrial genome characterization

The complete mitochondrial genome sequence of *P. roborovskii* was 16,273 bp long, longer than that of *M. auratus* (16,264 bp), but shorter than that of *C. griseus* (16,284 bp). The *P. roborovskii* mitogenome was similar to that of other mammals in terms of gene quantity and organizational structure, containing 13 protein-coding genes (PCGs), 2 rRNA genes (12S rRAN and 16S rRNA), 22 tRNA genes, and 1 control region (Bibb et al., 1981; Partridge et al., 2007; Li et al., 2009) (Fig. 1). The heavy strand (H-strand) base composition was 33.9% A, 13.3% G, 25.3% C and 27.4% T (Table 2). The most common base was A (33.9%), and the least common base was G (13.3%). Furthermore, the base composition of *P. roborovskii*'s mitogenome was AT-rich (61.3% A + T content, 38.6% G + C content; Table 2). These characteristics are similar to other rodent mitogenomes (Partridge et al., 2007; Jiang et al., 2012), such as *C. griseus* (NC_007936.1), and *Microtus fortis calamorum* (JF261175). As shown in Fig. 1, the 22 tRNA genes were interspersed among rRNAs and PCGs, with few gaps between them, and some genes even overlapped with each other. The longest overlap of 43 nucleotides was found between ATPase 8 and 6 (Table 3). All genes were encoded on the H-strand except ND6 and eight tRNA genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser1}$, $tRNA^{Glu}$, and $tRNA^{Pro}$), which were encoded on the light strand (L-strand) (Table 3).

Table 1
Cricetidae species used in phylogenetic analyses and their mitogenome accession numbers.

Species	Common name	GenBank accession number
Tscherskia triton	Greater long-tailed hamster	NC_013068.1
Cricetulus griseus	Striped hamster	NC_007936.1
Cricetulus longicaudatus	Lesser long-tailed hamster	KM067270.1
Allocricetulus eversmanni	Eversmann's hamster	in this study
Mesocricetus auratus	Golden hamster	EU660218.1
Phodopus roborovskii	Desert hamster	in this study
Eospalax rothschildi	Rothschild's zokor	NC_018535.1
Myospalax aspalax	Steppe zokor	NC_026915.1
Meriones meridianus	Mid-day jird	NC_027684.1
Meriones unguiculatus	Clawed jird	NC_023263.1
Meriones libycus	Libyan jird	NC_027683.1
Microtus fortis fortis	Reed vole	JF261174.1
Microtus kikuchii	Taiwan vole	AF348082.1
Eothenomys chinensis	Chinese oriental vole	NC_013571.1
Myodes rufocanus	Grey red-backed vole	NC_029477.1
Neodon irene	Pine vole	NC_016055.1

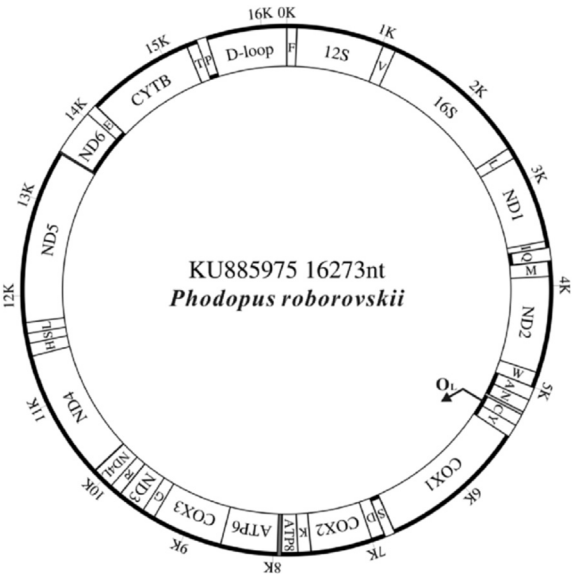

Fig. 1. Gene organization of the *Phodopus roborovskii* mitogenome. All tRNA genes are denoted by a single letter (amino acid abbreviation). A black thick line marks genes encoded on the H- and L-strands, with names listed outside and inside of the circle transcribed in the former and latter, respectively. The PCGs are indicated as follows: ND1-6, NADH dehydrogenase subunits 1–6; COX1-3, cytochrome *c* oxidase subunits 1–3; ATP6 and 8, ATP synthase F0 subunit 6 and 8; CYTB, cytochrome *b*; D-loop, control region; O_L, the putative origin of replication for the L-strand.

Table 2
Base composition (%) of the *Phodopus roborovskii* mitogenome.

Genes	T(U)	C	A	G	A + T	C + G
PCGS	27.5	27.1	33.4	12	60.9	39.1
1st	27	28.6	33.4	10.7	60.4	39.3
2nd	27	26.4	35.5	11	62.5	37.4
3rd	28	26.3	31.3	14.4	59.3	40.7
tRNA	28.8	20.5	35.3	15.5	64.1	36.0
12S rRNA	24.9	20.9	36.8	17.3	61.7	38.2
16S rRNA	25.0	19.7	37.6	17.7	62.6	37.4
D-loop	32.8	25.0	29.7	12.5	62.5	37.5
Complete sequence	27.4	25.3	33.9	13.3	61.3	38.6

Note: the triplet codon position is denoted by 1st (first), 2nd (second), 3rd (third).

3.2. Protein coding genes

The concatenated PCG was 11,392 bp long, accounting for approximately 70% of the total mitogenome. The PCG base composition was 33.4% A, 12% G, 27.1% C and 27.5% T. With 60.9% A + T content and 39.1% G + C content, *P. roborovskii* PCGs were AT-rich (Table 2). In contrast to other voles (Jiang et al., 2012) where the 3rd codon position in mitochondrial PCGs is more AT-rich than other positions, whereas the AT-content at the 3rd codon position in *P. roborovskii* was 59.3% lower than at the 1st (60.4%) or the 2nd (62.5%) codon position. In addition, the PCGs base feature of *P. roborovskii* showed that the G-content at the 3rd codon position was 14.4% higher than that at the 1st (10.7%) or the 2nd (11%) codon positions, and contrary to what other vertebrates were the G-content at the 3rd codon lower than that at the 1st or the 2nd codon positions, which was strongly biased against "G" at the 3rd codon position (Zhang et al., 2003a), such as *Andrias davidianus*, *M. f. calamorum*, *Vuples vuples*. The codon bias in *P. roborovskii* mtDNA was very apparent: CUA (Leu), AUA (Met), and ACA (Thr) were the three most frequently used codons (Table S2). All PCGs were encoded on the H-strand, except ND6 on the L-strand. Similar to all other Rodentia, ATPase 8 and 6 genes experienced the longest overlap of 43 nucleotides. The start codons were ATG (for COX1, COX2, ATP8, ATP6, COX3, ND4L, ND4, ND6, and CYTB), GTG (for ND1), ATT (for ND2, ND5), and ATA (for ND3). The stop codons were TAA (for COX1, ATP8, ATP6, ND3, ND4L, ND5, and CYTB), TAG (for COX2, ND6), and an incomplete stop codon (a single stop nucleotide "T") (for ND1, ND2, COX3, and ND4) (Table 3). Such incomplete stop codons are typically "completed" though polyadenylation (Chang and Tong, 2012); an incomplete stop codon has also been observed in other mammals (Gissi et al., 1998; Peng et al., 2007; Zhong et al., 2010).

3.3. Ribosomal and transfer RNA genes

The two *P. roborovskii* rRNA genes (12S rRNA and 16S rRNA) were annotated through alignment to homologous mitochondrial genes from other hamsters. As in other vertebrates, these rRNA genes were situated between $tRNA^{Phe}$ and $tRNA^{Leu}$

Table 3
Characterization of the *Phodopus roborovskii* mitogenome.

Gene	Position		Size(bp)	Strand	Codon		Intergenic nucleotides[a]
	From	To			Start	Stop	
tRNA-Phe	1	69	69	H			
12S rRNA	71	1016	946	H			1
tRNA-Val	1018	1089	72	H			1
16S rRNA	1090	2655	1566	H			0
tRNA-Leu[1]	2657	2731	75	H			1
ND1	2732	3686	955	H	GTG	T[b]	0
tRNA-Ile	3687	3754	68	H			0
tRNA-Gln	3752	3822	71	L			−3
tRNA-Met	3827	3902	76	H			4
ND2	3903	4935	1033	H	ATT	T[b]	0
tRNA-Trp	4936	5002	67	H			0
tRNA-Ala	5005	5074	70	L			2
tRNA-Asn	5076	5146	71	L			1
tRNA-Cys	5178	5245	68	L			31
tRNA-Tyr	5246	5312	67	L			0
COX1	5314	6858	1545	H	ATG	TAA	1
tRNA-Ser[1]	6856	6924	69	L			−3
tRNA-Asp	6928	6995	68	H			3
COX2	6997	7680	684	H	ATG	TAG	1
tRNA-Lys	7684	7750	67	H			3
ATP8	7752	7955	204	H	ATG	TAA	1
ATP6	7913	8593	681	H	ATG	TAA	−43
COX3	8593	9376	784	H	ATG	T[b]	−1
tRNA-Gly	9377	9444	68	H			0
ND3	9445	9792	348	H	ATA	TAA	0
tRNA-Arg	9796	9864	69	H			3
ND4L	9865	10,161	297	H	ATG	TAA	0
ND4	10,155	11,529	1375	H	ATG	T[b]	−7
tRNA-His	11,530	11,596	67	H			0
tRNA-Ser[2]	11,597	11,654	58	H			0
tRNA-Leu[2]	11,655	11,724	70	H			0
ND5	11,725	13,542	1818	H	ATT	TAA	0
ND6	13,526	14,050	525	L	ATG	TAG	−17
tRNA-Glu	14,051	14,119	69	L			0
CYTB	14,124	15,266	1143	H	ATG	TAA	4
tRNA-Thr	15,269	15,336	68	H			2
tRNA-Pro	15,337	15,403	67	L			0
D-loop	15,402	16,273	872	H			−2

[a] Numbers correspond to the nucleotides separating different genes. Negative numbers indicate overlapping nucleotides between adjacent genes.
[b] T(AA) stop codon was completed via polyadenylation.

and were separated by $tRNA^{Val}$ (Wolstenholme, 1992). The 12S and 16S rRNA genes were 946 bp and 1566 bp long, with a 61.7% and 62.6% A + T content (Table 2), respectively. The length of all tRNA genes combined was 1514 bp and accounted for approximately 9.3% of the complete *P. roborovskii* mitogenome. Individually, these genes ranged from 58 to 76 bp and were interspersed among rRNAs and PCGs. Nearly all tRNA genes (21 out of 22) could fold into a clover-leaf secondary structure. However, similar to other mammalian $tRNA^{ser2}$, the gene was only 58 bp long in *P. roborovskii* and didn't exhibit the dihydrouridine stem-loop structure. The secondary structure of lacking dihydrouridine arm is likely related to its structural compensation mechanism among tRNA arms (Steinberg and Cedergren, 1994) and is commonly found in other mammals (Kumazawa and Nishida, 1993; Gissi et al., 1998; Jiang et al., 2012).

3.4. Noncoding sequences

The largest non-coding sequence in the *P. roborovskii* mitogenome was the control region (D-loop), located between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes. This region contains the major regulatory elements for mitogenome replication and expression in mice (Sbisà et al., 1990), including the major sites of transcriptional initiation and origin of H-strand DNA replication in mammalian mtDNA (Clayton, 1991). However, prior to this study, D-loop structure remained fail to be determined for Cricetinae species, including *P. roborovskii*. The D-loop region is divided into three domains. First is the extended termination associated sequence (ETAS) domain that is adjacent to the $tRNA^{Pro}$ gene, where H-strand synthesis pauses, and associates with regulation of replication and transcription. Second is the central conserved domain. Third is the conserved sequence block (CSB) domain that is adjacent to the $tRNA^{Phe}$ gene, containing the origin of H-strand replication; two promoters and the CSB associate with the initiation of H-strand synthesis (Saccone et al., 1987; Sbisà et al., 1997; Jiang et al., 2012). The *Mus musculus* sequence was used as a reference to identify these three domains (Sbisà et al., 1997). The complete nucleotide

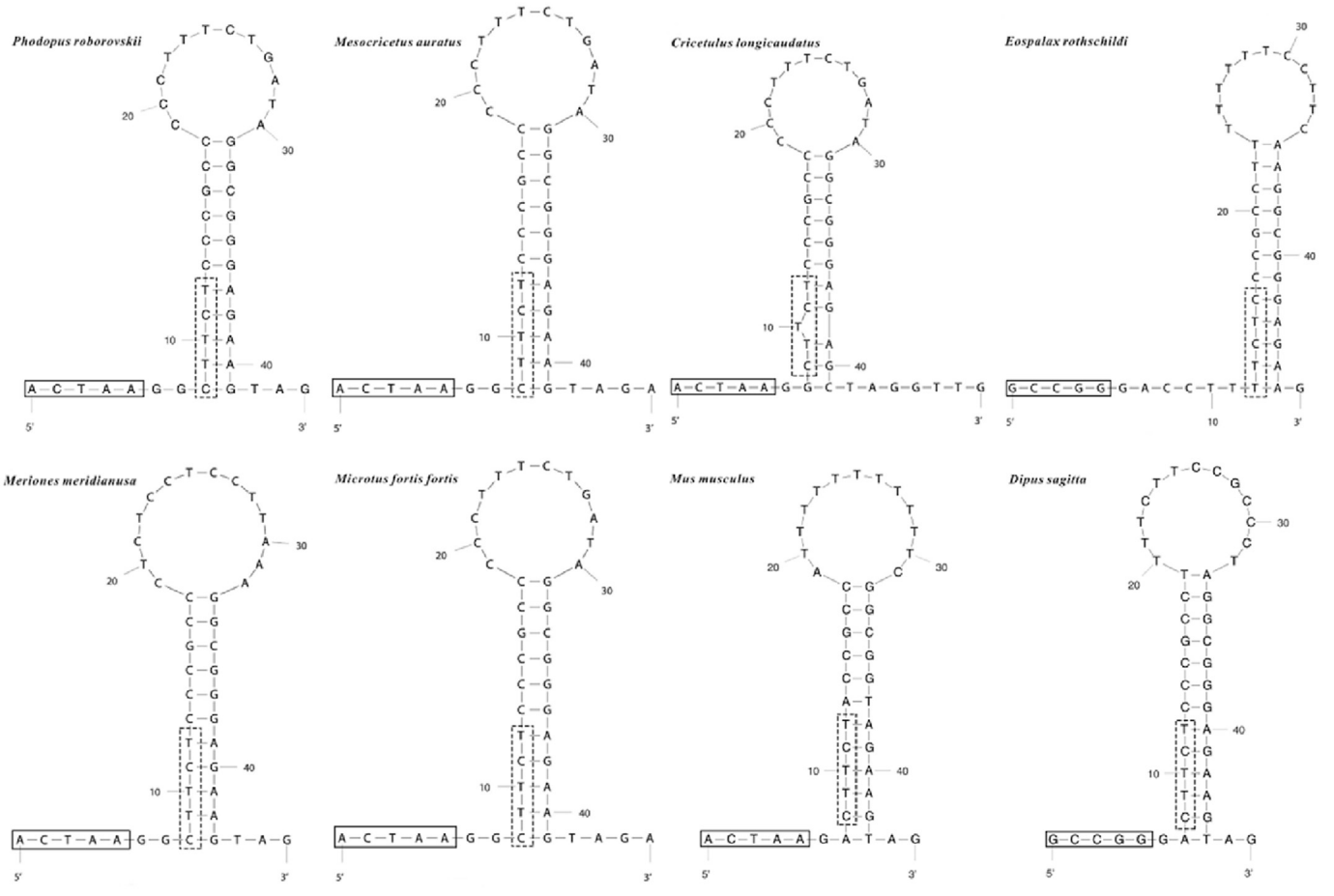

Fig. 2. Proposed stem-loop structures of the L-strand origin of replication in *Phodopus roborovskii* and other representative Rodents. Each sequence associated with the RNA-to-DNA transition is indicated with a solid box, and the conserved sequence (5′-CTTCT-3′) is indicated with a dashed box.

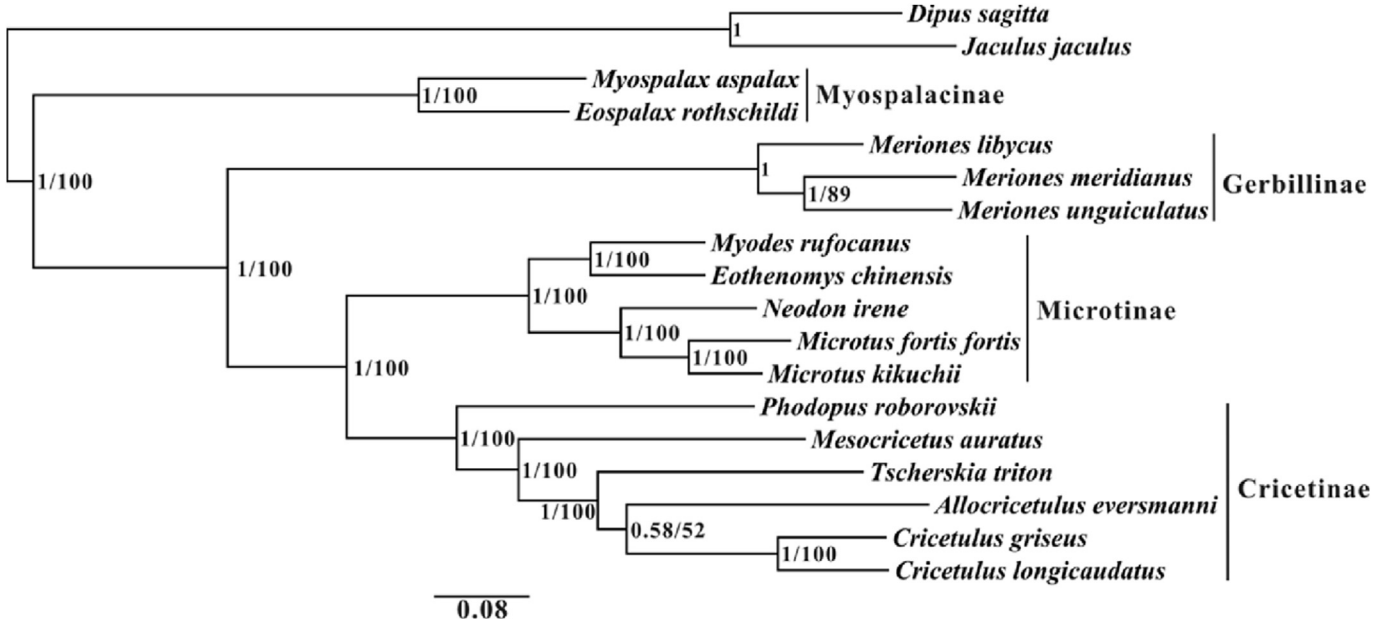

Fig. 3. Cricetidae phylogenetic relationships inferred from their mitogenome sequences. The BI/ML posterior probability and bootstrap values are indicated at the nodes.

sequence of the *P. roborovskii* D-loop was 872 bp long and its base composition was 32.8% T, 25.3% C, 33.9% A, and 13.3% G (Table 2). The D-loop's structural organization was similar to that of other Cricetinae species, but varied individual sequence length and base composition (Table S3). All species exhibited base frequencies of A + T > G + C in every D-loop domains (Table S3). When examining domains separately, ETAS and central domains possessed a base content of T > A in all hamster species, whereas A > T in the CSB domain and "G" was low in all the domains. The central domain appeared to differ in base-composition constraints compared with the two peripheral domains; we observed an increase of A + T and a decrease of G + C similar to results from previous comparisons of the D-loop in other mammals (Sbisà et al., 1997).

3.5. Putative origin of replication for the light strand (O_L)

Like most vertebrates, the O_L in the *P. roborovskii* mitogenome was located in the WANCY cluster. This region was approximately 45 bp long and had the potential to fold into a stem-loop secondary structure (Fig. 2). The replication starting sequence at the base of the O_L stem appeared commonly in Cricetidae, but with some minor differences between genera. Representative Cricetidae species (including members of Muroidea: mice) appeared to share the same beginning O_L sequence of 5′-ACTAA-3′, whereas in jerboa and zokor, O_L began with 5′-GCCGG-3′. Interestingly, despite the existence of mutations in the starting sequence (solid box) among other Rodentia, most stem-loop structures began with the 5′-CTTCT-3′ (dashed box), a strictly conserved DNA fragment that was used to locate the L-strand in the origin of replication (Fig. 2). It was more important that a representative species of the Myospalacinae group, *Eospalax rothschildi* (NC_018535.1), was an abnormal case which the stem-loop didn't begin with the 5′-CTTCT-3′ sequence. In addition, prior to this study, the motif 5′-GCCGG-3′ is found in other vertebrates (Zardoya et al., 1995) and has been reported to be critical for human mtDNA replication (Hixson et al., 1986).

3.6. Phylogenetic analysis

After the final alignment of the 18 study species (Table 1), the concatenated, 15,809 bp sequence (excluding sites with gaps and missing data) consisted of 8247 variable (polymorphic) sites that included 7230 potentially parsimony-informative sites and 1017 singleton sites. Among the 24 models tested, a GTR + I + G model was selected (-lnL = 150,149.6563 and AIC = 300,319.3125) as the most suitable model for the BI, ML and NJ analyses. BI and ML analyses yielded the same topological structure for mitogenome sequences (Fig. 3); the topological structure obtained from MP and NJ was also identical to each other (Supplementary Data, Fig. S1), although differing from BI and ML results. This difference mainly occurred in the *Cricetulus*-related subclade. However, all four methods yielded the same topological structure for *P. roborovskii*, representing a basal hamster taxon (BI = 1, ML = 100, MP = 100, and NJ = 100). This result agrees with previous molecular systematic research demonstrating that *P. roborovskii* was the most primitive species within Cricetinae (Norris et al., 2004; Neumann et al., 2006). The Cricetinae clade was strongly supported (BI = 1, ML = 100; Fig. 3) and the most basal genus was *Phodopus* (BI = 1, ML = 100) within Cricetinae. Within the Cricetinae clade, *P. roborovskii* separated into *M. auratus* (BI = 1, ML = 100); *M. auratus* has a close relationship with the *Cricetulus* group (BI = 1, ML = 100) whereas it was well separated from a group corresponding to *T. triton*, *A. eversmanni*, *C. griseus*, and *C. longicaudatus*. The phylogenetic trees derived from mitogenome sequences showed a clear partition of Palaearctic hamsters into three main evolutionary entities (*Phodopus*, *Mesocricetus*, and *Cricetulus* groups). This grouping was in line with previous systematic research based on mtDNA and nuclear gene markers (Neumann et al., 2006). Furthermore, previous karyotype research also considered *P. roborovskii* to be more primitive than *P. sungorus* and *P. campbelli*, or any other hamsters within Cricetinae (Schmid et al., 1986; Romanenko et al., 2007). Moreover, *P. roborovskii* also appeared highly divergent from the other hamsters, suggesting an ancient separation. Previous research found that *Phodopus* originated in the middle Miocene and diverged from *Mesocricetus* approximately 10.8–12.2 Myr (million years) ago, whereas radiation within the *Cricetulus*-related group began 6.5–7.5 Myr ago (Neumann et al., 2006). Therefore, all available evidence indicates that *P. roborovskii* is the most primitive species among extant hamsters within Cricetinae.

Acknowledgements

The work was supported by the National Natural Science Foundation of China (Nos.30870294, 31372179).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.bse.2016.10.010.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Arèvalo, E., Davis, S.K., Sites, J.W., 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. 43, 387–418.
Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W., Clayton, D.A., 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell 26, 167–180.
Boore, J.L., 2006. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol. Evol. 21, 439–446.
Chang, J.H., Tong, L., 2012. Mitochondrial poly(A) polymerase and polyadenylation. Biochimica Biophysica Acta 1819, 992–997.
Churakov, G., Sadasivuni, M.K., Rosenbloom, K.R., Huchon, D., Brosius, J., Schmitz, J., 2010. Rodent evolution: back to the root. Mol. Biol. Evol. 27, 1315–1326.
Clayton, D.A., 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. 7, 453–478.
Corbet, G.B., 1978. The Mammals of the Palaearctic Region: a Taxonomic Review. British Museum (Natural History). Cornell University Press, London and Ithaca, pp. 88–140.
Fan, Z.X., Liu, S.Y., Liu, Y., Liao, L.H., Zhang, X.Y., Yue, B.S., 2012. Phylogeography of the South China field mouse (*Apodemus draco*) on the Southeastern Tibetan plateau reveals high genetic diversity and Glacial Refugia. Plos One 7, e38184.
Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. 20, 406–416.
Flint, V.E.e., 1966. Die Zwerghamster der paläarktischen Fauna. Ziemsen Verlag, Lutherstadt-Wittenberg.
Gissi, C., Gullberg, A., Arnason, U., 1998. The complete mitochondrial DNA sequence of the rabbit, *Oryctolagus cuniculus*. Genomics 50, 161–169.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hao, H.B., Liu, S.Y., Zhang, X.Y., Chen, W.C., Song, Z.B., Peng, H.Y., Liu, Y., Yue, B.S., 2011. Complete mitochondrial genome of a new vole *Proedromys liangshanensis* (Rodentia: Cricetidae) and phylogenetic analysis with related species: are there implications for the validity of the genus *Proedromys*? Mitochondrial DNA 22, 28–34.
Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4, 6–11.
Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.
Hixson, J., Wong, T., Clayton, D.A., 1986. Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication. J. Biol. Chem. 261, 2384–2390.
Hutchison, C.A., Newbold, J.E., Potter, S.S., Edgell, M.H., 1974. Maternal inheritance of mammalian mitochondrial DNA. Nature 251, 536–538.
Jiang, X.H., Gao, J., Ni, L.J., Hu, J.H., Li, K., Sun, F.P., Xie, J.Y., Bo, X., Gao, C., Xiao, J.H., Zhou, Y.X., 2012. The complete mitochondrial genome of *Microtus fortis calamorum* (Arvicolinae, Rodentia) and its phylogenetic analysis. Gene 498, 288–295.
Kumazawa, Y., Nishida, M., 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37, 380–398.
Li, D.M., Fan, L.Q., Zeng, B., Yin, H.L., Zou, F.D., Wang, H.X., Meng, Y., King, E., Yue, B.S., 2009. The complete mitochondrial genome of *Macaca thibetana* and a novel nuclear mitochondrial pseudogene. Gene 429, 31–36.
Librado Sanz, P., Rozas Liras, J.A., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Lowe, T.M., Eddy, S.R., 1997. tRNA scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Lu, L., Ge, D.Y., Chesters, D., Ho, S.Y.W., Ma, Y., Li, G.C., Wen, Z.X., Wu, Y.J., Wang, J., Xia, L., Liu, J.L., Guo, T.Y., Zhang, X.L., Zhu, C.D., Yang, Q.S., Liu, Q.Y., 2015. Molecular phylogeny and the underestimated species diversity of the endemic white-bellied rat (Rodentia: Muridae: *Niviventer*) in Southeast Asia and China. Zool. Scr. 44, 475–494.
Luo, Z.X., Chen, W., Gao, W., et al., 2000. Fauna Sinica: Mammalia. In: Rodentia Part III: Cricetidae, vol. 6. Chinese Academy of Sciences, Science Press, Beijing, pp. 81–86.
Neumann, K., Michaux, J., Lebedev, V., Yigit, N., Colak, E., Ivanova, N., Poltoraus, A., Surov, A., Markov, G., Maak, S., 2006. Molecular phylogeny of the Cricetinae subfamily based on the mitochondrial cytochrome b and 12S rRNA genes and the nuclear vWF gene. Mol. Phylogenetics Evol. 39, 135–148.
Newkirk, K.D., Cheung, B.L., Scribner, S.J., Wynne-Edwards, K.E., 1998. Earlier thermoregulation and consequences for pup growth in the Siberian versus Djungarian dwarf hamster (*Phodopus*). Physiology Behav. 63, 435–443.
Norris, R.W., Zhou, K.Y., Zhou, C.Q., Yang, G., Kilpatrick, C.W., Honeycutt, R.L., 2004. The phylogenetic position of the zokors (Myospalacinae) and comments on the families of muroids (Rodentia). Mol. Phylogenetics Evol. 31, 972–978.
Nylander, J.A.A., 2004. MrModeltest v2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.
Partridge, M.A., Davidson, M.M., Hei, T.K., 2007. The complete nucleotide sequence of Chinese hamster (*Cricetulus griseus*) mitochondrial DNA. DNA Seq. 18, 341–346.
Peng, R., Zeng, B., Meng, X.X., Yue, B.S., Zhang, Z.H., Zou, F.D., 2007. The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*). Gene 397, 76–83.
Robins, J.H., Hingston, M., Matisoo-Smith, E., Ross, H.A., 2007. Identifying *Rattus* species using mitochondrial DNA. Mol. Ecol. Notes 7, 717–729.
Romanenko, S.A., Volobouev, V.T., Perelman, P.L., Lebedev, V.S., Serdukova, N.A., Trifonov, V.A., Biltueva, L.S., Nie, W., O'Brien, P.C., Bulatova, N.S., 2007. Karyotype evolution and phylogenetic relationships of hamsters (Cricetidae, Muroidea, Rodentia) inferred from chromosomal painting and banding comparison. Chromosome Res. 15, 283–298.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Ross, P.D., 1994. Mammalian species: Phodopus roborovskii. Am. Soc. Mammal. 459, 1–4.
Saccone, C., Attimonelli, M., Sbisà, E., 1987. Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. J. Mol. Evol. 26, 205–211.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Sbisà, E., Nardelli, M., Tanzariello, F., Tullo, A., Saccone, C., 1990. The complete and symmetric transcription of the main non coding region of rat mitochondrial genome: in vivo mapping of heavy and light transcripts. Curr. Genet. 17, 247–253.
Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., Saccone, C., 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205, 125–140.
Schmid, M., Haaf, T., Weis, H., Schempp, W., 1986. Chromosomal homoeologies in hamster species of the genus *Phodopus* (Rodentia, Cricetinae). Cytogenet. Genome Res. 43, 168–173.
Steinberg, S., Cedergren, R., 1994. Structural compensation in atypical mitochondrial tRNAs. Nat. Struct. Mol. Biol. 1, 507–510.
Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis Using Parsimony, version 4.0 b10. Sinauer Associates, Sunderland, Massachusetts.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tu, F.Y., Fan, Z.X., Murphy, R.W., Chen, S.D., Zhang, X.Y., Yan, C.C., Liu, Y., Sun, Z.Y., Fu, J.R., Liu, S.Y., Yue, B.S., 2014. Molecular phylogenetic relationships among Asiatic shrewlike moles inferred from the complete mitogenomes. J. Zoological Syst. Evol. Res. 53, 155–160.
Wang, S., Zheng, C.L., 1973. Notes on Chinese hamsters (Cricetinae). Acta Zool. Sin. 19, 61–68.
Wang, Y., Zhao, L.M., Fang, F.J., Liao, J.C., Liu, N.F., 2013. Intraspecific molecular phylogeny and phylogeography of the *Meriones meridianus* (Rodentia: Cricetidae) complex in northern China reflect the processes of desertification and the Tianshan Mountains uplift. Biol. J. Linn. Soc. 110, 362–383.
Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. 141, 173–216.

Wynne-Edwards, K.E., Lisk, R.D., 1987. Behavioral interactions differentiate Djungarian (*Phodopus campbelli*) and Siberian (*Phodopus sungorus*) hamsters. Can. J. Zoology 65, 2229–2235.
Yang, Z.H., Rannala, B., 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. Mol. Biol. Evol. 14, 717–724.
Zardoya, R., Garrido-Pertierra, A., Bautista, J.M., 1995. The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. J. Mol. Evol. 41, 942–951.
Zardoya, R., Meyer, A., 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. 13, 933–942.
Zhang, P., Chen, Y.Q., Liu, Y.F., Zhou, H., Qu, L.H., 2003a. The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). Gene 311, 93–98.
Zhang, P., Chen, Y.Q., Zhou, H., Wang, X.L., Qu, L.H., 2003b. The complete mitochondrial genome of a relic salamander, *Ranodon sibiricus* (Amphibia: Caudata) and implications for amphibian phylogeny. Mol. Phylogenetics Evol. 28, 620–626.
Zhong, H.M., Zhang, H.H., Sha, W.L., Zhang, C.D., Chen, Y.C., 2010. Complete mitochondrial genome of the red fox (*Vuples vuples*) and phylogenetic analysis with other canid species. Zoological Res. 31, 122–130.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.