Complete Nucleotide Sequence and Gene Organization of the Mitochondrial Genome of Common House Gecko, *Hemidactylus Frenatus*

Yan Jie, Zhou Jianli, Tian Chao, Zhou Kaiya

(Jiangsu Key Laboratory for Biodiversity and Biotechnology, School of Life Sciences, Nanjing Normal University, Nanjing 210046, China)

Abstract We determined the complete mitochondrial sequence of the common house gecko, *Hemidactylus frenatus*, by using the polymerase chain reactions (PCRs). The entire mtDNA sequence is 16 891 bp in length and contains 13 protein-coding genes, 2 rRNA genes, 22 tRNAs genes and the noncoding control region. Most genes are encoded by the H strand, except for one protein-coding and eight tRNA genes. The base composition is skewed similar to other vertebrate mtDNAs. No long intergenic spacers are found indicating a compact composition of this genome. In gene organization, the mitochondrial genome of *H. frenatus* is identical to that of typical vertebrate, that is to say three main atypical features (rearrangements, duplication of genes and of the control region) found in some squamates are not found in this genome. Except the parthenogenetic geckos, known geckos have consistent gene contents and orders. The new complete sequence can be useful to improve the future phylogeny by broadening the coverage of squamate diversity.

Key words *Hemidactylus frenatus*; Gekkonidae; Gekkota; mitochondrial genome

CLC number Q959.3 **Document code** A **Article ID**: 1001-4616(2009)04-0077-06

疣尾蜥虎线粒体基因组全序列及其基因组成

严 洁, 周建丽, 田 超, 周开亚

(江苏省生物多样性和生物技术重点实验室, 南京师范大学生命科学学院, 江苏 南京 210046)

[摘要] 测定了疣尾蜥虎(*Hemidactylus frenatus*)的线粒体基因组全序列. 序列全长为 16 891 bp, 包括 13个蛋白质编码基因、2个 rRNA 基因和 22个 tRNA 基因. 除大部分基因由重链编码外, 仅 1个蛋白编码基因和 8个 tRNA 基因由轻链编码. 碱基组成的偏好与其他脊椎动物线粒体 DNA 接近. 没有发现长的基因间隔区, 说明该基因组的结构十分紧凑. 基因组的组成与典型脊椎动物的相近, 即没有发现重排、基因或控制区的重复等在其它有鳞类动物中出现过的异常特征. 除单性生殖的壁虎外, 现有的壁虎类线粒体基因组在基因含量和顺序上是一致的. 作为蜥虎属线粒体基因组全序列的惟一代表, 该序列有望在有鳞类系统发生的推断上发挥一定的作用.

[关键词] 疣尾蜥虎, 壁虎科, 壁虎类, 线粒体基因组

Animal mitochondrial DNA (mtDNA) is a small, extrachromosomal genome, typically ~ 16 kb in size. The relative conservation in gene contents and arrangements, compact organization, abundance in animal cells, faster evolutionary rate made mtDNA genome successful in phylogenetic analyses. In the past 10 years, more and more mtDNA genomes have been sequenced. Till now, there are 65 squamate mtDNAs deposited in GenBank. Based on these known genomes, phylogenies have been reconstructed among different taxa, such as Serpentes[1,2], Amphisbaenians[3], Iguanians[4,5] and the whole Squamata[6,7], intending to resolve phylogenetic issues including origin of snakes[6], monophyly of Sauria[7], phylogenetic position of individual taxa[8,9] and so on. However, in-

Received date 2009-07-20

Foundation item Supported by the National Natural Science Foundation of China (30870286), Nanjing Normal University Innovative Team Project

Corresponding author Zhou Kaiya, professor, majored in zoology, E-mail: kyzhoun@126.com

complete sampling and the long-branch attraction (LBA) may distub the actual phylogeny[6, 10] and give out wrong results. To resolve these, sufficient sampling is one of the ways[11].

Changes on gene organization in mtDNA have been reported in almost all major lineages in animals[12]. Since rearrangements appear to be unique, generally rare events are unlikely to arise independently in separate evolutionary lineages. The comparison of animal mitochondrial gene arrangements has become a very powerful means for inferring ancient evolutionary relationships[12 13]. Several kinds of rearrangements and organization changes have been found in squamate mtDNAs[14 15], and have been used to infer evolutionary processes of the mitochondrial genome in snakes[2]. More mtDNAs will be useful for further study of genome evolution.

Hemidactylus (Gray, 1845) is a species-rich genus comprises about 80 species widely distributed in tropical and subtropical Asia, Europe, Africa and America[16]. To date, mtDNA genomic information from this genus is blank. In this study, we sequenced mtDNA genome from one of the four *Hemidactylus* species in China, *H. frenatus*, with the aim of providing information for studies both on phylogeny and mitogenome evolution.

1 Material and Methods

1.1 Sample, DNA amplification, and sequencing

The specimen of *Hemidactylus frenatus* sequenced in this study was sampled from Jianfengling, Hainan Province. Total DNA was extracted from a small quantity (20 mg) of tissues by DNeasy Tissue Kit (Qiagen). Several short mtDNA fragments were amplified using Ex-Taq DNA polymerase (Takara) and sequenced in order to design taxon-specific primers. PCRs were performed in a MJ PTC-200 thermal cycler under the profile: 5 min at 95℃ followed by 35 cycles of 95℃ for 30 s, 50~55℃ for 30 s, and 72℃ for 90 s. PCR products of 1~2.5 kb were purified and then sequenced employing an ABI 310 or 3 700 system with bi-directional and several internal primers. Short fragments were assembled into a continuous sequence.

1.2 Sequence analysis

Sequence assembling was implemented by SeqMan II in DNASTAR package. In the mtDNA sequences thus obtained, protein-coding genes were identified by searching the start and stop codons based on corresponding homologues from other geckos. The transfer RNA genes were located with tRNAscan-SE 1.21 (http://lowelab.ucsc.edu/), and secondary structures were inferred using DNASIS version 2.5 (Hitachi Engineering, Tokyo, Japan) and RNAstucture 4.6[17]. Boundaries of rRNA genes and control regions were tentatively defined by the boundaries of adjacent coding genes. The mtDNA sequence has been deposited at GenBank (GQ245970).

Comparisons with other five gekkotan mtDNAs (*Gekko gecko*: AY282753; *Gekko vittatus*: NC_008772; *Tarentola mauritanica*: EU443255; *Coleonyx variegatus*: NC_008774; *Teratoscincus keyserlingii*: NC_007008) in gene lengths and G+C content was conducted. Another gekkotan mtDNA in GenBank (*Heteronotia binoei*: NC_010292) was not added into comparison because of the relatively longer length due to multiple gene duplications. Nucleotide composition and codon usage bias were calculated using software MEGA 4.0[18].

2 Results

2.1 Genome content and organization

The mitochondrial genome presented here for *H. frenatus* is 16 891 bp in length. This length is within the range reported for other squamate mt genomes[19]. This new sequence includes 2 rRNA genes, 13 protein-coding genes, 22 tRNA genes and the noncoding control region. Position 1 of the sequence corresponds to the first nucleotide acid in the $tRNA^{Phe}$ gene. The gene organization conformed to that of typical vertebrate (Fig. 1, Table 1). Most genes are encoded on the H strand, except for the ND6 gene and eight tRNA genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser(UCN)}$, $tRNA^{Pro}$, $tRNA^{Glu}$), which are encoded on the L strand. Eight overlapping regions and sixteen intergenic spacers are found and shown in Table 1. The base compositions in these mtDNAs are skewed similarly to other vertebrate mtDNAs.

Table 1 Characterisitics of the mitochondrial genome of *Hemidactylus frenatus*

表 1 疣尾蜥虎线粒体基因组的结构特征

Gene[a]	Position		Size		Codons		Spacer(+)
	From	To	bp	aa	start	stop[b]	Overlap(−)
tRNA^{Phe}	1	75	75				
12sRNA	76	1 033	958				
tRNA^{Val}	1034	1 102	69				
16sRNA	1 103	2 641	1 539				
tRNA$^{Leu(UUR)}$	2642	2 716	75				
NADH 1	2717	3 670	954	317	GTG	TAA	7
tRNA^{Ile}	3678	3 748	71				1
tRNA^{Gln}(L)	3750	3 821	72				−1
tRNA^{Met}	3821	3 890	70				
NADH 2	3891	4 929	1 039	346	ATA	Taa	
tRNA^{Trp}	4930	4 997	68				−2
tRNA^{Ala}(L)	4996	5 064	69				1
tRNA^{Asn}(L)	5066	5 138	73				
Origin-L	5139	5 167	29				−3
tRNA^{Cys}(L)	5165	5 230	66				4
tRNA^{Tyr}(L)	5235	5 300	66				4
COI	5305	6 852	1 548	515	ATG	AGA	−7
tRNA$^{Ser(UCN)}$(L)	6846	6 916	71				3
tRNA^{Asp}	6920	6 987	68				1
COII	6989	7 676	6 88	229	ATG	Taa	
tRNA^{Lys}	7677	7 743	67				
ATPase8	7744	7 908	165	54	GTG	TAG	−10
ATPase6	7899	8 578	680	226	ATG	TAa	
COIII	8579	9 362	784	261	ATG	TAa	
tRNA^{Gly}	9363	9 431	69				2
NADH 3	9434	9 780	347		ATG	TAa	1
tRNA^{Arg}	9782	9 849	68				2
NADH 4L	9852	10 145	294	97	ATA	TAA	−7
NADH 4	10139	11 506	1 368	455	ATG	TAA	5
tRNA^{His}	11512	11 579	68				
tRNA$^{Ser(AGY)}$	11580	11 639	60				−1
tRNA$^{Leu(CUN)}$	11639	11 711	73				2
NADH 5	11714	13 522	1 809	602	ATG	AGA	−19
NADH 6(L)	13504	14 025	522	173	ATG	AGG	1
tRNA^{Glu}(L)	14027	14 098	72				2
Cyt b	14101	15 234	1 134	377	ATG	TAA	5
tRNA^{Thr}	15240	15 306	67				10
tRNA^{Pro}(L)	15317	15 382	66				
D-loop	15383	16 891	1 509				

[a] (L) indicates a gene encoded on the L-strand [b] TA a and Taa represent incomplete stop codons

2.2 Ribosomal, transfer RNA and protein-coding genes

The 12S and 16S rRNA genes of *H. frenatus* are 958 and 1 539 bp, respectively. The size is similar to those of other determined mtDNAs from Gekkonidae (Table 2). They are located between the tRNA^{Phe} and tRNA$^{Leu(UUR)}$ genes, and separated by the tRNA^{Val} gene with the same situation found in other geckos[7].

Total of the 22 tRNA genes are interspersed in the genome and range in size from 60 to 75 bp. All could be folded into the canonical cloverleaf secondary structure (such as Fig. 2b), except tRNA$^{Ser(AGY)}$ in which the complete dihydrouridine arm (D-arm) was lacking (Fig. 2a). This is a common feature of metazoan mtDNAs[20], and the aberrant tRNA can also fit the ribosome by adjusting its structural conformation and function in a similar way to that of usual tRNAs in the ribosome[21].

Table 2　General characteristics of six gekkotan mitochondrial genomes

表 2　壁虎类动物线粒体基因组的基本特征

Taxa	Genome Length/bp					G+C nucleotide content/%				
	Total	Protein coding	tRNAs	RNAs	Control region	Total	Protein coding	tRNAs	RNAs	Control region
Hemidactylus frenatus	16 891	11332	2 497	1 523	1 509	45.8	45.9	47.6	45.1	43.8
Gekko gecko	16 435	11 335	2 547	1 547	1 014	41.2	41.1	42.4	41.7	37.5
Gekko vittatus	16 946	11 319	2 539	1 529	1 554	37.5	36.7	41.4	40.3	34.7
Tarentola mauritanica	16 593	11 282	2 492	1 541	1 180	44.7	45.6	47.8	44.7	29.2
Coleonyx variegatus	17 110	11 365	2 471	1 544	1 727	43.8	44.0	44.4	42.5	42.3
Teratoscincus keyserlingii	17 199	11 101	2 472	1 525	2 117	44.1	44.1	46.4	43.2	41.8

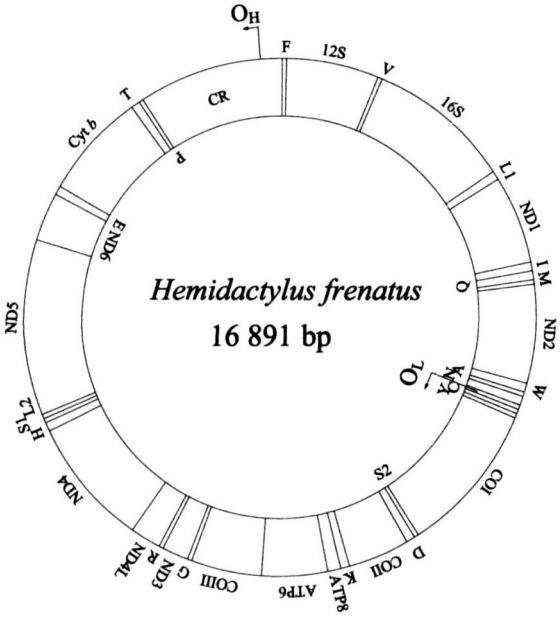

Fig.1　Complete mitochondrial (mt)DNA organization of *Hemidactylus frenatus*. Transfer RNA genes were identified by a single letter of the amino acid code. O_H and O_L represent the replication origins of H-strand and L-strand

图 1　疣尾蜥虎线粒体 DNA 基因组结构简图(tRNA 基因由单个字母的氨基酸代码表示,O_H 和 O_L 为直链和轻链复制起点)

Fig.2　Proposed secondary structure of the $tRNA^{Ser(AGY)}$ (a) and $tRNA^{Val}$ genes (b) in *H. frenatus* mitochondrial genome. a, structure lacking the complete dihydrouridine arm (D-arm); b, representative of the standard cloverleaf secondary structure

图 2　根据序列推测的 $tRNA^{Ser(AGY)}$(a)和 $tRNA^{Val}$ 基因(b)的二级结构. a: 缺失二氢尿嘧啶环所在的 D 臂, b: 代表标准的三叶草结构

Positions of protein-coding genes were determined by the occurrence of start and stop codons and by analogy with other complete squamate mtDNA sequences. In the 13 protein-coding genes, nine of them use ATG as the start codon, while ND1 and ATP8 use GTG (Table 1). ND2 and ND4L genes start with ATA. Five protein-coding genes—ND2, COII, ATP6, COIII and ND3—have incomplete stop codons (TA or T), which were presumably completed as TAA by post-transcriptional polyadenylation[22]. The TAG and AGR are observed as terminator in other genes (COI, ATP8, ND5 and ND6). Codon usage in the *H. frenatus* mitochondrial genome indicated that CUA (Leu), ACA (Thr) and GCC (Ala) occur at a higher frequency than others (data not shown). The lowest proportion (10.1%) of G at the third codon position, represent a typical bias against the use of 'G' in vertebrates. Two pairs of protein-coding genes (ATP8 and ATP6, ND4L and ND4), both located on the H-strand, overlapped in reading-frames for 10 and 7 bp.

2.3　Noncoding regions

The noncoding regions of *H. frenatus* include a control region, a putative light-strand origin of replication (O_L) and a few intergenic spacers.

The single control region spans 1 509 bp and is located between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes. Within this sequence, conserved sequence blocks (CSB1-3) were identified in the positions 963-983, 1 234-1 251, and

1 287–1 306, respectively. These conserved sequences were thought to be involved in positioning RNA polymerase both for transcription and for priming replications[23 24]. Comparing to other three geckos, the C+G content of control region in *H. frenatus* is higher (Table 2).

As in most vertebrates, the putative light-strand origin of replication (O_L) of *H. frenatus* comprises 29 bp and located between $tRNA^{Asn}$ and $tRNA^{Cys}$ in W<u>ANCY</u> cluster ($tRNA^{Trp}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$ and $tRNA^{Tyr}$). It has potential to fold into a stable stem-loop secondary structure with 9 bp in the stem and 11 bp in the loop. This region overlaps the $tRNA^{Cys}$ gene by 3 bp. A similar sequence (5′–CCCGG–3′) to the 5′–GCCGG–3′ conserved motif, which was thought to be associated with the transition from RNA synthesis to DNA synthesis in human mtDNA[25], is found at the base of the stem of $tRNA^{Cys}$.

The intergenic spacers found in *H. frenatus* is short (≤10 bp), indicating a compact structure.

3 Discussion

In *H. frenatus* mtDNA, identical gene organization to typical vertebrate is found. Three main atypical features found in some squamate reptiles, gene rearrangements in amphisbaenian reptiles[3] and snakes[1 2], duplicated genes in *Heteronotia* geckos[26] and cuplicated control regions in advanced snakes[1 2 27] and the Komodo dragon[28], are not observed in this new genome. In all examined gekkotan mitochondrial genomes to date, gene contents and orders are identical, except the parthenogenetic geckos, *Heteronotia binoei* complex[26] in which random duplications varying from 5.7 to 9.4 kb were found. In addition, the light-strand origin of replication (O_L), which is lost in many squamate reptiles[14], can be found in all these genomes.

Within Squamata, phylogenetic relationships are not well resolved. Several mitochondrial and nuclear genes have been involved[10 29 30], while controversies continue. The complete mitochondrial genomes, although powerful, are limited in amounts and covering only a few squamate taxa. Besides, differences in evolutionary rates of individual lineages increase the risk of mistake. As discussed above, insufficient sampling and the long-branch attraction have been reported possibly relative to unexpected and wrong relationships, and thorough taxon coverage of squamate diversity is necessary[31]. In this study, the first mitochondrial genome from *Hemidactylus* (Gekkonidae, Gekkota) and meanwhile the seventh in gekkotans was reported. Though distinction in gene organization, which may provide important evolutionary information, was not found in this new genome, sequence information could be well used in future phylogenetic discussions at different levels.

[References]

[1] Dong S, Kumazawa Y. Complete mitochondrial DNA sequences of six snakes: phylogenetic relationships and molecular evolution of genomic features[J]. Journal of Molecular Evolution, 2005, 61(1): 12–22.

[2] Yan J, Li H, Zhou K. Evolution of the mitochondrial genome in snakes: Gene rearrangements and phylogenetic relationships [J]. BMC Genomics, 2008, 9(1): 569.

[3] Macey J R, Papenfuss T J, Kuehl J V, et al. Phylogenetic relationships among amphisbaenian reptiles based on complete mitochondrial genomic sequences[J]. Molecular Phylogenetics and Evolution, 2004, 33(1): 22–31.

[4] Macey J R, Schulte J A, Fong J J, et al. The complete mitochondrial genome of an agamid lizard from the Afro-Asian subfamily agaminae and the phylogenetic position of *Bufoniceps* and *Xenagama* [J]. Molecular Phylogenetics and Evolution, 2006, 39(3): 881–886.

[5] Macey J R, Kuehl J V, Larson A, et al. Socotra Island the forgotten fragment of Gondwana: unmasking chameleon lizard history with complete mitochondrial genomic data[J]. Molecular Phylogenetics and Evolution, 2008, 49(3): 1 015–1 018.

[6] Kumazawa Y. Mitochondrial DNA sequences of five squamates: phylogenetic affiliation of snakes[J]. DNA Research, 2004, 11(2): 137–144.

[7] Zhou K, Li H, Han D, et al. The complete mitochondrial genome of *Gekko gecko* (Reptilia: Gekkonidae) and support for the monophyly of Sauria including Amphisbaenia[J]. Molecular Phylogenetics and Evolution, 2006, 40(3): 887–892.

[8] Macey J R, Fong J J, Kuehl J V, et al. The complete mitochondrial genome of a gecko and the phylogenetic position of the

Middle Eastern *Teratoscincus keyserlingii*[J]. Molecular Phylogenetics and Evolution, 2005, 36(1): 188-193.
[9] Bohme M U, Fritzsch G, Tippmann A, et al. The complete mitochondrial genome of the green lizard *Lacerta viridis viridis* (Reptilia: Lacertidae) and its phylogenetic position within squamate reptiles[J]. Gene, 2007, 394(1-2): 69-77.
[10] Townsend T M, Larson A, Louis E, et al. Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree[J]. Systematic Biology, 2004, 53(5): 735-757.
[11] Page R D M, Holmes E C. Molecular Evolution, a Phylogenetic Approach[M]. Oxford: Blackwell Science, 1998: 1-352.
[12] Boore J L. Animal mitochondrial genomes[J]. Nucleic Acids Research, 1999, 27(8): 1767-1780.
[13] Boore J L, Brown W M. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool[J]. Current Opinion in Genetics and Development, 1998, 8(6): 668-674.
[14] Macey J R, Schulte J A, Larson A. Evolution and phylogenetic information content of mitochondrial genomic structural features illustrated with acrodont lizards[J]. Systematic Biology, 2000, 49(2): 257-277.
[15] Amer S A, Kumazawa Y. Mitochondrial genome of *Pogona vitticepes* (Reptilia: Agamidae): control region duplication and the origin of Australasian agamids[J]. Gene, 2005, 346: 249-256.
[16] Carranza S, Arnold E N. Systematics, biogeography, and evolution of *Hemidactylus* geckos (Reptilia: Gekkonidae) elucidated using mitochondrial DNA sequences[J]. Molecular Phylogenetics and Evolution, 2006, 38(2): 531-545.
[17] Mathews D H, Disney M D, Childs J L, et al. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure[J]. Proceedings of the National Academy of Sciences, 2004, 101(19): 7287-7292.
[18] Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0[J]. Molecular Biology and Evolution, 2007, 24(8): 1596-1599.
[19] Kumazawa Y. Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations[J]. Gene, 2007, 388(1-2): 19-26.
[20] Wolstenholme D R. Animal mitochondrial DNA: structure and evolution[J]. International Review of Cytology, 1992, 141: 173-216.
[21] Ohtsuki T, Kawai G, Watanabe K. The minimal tRNA: unique structure of *Ascaris suum* mitochondrial tRNA(Ser)(UCU) having a short T arm and lacking the entire D arm[J]. FEBS Letters, 2002, 514(1): 37-43.
[22] Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria[J]. Nature, 1981, 290(5806): 470-474.
[23] Clayton D A. Replication and transcription of vertebrate mitochondrial DNA[J]. Annual Review of Cell Biology, 1991, 7(1): 453-478.
[24] Shadel G S, Clayton D A. Mitochondrial DNA maintenance in vertebrates[J]. Annual Review of Biochemistry, 1997, 66(1): 409-435.
[25] Hixson J E, Wong T W, Clayton D A. 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]. Journal of Biological Chemistry, 1986, 261(5): 2384-2390.
[26] Fujita M K, Boore J L, Moritz C. Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (*Heteronotia binoei*; Squamata, Gekkonidae)[J]. Molecular Biology and Evolution, 2007, 24(12): 2775-2786.
[27] Kumazawa Y, Ota H, Nishida M, et al. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster[J]. Molecular Biology and Evolution, 1996, 13(9): 1242-1254.
[28] Kumazawa Y, Endo H. Mitochondrial genome of the Komodo dragon: efficient sequencing method with reptile-oriented primers and novel gene rearrangements[J]. DNA Research, 2004, 11(2): 115-125.
[29] Vidal N, Hedges S B. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes[J]. Comptes Rendus Biologies, 2005, 328(10/11): 1000-1008.
[30] Vidal N, Hedges S B. The molecular evolutionary tree of lizards, snakes, and amphisbaenians[J]. Comptes Rendus Biologies, 2009, 332(2/3): 129-139.
[31] Albert E M, San Mauro D, Garcia-Paris M, et al. Effect of taxon sampling on recovering the phylogeny of squamate reptiles based on complete mitochondrial genome and nuclear gene sequence data[J]. Gene, 2009, 441(1/2): 12-21.

[责任编辑: 孙德泉]