Biochemical Systematics and Ecology 69 (2016) 195–203

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

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

Characterization and phylogenetic analysis of the complete mitogenome of *Allactaga sibirica* (Rodentia: Dipodidae)

Li Ding, Guangjie Luo, 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 5 October 2016
Accepted 8 October 2016

Keywords:
Allactaga sibirica
Mitochondrial genome
Sequence analysis
Dipodidae
Phylogenetic analysis

ABSTRACT

Allactaga sibirica (Dipodidae) is widely distributed in the northwestern arid regions of China. The complete mitochondrial genome (mitogenome) of *A. sibirica* was 16,685 bp in length; included 13 protein-coding genes, 2 ribosome genes, 22 tRNA genes, and one control region; and had a structure that was typical of vertebrates. The base composition and codon usage of the mitogenome are described, and the structure of the non-coding sequence in the *A. sibirica* is reported for the first time. The putative origin of replication for the light strand of *A. sibirica* was approximately 45 bp long, and was highly conserved in the stem-loop and adjacent sequences. Phylogenetic analyses showed high resolution in each of the main divergent clades within Dipodoidea using mitogenomes data. The results indicated that the Zapodidae group was a representative of very basal taxon in Dipodoidea, and shared a common ancestor with Dipodidae species. Within Dipodidae clade, Allactaginae species was at basal position, and this result was in line with previous molecular systematic and morphological studies. Furthermore, *Euchoreutes naso* and *A. sibirica* had a close relationship could implicate a sister-group relationship between Euchoreutinae and Allactaginae. Meanwhile, this work also provided a set of useful data on phylogeny and molecular evolution in Dipodidae species.

1. Introduction

Dipodoidea originated in the Eocene and arose from a common ancestor shared with Muroidea (Wang, 1985; Zhang et al., 2013). According to the latest research, the group contains 51 extant species that are classified into 16 genera and are widely distributed in the Holarctic region (Holden and Musser, 2005). Dipodoidea comprises of the Zapodidae and Dipodidae (Vinogradov, 1925). The Zapodidae contains two subfamilies (Sicistinae and Zapodinae), members of which live in woodlands, grasslands, and meadows and are disjunctively distributed in China and North America (Zhang et al., 2013). On the other hand, the Dipodidae contains four subfamilies (Allactaginae, Cardiocraniinae, Dipodinae, and Euchoreutinae), members of which are desert dwellers and specialized for saltatorial locomotion, especially in regard to their hindlimb (Zhang et al., 2013). Based on morphological characteristics and color variation in China, extant Dipodoidea species consist of 11 recognized species across seven genera within Dipodidae, and 2 species are classified into two genera within Zapodidae (Zhao, 1991; Tan, 1992). The widespread distribution and distinct adaptations of jerboas makes the species ideal for studying the relationships between environmental change, geological events, and the origin of Rodentia. Moreover, the use of increasingly accurate

* Corresponding author.
E-mail addresses: dingl14@lzu.edu.cn (L. Ding), guangjieluo702@163.com (G. Luo), 1132905461@qq.com (W. Li), liaojch@lzu.edu.cn (J. Liao).

http://dx.doi.org/10.1016/j.bse.2016.10.004
0305-1978/

molecular markers can facilitate the detailed description of phylogenetic relationships that link evolutionary history, and demographic processes.

Previous studies have found that mitochondrial DNA (mtDNA) is a useful marker system for the phylogenetic analysis of vertebrate relationships (Zhang et al., 2003b), and the mtDNA of animals has been widely used for studying animal origin and evolution, population genetics, and phylogenetic relationships, owing to their maternal inheritance, relatively low rates of recombination, and high substitution rates (Harrison, 1989; Wang et al., 2013; Lu et al., 2015). However, single mitochondrial genes may be insufficient for resolving phylogenetic relationships because they provide a limited amount of phylogenetic information (Cummings et al., 1995; Boore, 2006; Tu et al., 2014). Therefore, the complete mitochondrial genome (mitogenome) is expected to provide more reliable estimates of evolutionary relationships. However, complete mitogenome sequences have only been determined for a few Dipodoidea species so far.

Allactaga sibirica is the largest jerboa species and has a wide distribution in the northern and northwestern arid regions of China (Wang, 1991). The species is adaptable and abundant in a variety of habitats. In addition, *A. sibirica* can endure extreme arid desert conditions and preferentially inhabits the desert and semi-desert regions of the Gobi Desert, and the range further extends into the grasslands and mountains (Liao et al., 2016). Furthermore, the species inhabits habitats that range in altitude from 1000 to 2500 m above sea level (a.s.l.), and even to 4000 m a.s.l. in the Qinghai-Tibetan plateau (Zhang and Wang, 1963; Zhao, 1991). To date, the complete mitogenome sequence of *A. sibirica* was not reported, and yet didn't know the taxonomic status and the phylogenetic relationships between Allaetaginae and Dipodinae. Therefore, thus hampering the development of well-supported phylogenetic inferences based on the analyses of full mitogenomes.

Previous molecular systematic studies, they only used fewer mitogenome data or partial gene marker to reconstruct phylogenetic trees among Dipodoidea (Montgelard et al., 2008; Yue et al., 2015), and failed to resolve the phylogenetic relationships between Allaetaginae and Dipodinae using mitogenome data. In the present study, we determined the complete mitochondrial sequence of *A. sibirica* and characterized the complete sequence by comparing it with the mitogenomes of other jerboa species. Phylogenetic analysis was carried out using data from the present study, as well that from previously published study of Dipodoidea species, with the aim of elucidating phylogenetic relationships within Dipodoidea, and provided a set of useful data on phylogeny and molecular evolution in Dipodidae species.

2. Materials and methods

2.1. Sampling and ethics statement

A specimen of *A. sibirica* was identified based on external characteristics and was further confirmed according to skull morphology and molecular data in the laboratory following the morphological characteristics (Wang, 1991; Robins et al., 2007). An *A. sibirica* specimen was caught using the snap-trap method in Jingtai County (N 37°11′1.78″, E 104°3′47.17″, 1629 m a.s.l.), Gansu Province, China, then both muscle samples and voucher specimens were preserved in absolute ethanol and formalin solutions. The sampling collections and study had no conflict with ethical guidelines, religious beliefs, or legal requirements of China. In addition, 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, Lanzhou, China.

2.2. DNA extraction, amplification, and sequencing

The total genomic DNA was extracted from an *A. sibirica* muscle sample using the TIANamp Genomic DNA Kit (Tiangen, Beijing, China), and the partial conserved sequences were amplified using polymerase chain reaction (PCR) with three primers (Kocher et al., 1989; Arèvalo et al., 1994). The remaining long fragments were sequenced using long-and-accurate PCR with primers that were specifically designed to produce fragments that overlapped by > 100 bp (Table S1) (Zhang et al., 2003a). The amplification was performed in 50 μl reaction mixtures that contained 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 conditions included a pre-denaturing step of 3 min at 98 °C, which was followed by 35 cycles of denaturing 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 with ethidiumbromide and then directly sequenced using the Sanger dideoxy sequencing method (Genewiz Biotech Co., Ltd., Suzhou, China).

2.3. Mitogenome sequence analysis and gene identification

The DNA sequence was edited and assembled manually using the software Lasergene version 5.0 (DNASTAR) aligned with BioEdit version 7.0.5.2 program (Hall, 1999). Contig assembly was performed using Seqman program (Lasergene subprogram). Nucleotide base composition and codon usage of protein-coding genes (PCGs) were calculated using MEGA 6.0 (Tamura et al., 2013). The locations of PCGs, tRNA genes, and rRNA genes were identified using the basic local alignment search tool (BLAST) (Altschul et al., 1990) and BioEdit to compare the *A. sibirica* sequence with homologous reference sequences from *D. sagitta* and *E. naso*. Most tRNA genes were further validated using tRNA scan-SE1.2.1 (Lowe and Eddy, 1997) with the default search mode using the vertebrate mitochondrial genetic code source. We also used reference sequence to identify the D-loop region

in the *A. sibirica* sequence (Sbisà et al., 1997). However, since the D-loop structure had not yet been reported for *A. sibirica* and other jerboas. We used the DNA sequences of the previously characterized from *Glis glis* as the reference sequence to locate three blocks by local alignment with BioEdit program (Sbisà et al., 1997). A circular map of the *A. sibirica* mitogenome was drawn using the Mtv online tool (http://pacosy.informatik.uni-leipzig.de/mtviz/mtviz). The putative origin of light 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 original replication sequence 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 Dipodoidea species (Zuker, 2003).

2.4. Phylogenetic analysis

The phylogenetic relationships among *A. sibirica* and several other Dipodoidea species were estimated by analyzing their mitogenome sequence (Table 1). The polymorphic sites were analyzed using DnaSP v5 program (Librado Sanz and Rozas Liras, 2009) and multiple sequence aligned with BioEdit. Phylogenetic analyses were performed using PAUP 4.0b10 (Swofford, 2003) and MrBayes 3.2.6 (Ronquist et al., 2012) with Bayesian inference (BI) (Yang and Rannala, 1997), maximum likelihood (ML) (Felsenstein, 1981), maximum parsimony (MP) (Fitch, 1971), and distance-based neighbor-joining analyses (NJ) (Saitou and Nei, 1987) methods. We used MrModeltest 2.3 (Nylander, 2004) to identify the best-fit model of evolution under the Akaike Information Criterion (AIC) (Akaike, 1974), and used the General Time Reversible (GTR), accordingly for the BI, NJ, and ML analyses. Both ML and NJ employed the same maximum likelihood model, whereas the MP analyses were performed using a heuristic search and a tree-bisection-and-reconnection (TBR) branch swapping option. The robustness of inferences was assessed by bootstrapping (BT) (1000 random repetitions for MP and NJ and 100 random repetitions for ML) (Felsenstein, 1985). For all analyses, only clades with 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. According to previous studies (Yue et al., 2015), *Ochotona curzoniae* (NC_011029.1) and *Lepus capensis* (NC_015841.1) 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 *A. sibirica* mitogenome sequence was 16,685 bp in length (GenBank accession number: KX058130), which was longer than that of *D. sagitta* (16,664 bp), but shorter than that of *E. naso* (16,705 bp). The *A. sibirica* mitogenome contained 13 PCGs, two rRNA genes (12S rRNA and 16S rRNA), 22 tRNA genes, and one control region, and the mitogenome was similar to those of other mammals, in regard to both gene quantity and organizational structure (Bibb et al., 1981; Jiang et al., 2012; Yue et al., 2015) (Fig. 1). The heavy strand (H-strand) base composition was 31.4% adenine (A), 13.8% guanine (G), 28.5% cytosine (C), and 26.2% thymine (T) (Table 2). Thus, A and G were the most (31.4%) and least (13.8%) common base, respectively, and the A + T content (57.6%) was greater than the G + C content (42.3%) (Table 2). These characteristics were similar to those of other mammalian mitogenome sequences (Partridge et al., 2007; Jiang et al., 2012). Moreover, the 22 tRNA genes were interspersed among the rRNA genes and PCGs (Fig. 1), with few gaps between them, and some genes even overlapped with each other, with the longest overlap (43 bp) observed between the ATPase 8 and 6 genes (Table 3). All genes were encoded on the H-strand, except the ND6 gene and nine tRNA genes ($tRNA^{Ile}$, $tRNA^{Gln}$, $tRNA^{Trp}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Ser1}$, $tRNA^{Glu}$, and $tRNA^{Pro}$), which were encoded on the light strand (L-strand) (Table 3).

Table 1
Dipodoidea species and their mitogenome sequence accession numbers used for phylogenetic analysis.

Species	Common name	GenBank accession number
Dipus sagitta	Northern Three-toed Jerboa	NC_027499.1
Euchoreutes naso	Long-eared Jerboa	NC_027500.1
Jaculus jaculus	Lesser Egyptian Jerboa	NC_005314.1
Allactaga sibirica	Mongolian five-toed Jerboa	in this study
Stylodipus telum	Feather-tailed Jerboa	NC_027692.1
Sicista concolor	Chinese Birch Mouse	NC_027579.1
Eozapus setchuanus	Chinese Jumping Mouse	NC_027578.1

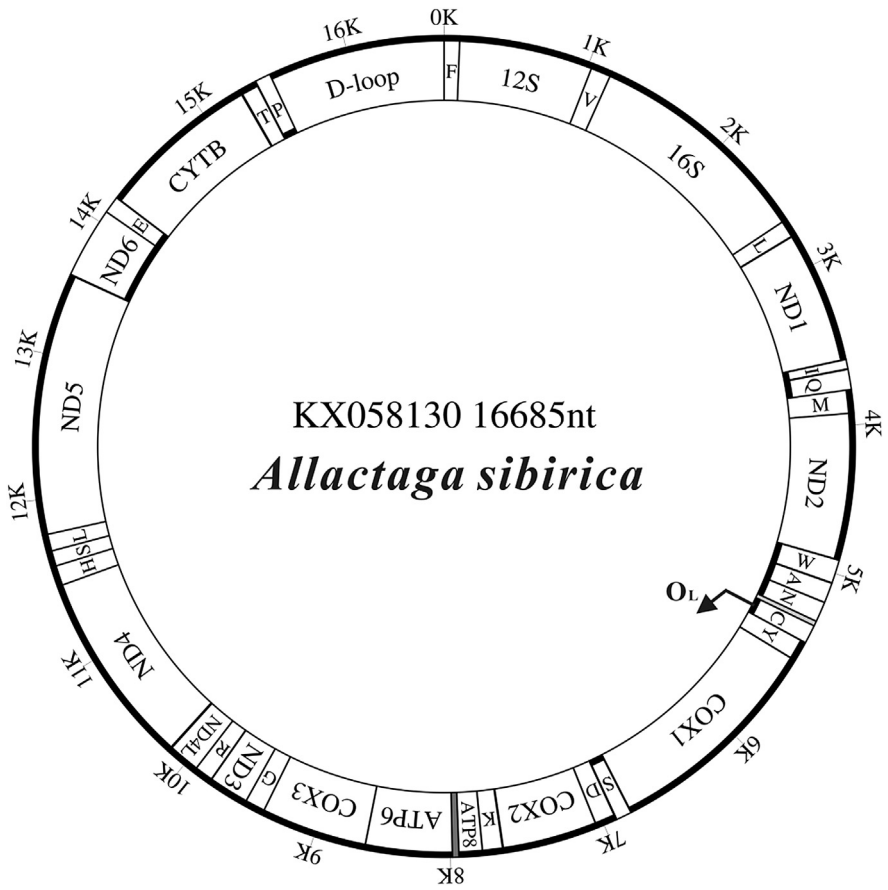

Fig. 1. Gene organization of the *Allactaga sibirica* mitogenome. All tRNA genes were denoted by a single letter (amino acid abbreviation). The genes encoded on the H- and L-strands are marked with thick black lines, and gene symbols listed outside and inside of the circle were transcribed in the H- and L-strands, respectively. ND1-6, NADH dehydrogenase subunits 1–6; COX1-3, cytochrome *c* oxidase subunits 1–3; ATP 6 and 8, ATP synthase F0 subunit 6 and 8; CYTB, cytochrome *b*; and D-loop, control region; and O_L, the putative origin of replication for the L-strand (O_L).

Table 2
Base composition (%) of the *Allactaga sibirica* mitogenome.

Genes	T(U)	C	A	G	A + T	C + G
PCGs	26.6	31.0	30.0	12.3	56.6	43.3
1st	26.0	28.3	29.5	16.1	55.5	44.4
2nd	30.0	32.7	36.3	11.0	56.3	43.7
3rd	24.0	32.0	34.1	10.0	58.1	42.0
tRNA	26.7	22.8	34.3	16.3	61.0	39.1
12S rRNA	23.9	23.0	33.4	19.7	57.3	42.7
16S rRNA	24.2	20.3	37.1	18.3	61.3	39.6
D-loop	26.4	26.6	33.0	14.0	59.4	40.6
Complete sequence	26.2	28.5	31.4	13.8	57.6	42.3

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

3.2. Protein coding genes

The concatenated PCGs were 11,396 bp long, thus accounting for approximately 68.3% the total mitogenome. The PCG base composition was 30.0% A, 12.3% G, 31.0% C and 26.6% T, and the A + T content (56.6%) of the PCGs was greater than the G + C content (43.3%) (Table 2). It was common that the 3rd codon position in the mitochondrial PCGs was AT-rich; the AT-content of the 3rd position was 58.1%, whereas that of the 1st and 2nd positions was 55.5% and 56.3%, respectively (Table 2). The AT-content of the *A. sibirica* PCGs was the same as that of other Rodentia (Jiang et al., 2012). In addition, the PCGs of *A. sibirica* had a strong bias against the use of 'G' at the 3rd codon position (10.0%) lower than that at the 1st (16.1%) or the 2nd (11%) codon position, which was typical of vertebrates (Zhang et al., 2003a), such as *Andrias davidianus*, *Microtus fortis calamorum*, *Vuples*

Table 3
Characterization of the *Allactaga sibirica* mitogenome.

Gene	Position		Size (bp)	Strand	Codon		Intergenic nucleotides[a]
	From	To			Start	Stop	
tRNA-Phe	1	64	64	H			
12S rRNA	65	1030	966	H			0
tRNA-Val	1031	1099	69	H			0
16S rRNA	1100	2677	1578	H			0
tRNA-Leu[1]	2678	2752	75	H			0
ND1	2756	3712	957	H	ATG	TAG	3
tRNA-Ile	3711	3779	69	L			−2
tRNA-Gln	3777	3848	72	L			−3
tRNA-Met	3849	3918	70	H			0
ND2	3919	4962	1044	H	ATG	TAG	0
tRNA-Trp	4961	5028	68	L			−2
tRNA-Ala	5032	5100	69	L			3
tRNA-Asn	5102	5175	74	L			1
tRNA-Cys	5209	5273	65	L			33
tRNA-Tyr	5275	5341	67	H			1
COX1	5343	6887	1545	H	ATG	TAA	1
tRNA-Ser[1]	6885	6954	70	L			−3
tRNA-Asp	6959	7024	66	H			4
COX2	7026	7709	684	H	ATG	TAG	1
tRNA-Lys	7715	7782	68	H			−5
ATP8	7784	7987	204	H	ATG	TAG	1
ATP6	7945	8625	681	H	ATG	TAA	−43
COX3	8625	9408	784	H	ATG	T[b]	−1
tRNA-Gly	9409	9477	69	H			0
ND3	9478	9825	348	H	ATA	TAA	0
tRNA-Arg	9828	9895	68	H			2
ND4L	9896	10,192	297	H	ATG	TAA	0
ND4	10,186	11,563	1378	H	ATG	T[b]	−7
tRNA-His	11,564	11,630	67	H			0
tRNA-Ser[2]	11,631	11,687	57	H			0
tRNA-Leu[2]	11,688	11,757	70	H			0
ND5	11,758	13,566	1809	H	ATT	TAA	0
ND6	13,563	14,087	525	L	ATG	TAG	−4
tRNA-Glu	14,088	14,156	69	L			0
CYTB	14,159	15,298	1140	H	ATG	AGA	2
tRNA-Thr	15,305	15,371	67	H			6
tRNA-Pro	15,375	15,441	67	L			3
D-loop	15,442	16,685	1244	H			0

[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.

vuples. Furthermore, when the codon bias of the *A. sibirica* mtDNA was analyzed, we found that CUA (Leu), AUC (Ile), AUA (Met), and UUC (Phe) occurred most frequently used codons (Table S2). All the PCGs were encoded on the H-strand except the ND6 gene, which was encoded on the L-strand, and the longest overlap (43 bp) was by the ATPase 8 and 6 genes, as in other Rodentia (Yue et al., 2015). Although the ATG start codon was used for ND1, ND2, COX1, COX2, ATP6, ATP8, COX3, ND4L, ND4, ND6, and CYTB, the ATA codon was used for ND3, and the ATT codon was used for ND5. Moreover, the stop codon for COX1, ATP6, ND3, ND4L, and ND5 was TAA; whereas the stop codon for ND1, ND2, COX2, ATP8, and ND6 was TAG, the AGA stop codon was used for CYTB, and both COX3 and ND4 had an incomplete stop codon (a single stop nucleotide "T") (Table 3). Commonly, when incomplete stop codons were present, and they were supposed to be "completed" by polyadenylation (Chang and Tong, 2012). Incomplete stop codons have also been observed in other mammals, including *Oryctolagus cuniculus*, *Ailuropoda melanoleuca*, *Equus asinus* and *V. vuples* (Xu et al., 1996; Gissi et al., 1998; Peng et al., 2007).

3.3. Ribosomal and transfer RNA genes

The two rRNA genes (12S rRNA and 16S rRNA) of *A. sibirica* were annotated through alignment to homologous sequences from other jerboa mitogenomes. As in other vertebrates, the rRNA genes were situated between $tRNA^{Phe}$ and $tRNA^{Leu}$ and were separated by the $tRNA^{Val}$ (Wolstenholme, 1992). The 12S rRNA gene was 966 bp long, whereas the 16S rRNA gene was 1578 bp long; the A + T content of the two genes were 57.3% and 61.3%, respectively (Table 2). Together, the tRNA genes were 1500 bp long and accounted for approximately 9% of the complete mitogenome. The length of the *A. sibirica* tRNA genes ranged from 57 to 75 bp, and the genes were interspersed among the rRNA genes and PCGs. Twenty-one out of twenty-two tRNA genes could fold into clover-leaf secondary structures. However, the $tRNA^{ser2}$ was only 57 bp long and had lost the dihydrouridine stem and loop, and the tRNA structure of lacking dihydrouridine arm was found in other mammals, such as *O.*

cuniculus, *V. vuples*, *Homo sapiens* and *Bos taurus* (de Bruijn et al., 1980; Kumazawa and Nishida, 1993; Zhong et al., 2010). This exceptional secondary structure could work well related to its structural compensation mechanism among tRNA arms (Steinberg and Cedergren, 1994).

3.4. Non-coding sequences

The control region (D-loop) of *A. sibirica* was the largest non-coding sequence in the mitochondrial genome and was located between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes. This region contains the major regulatory elements for the replication and expression of the mitogenome in mice (Sbisà et al., 1990), and the region is the major site of transcriptional initiation and the origin of H-strand DNA replication in mammalian mtDNA (Clayton, 1991). However, such regions have not yet been clearly determined in *A. sibirica*, or other jerboa species. In general, the region can be divided into three domains. First is the extended termination-associated sequence (ETAS) domain that is adjacent to the $tRNA^{Pro}$ gene where the synthesis of the heavy strand 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, which contains the H-strand origin of replication; the two promoters and the CSB, which have been associated with the initiation of H-strand synthesis (Saccone et al., 1987; Sbisà et al., 1997; Jiang et al., 2012). The complete nucleotide sequence of the *A. sibirica* D-loop was 1244 bp in length, and the base composition of this region was shown in Table 2. Like other genes in the *A. sibirica* mitogenome, A and G were the most (33.0%) and least (14.0%) common base, respectively, and the AT-content was still AT-rich. The structure organization of the D-loop region was similar in other mammals but the sequence length varied between genera and species. Similarly, the D-loop region of *A. sibirica* contained three blocks (the ETAS, central, and CSB domains), which were 439, 336, and 469 bp long (Table S3), respectively. We observed that the central domain had a higher 'G' content (16.7%) than the ETAS (14.6%) or CSB domains (11.5%). This characteristic is also common to other species and might be related to its conserved function (Sbisà et al., 1997).

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

As in most vertebrates, the O_L of the *A. sibirica* mitogenome was located in the WANCY region of the tRNA gene. The region was approximately 45 bp long and had the potential to fold into a stem-loop secondary structure (Fig. 2). The 5′-GCCGG-3′ replication starting sequence at the base of the O_L stem was commonly found in Dipodoidea, but it was not strict congruence (5′-GCCGG-3′) sequence with the differences existed (dashed box) between genera and species (Fig. 2), such as *E. naso*, *S. telum*. Interestingly, most of the stem-loop structures began with the 5′-CTTCT-3′ (solid box) sequence and it was also a strictly conserved DNA fragment that could be used to locate the O_L even if the synthesis of starting sequences with the mutation loci existed. Furthermore, we also found that closely related species tended to possess similar O_L structures (Figs. 2 and 3), such as *J. jaculus* and *S. telum*. Previous about O_L studies, the motif 5′-GCCGG-3′ is found in other vertebrates, such as *Xenopus laevis*, *Phoca vitulina*, *Homo sapiens*, *Crossostoma lacustre*, *Gadus rnorhua*, *Cyprinus carpio* and *Oncorhynchus mykiss* (Zardoya et al., 1995), and it has been reported to be critical for human mtDNA replication (Hixson et al., 1986).

3.6. Phylogenetic analysis

A total of 16,138 bp of concatenated sequences (excluding sites with gaps and missing data) contained 7770 variable (polymorphic) sites, with 5567 potentially parsimony informative sites, and 2203 singleton sites after the final alignment for the nine species listed in Table 1. Among the 24 models tested, a GTR + I + G model (-lnL = 98,336.4531 and AIC = 196,692.9063) that was considered to be the most appropriate model for the BI, ML and NJ analyses. Both BI and ML analyses of the data set produced the same topological structures (Fig. 3). However, the MP and NJ analyses produced a different topology, in which *E. naso* and *A. sibirica* were not clustered in a clade (Supplementary Data, Figs. S1 and S2). In addition, *S. concolor* and *E. setchuanus* were not grouped together, but fell into different clades (BI = 1, ML = 100) within the Zapodidae taxon (Fig. 3). It was clear that the Zapodidae arose some time earlier than the Dipodidae and had a long evolutionary history, which was in line with previous conclusions (Wang, 1985; Zhang et al., 2013). In the Zapodidae clade, *S. concolor* was the most primitive species, which indicated that the *Sicista* gave rise to the clade that included the *Eozapus* clade. It should be noted that Dipodidae originated from Zapodidae (BI = 1, ML = 82) in keeping with a previous molecular systematic studies (Montgelard et al., 2008). The Dipodidae species were grouped in a clade (BI = 1, ML = 82) that shared a common ancestor with the Zapodidae species. Interestingly, within the Dipodidae, the Dipodinae (*D. sagitta*, *J. jaculus*, and *S. telum*) formed a sub-clade (BI = 1, ML = 100), and *E. naso* and *A. sibirica* formed another sub-clade (BI = 1, ML = 100), each with high support values. *E. naso*, which was the representative of Euchoreutinae, and *A. sibirica* clustered in a sub-clade, showed a close evolutionary relationship which could implicate a sister-group relationship between Euchoreutinae and Allactaginae (Fig. 3). Furthermore, our results indicated that *A. sibirica*, a representative of Allactaginae, occurred time earlier than the Dipodinae species (BI = 1, ML = 100) (Fig. 3). Allactaginae species appeared highly divergent from Dipodinae species, suggesting an ancient separation between them (Zhang et al., 2013). Both Allaetaginae and Dipodinae species first occurred in late Miocene; however, Allactaginae species preceded the Dipodinae, and Allactaginae taxa had broader distributions than those of the Dipodinae, expect for in Asia (Qiu and Wang, 1999; Zheng et al., 2006; Shenbrot et al., 2008). On the other hand, Dipodinae species also made major changes in their toe number (from five to three), which made Dipodinae species better

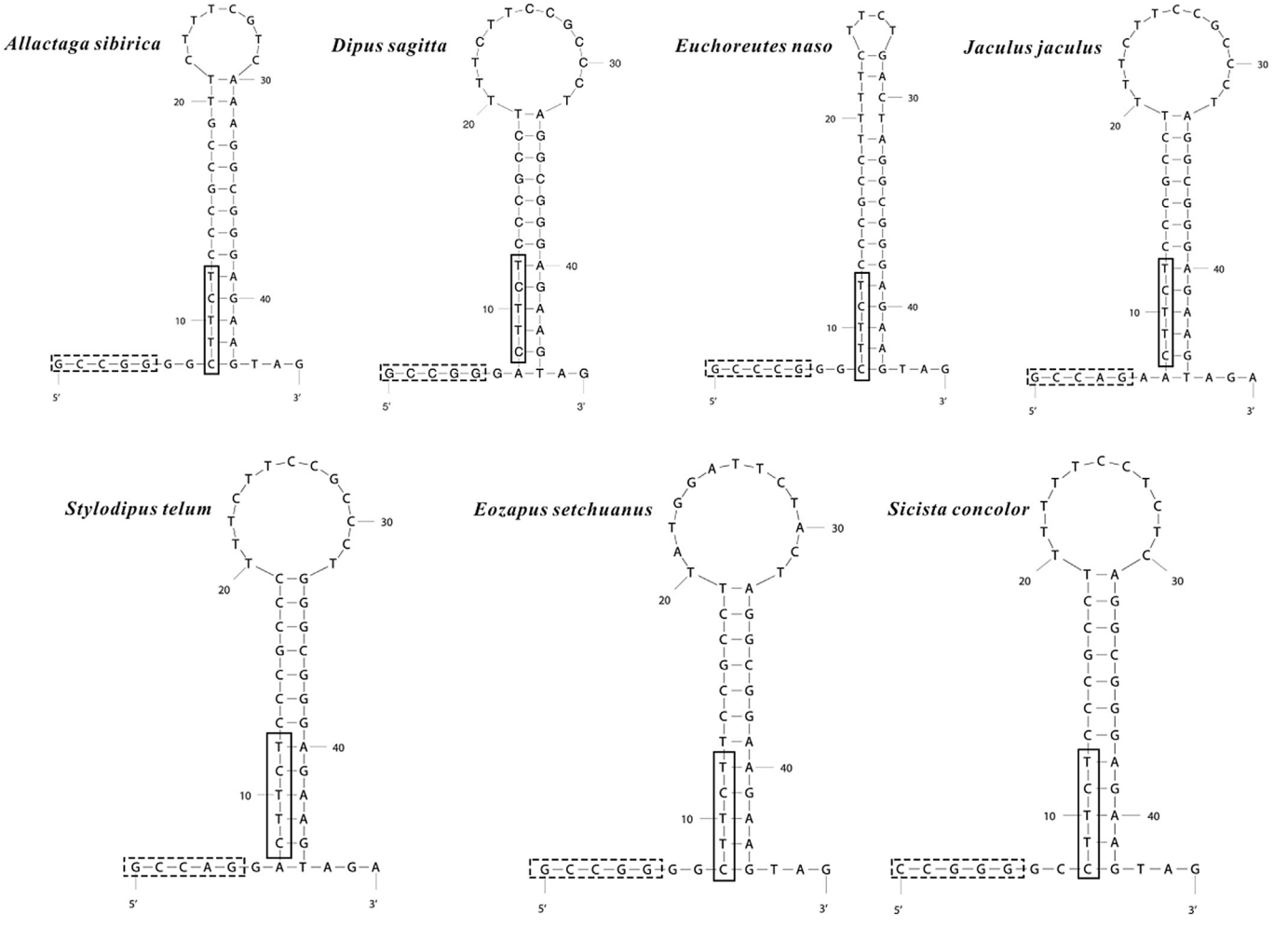

Fig. 2. Proposed stem-loop structures for the origin of replication in the mtDNA L-strand of *Allactaga sibirica* and other representative Dipodoidea species. The starting replication motif (5′-GCCGG-3′) associated with the transition of RNA synthesis to DNA synthesis was indicated by a dashed box, and the conserved stem-loop sequence (5′-CTTCT-3′) by a solid box.

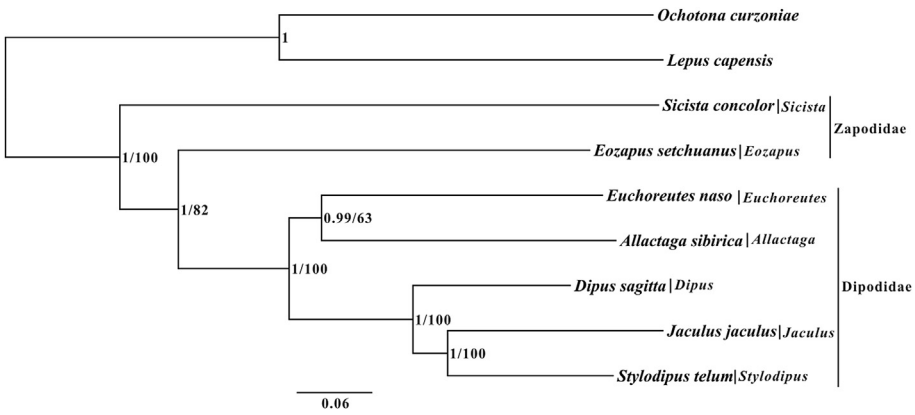

Fig. 3. Phylogenetic relationships inferred from the mitogenome sequences of Dipodoidea species. The posterior probability and bootstrap values were indicated at nodes as follows: BI/ML.

adapted to jumping in the wide habitats (Zhao, 1991). The morphological evidences also demonstrated a point at which Allactaginae species were older than Dipodinae species.

Declaration of interest

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

Acknowledgements

This 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.2016.10.004.

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. Biochim. Biophys. Acta (BBA) - Gene Regul. Mech. 1819, 992–997.
Clayton, D.A., 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. 7, 453–478.
Cummings, M.P., Otto, S.P., Wakeley, J., 1995. Sampling properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. 12, 814–822.
de Bruijn, M.H., Schreier, P.H., Eperon, I.C., Barrell, B.G., Chen, E.Y., Armstrong, P.W., Wong, J.F., Roe, B.A., 1980. A mammalian mitochondrial serine transfer RNA lacking the "dihydrouridine" loop and stem. Nucleic Acids Res. 8, 5213–5222.
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.
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 Symposium Series 41, 95–98.
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.
Holden, M., Musser, G., 2005. Family Dipodidae. Mammal Species of the World: a Taxonomic and Geographic Reference. Johns Hopkins University Press, Baltimore, pp. 871–893.
Jiang, X.H., Gao, J., Ni, L.J., Hu, J., 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.
Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. 86, 6196–6200.
Kumazawa, Y., Nishida, M., 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37, 380–398.
Liao, J., Jing, D.D., Luo, G.J., Wang, Y., Zhao, L.M., Liu, N.F., 2016. Comparative phylogeography of *Meriones meridianus, Dipus sagitta*, and *Allactaga sibirica*: potential indicators of the impact of the Qinghai-Tibetan Plateau uplift. Mamm. Bio. Z. für Säugetierkd. 81, 31–39.
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. tRNAscan-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.
Montgelard, C., Forty, E., Arnal, V., Matthee, C.A., 2008. Suprafamilial relationships among Rodentia and the phylogenetic effect of removing fast-evolving nucleotides in mitochondrial, exon and intron fragments. BMC Evol. Biol. 8, 1–16.
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.
Qiu, Z.D., Wang, X.M., 1999. Small mammal faunas and their ages in Miocene of central Nei Mongol (Inner Mongolia). Vertebr. Palasiat. 37, 120–139.
Robins, J.H., Hingston, M., Matisoo-Smith, E., Ross, H.A., 2007. Identifying *Rattus* species using mitochondrial DNA. Mol. Ecol. Notes 7, 717–729.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, 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.
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.
Shenbrot, G.I., Sokolov, V.E., Heptner, V.G., 2008. Jerboas: Mammals of Russia and Adjacent Regions. Science Publishers, Enfield, Jersey & Plymouth.
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.
Tan, B.J., 1992. A Systematic List of the Mammals. China Medical Science Press, Beijing, pp. 304–306.
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., 2014. Molecular phylogenetic relationships among Asiatic shrewlike moles inferred from the complete mitogenomes. J. Zool. Syst. Evol. Res. 53, 155–160.
Vinogradov, B.S., 1925. On the structure of the external Genitalia in Dipodidae and Zapodidae (Rodentia) as a classificatory character. J. Zool. 95, 577–585.
Wang, B.Y., 1985. Zapodidae (Rodentia, Mammalia) from the lower oligocene of Qujing, Yunnan, China. Mainz. Geowiss. Mittl. 14, 345–367.
Wang, X.T., 1991. Vertebrate Fauna of Gansu. Gansu Science and Technology Publishing House, Lanzhou, pp. 1115–1118.
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.
Xu, X.F., Gullberg, A., Arnason, U., 1996. The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. J. Mol. Evol. 43, 438–446.
Yang, Z.H., Rannala, B., 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. Mol. Biol. Evol. 14, 717–724.
Yue, H., Yan, C.C., Tu, F.Y., Yang, C.Z., Ma, W.Q., Fan, Z.X., Song, Z.B., Owens, J., Liu, S.Y., Zhang, X.Y., 2015. Two novel mitogenomes of Dipodidae species and phylogeny of Rodentia inferred from the complete mitogenomes. Biochem. Syst. Ecol. 60, 123–130.
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.
Zhang, J., Wang, Z.Y., 1963. Faunistic studies of mammals of the chinghai province. Acta Zool. Sin. 15, 125–127.
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.
Zhang, Q., Xia, L., Kimura, Y., Shenbrot, G., Zhang, Z.Q., Ge, D.Y., Yang, Q.S., 2013. Tracing the origin and diversification of Dipodoidea (Order: Rodentia): evidence from fossil record and molecular phylogeny. Evol. Biol. 40, 32–44.
Zhao, K.T., 1991. China's Jerboas. J. Suzhou Railw. Teach. Coll. 8, 29–36.
Zheng, S.H., Cai, B.Q., Li, Q., 2006. The plio-pleistocene small mammals from donggou section of Nihewan basin, Hebei, China. Vertebr. Palasiat. 44, 320–331.
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. Zool. Res. 31, 122–130.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.