Mitochondrial DNA, February 2010; 21(1): 8–18

informa
healthcare

The complete mitochondrial genome of the Sichuan hot-spring keel-back (*Thermophis zhaoermii*; Serpentes: Colubridae) and a mitogenomic phylogeny of the snakes

MIAO HE[1,2], JINCHAO FENG[3], & ERMI ZHAO[1,4]

[1]*Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, School of Life Sciences, Sichuan University, Chengdu, P. R. China,* [2]*Institute of Blood Transfusion, The Chinese Academy of Medical Sciences, Chengdu, P. R. China,* [3]*College of Life and Environmental Science, The Central University for Nationalities, Beijing, P. R. China, and* [4]*Chengdu Institute of Biology, The Chinese Academy of Sciences, Chengdu, P. R. China*

(*Received 25 March 2009; revised 11 November 2009; accepted 23 November 2009*)

Abstract

The complete mitochondrial genome sequence of the Sichuan hot-spring keel-back (*Thermophis zhaoermii*) was determined in the present study. The genome is 17,322 bp in size, containing 2 ribosomal RNA (rRNA) genes, 13 protein-coding genes, 22 transfer RNA (tRNA) genes, and 2 control regions, similar to other alethinophidian snakes. A special 40 bp non-coding region, which was highly homologous to the start of control regions I (CR I) and II (CR II), containing a 16 bp C-rich segment, was identified upstream of the pseudo-tRNA^{Pro} gene that had been observed across Colubridae and Homalopsidae. Twelve concatenated heavy-strand encoded protein-coding genes were used for phylogenetic reconstruction employing Bayesian and maximum likelihood inference. Both analyses yielded identical topologies, demonstrating that *T. zhaoermii* can solidly be placed within Colubridae as a sister group to Colubrinae. The paraphyly of Scolecophidia and monophyly of Henophidia and Caenophidia were also supported. A relaxed clock molecular divergence time analysis was carried out to estimate the temporal origin of each clade. Our results indicate that the Alethinophidia began to diverge from the paraphyletic Scolecophidia approximately 130 million years ago in the early Cretaceous; the divergence of living alethinophidian snakes, the radiation of the Caenophidia, and the separation between *Acrochordus* and the Colubroidea might have been caused by the K/T event.

Keywords: Thermophis zhaoermii, *hot-spring keel-back, mitochondrial genome, phylogeny*

Introduction

The hot-spring keel-back (*Thermophis*), which attains the highest elevation (over 4000 m) of any known snake species, is commonly found near hot springs in the Qinghai–Xizang Plateau and was first described by Wall (1907). Although the genus has a broad distribution, ranging from Tibet to western Sichuan, the inclement climate on the plateau could restrict the snakes of this genus to inhabit only places in the proximity of hot springs. Currently, there are only a few localities where *Thermophis* snakes have been reported (Guo et al. 2008; He et al. 2009; Huang et al. 2009).

For nearly a century, only one species in the genus *Thermophis* (*Thermophis baileyi*) was recognized and considered to be restricted to the Lhasa area (Malnate 1953), until Guo et al. (2008) claimed that the hot-spring keel-back, which was found in Litang, Sichuan Province, so far away from Lhasa area, should be elevated to species rank as *Thermophis zhaoermii*. The validity of *T. zhaoermii* was supported by both morphological analysis based on subcaudal scales, maxillary teeth, and hemipenial characters (Guo et al. 2008) and molecular phylogeny based on the mitochondrial cytochrome *b* (*CYTB*) and nuclear oocyte maturation factor (c-mos) genes

Correspondence: E. Zhao, Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, School of Life Sciences, Sichuan University, Chengdu 610064, P. R. China. Tel: 86 28 85249806. Fax: 86 28 85222753. E-mail: zhaoermi@cib.ac.cn

ISSN 1940-1736 print/ISSN 1940-1744 online © 2010 Informa UK Ltd.
DOI: 10.3109/19401730903505867

(He et al. 2009), and on the *CYTB* and mitochondrial *ND4* genes (Huang et al. 2009). All these studies have successfully placed the genus *Thermophis* into Xenodontinae, whose members are only distributed in the New World (Guo et al. 2008; He et al. 2009; Huang et al. 2009).

Mitochondrial genome sequences have been widely used in molecular phylogenetic analyses of snakes (Kumazawa et al. 1998; Dong and Kumazawa 2005; Douglas et al. 2006; Jiang et al. 2007; Castoe et al. 2008; Yan et al. 2008) since the first complete mitochondrial DNA (mtDNA) sequence of *Dinodon semicarinatus* was reported in 1998 (Kumazawa et al. 1998). Other than universal mtDNA structures, many unique phenomena were discovered in the mitochondrial genomes of snakes, such as the duplication of the control region and its adjacent area in alethinophidian snakes, length reduction of all genes (Kumazawa et al. 1998; Dong and Kumazawa 2005; Jiang et al. 2007; Yan et al. 2008), and a burst of unique amino acids substitutions (Castoe et al. 2008). Here, we present the complete nucleotide sequence of the *T. zhaoermii* mitochondrial genome, clarify its genome structures, and further discuss the phylogenetic relationships of the snakes based on protein-coding sequences.

Materials and methods

Taxon sampling and DNA extraction

Specimens of *T. zhaoermii* were collected from Litang, Sichuan Province, China. Total genomic DNA was extracted from muscle tissue using the standard phenol/chloroform protocol of Sambrook et al. (1989). Table I presents all the taxa used in the present study, voucher numbers, and GenBank accession numbers.

DNA amplification and sequencing

Thirteen pairs of primers were employed to amplify contiguous, overlapping segments covering the complete mitochondrial genome of *T. zhaoermii*. The primers used in this study are presented in Table II.

The PCR was performed with *Taq* DNA Polymerase (Promega, Madison, WI, USA) in a total volume of 45 μl, with the following reaction parameters: predenaturation at 94°C for 4 min, followed by 35 cycles of 40 s denaturation (95°C), 30 s annealing (55–68°C), and a 1 min extension (72°C) with a final extension at 72°C for 7 min. PCR cycling was performed on an MJ PTC-100 thermal cycler (MJ Research, Waltham, MA, USA). The PCR products were purified using the DNA Agarose Extraction kit (Omega, Doraville, GA, USA) according to the manufacturer's instructions. Purified products were used as templates in cycle-sequencing reactions, and the resulting sequences were directly read with an ABI 3730 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA). Sequences were determined for both directions and deposited on GenBank (Table I for accession numbers).

Table I. Complete mtDNA sequences used in the present study.

Species	GenBank accession number	Source
A. meiguensis	NC_011576	Wang et al. (2009)
A. granulatus	NC_007400	Dong and Kumazawa (2005)
A. piscivorus	EF669477	Jiang et al. (2007)
Boa constrictor	NC_007398	Dong and Kumazawa (2005)
Bungarus fasciatus	NC_011393	Chen and Zhao (2009)
Bungarus multicinctus	NC_011392	Chen and Zhao (2009)
Cylindrophis ruffus	NC_007401	Dong and Kumazawa (2005)
D. russellii	NC_011391	Chen and Zhao (2009)
D. acutus	NC_010223	Yan et al. (2008)
D. semicarinatus	NC_001945	Kumazawa et al. (1998)
E. plumbea	NC_010200	Yan et al. (2008)
Eunectes notaeus	AM236347	Douglas et al. (2006)
G. blomhoffi brevicaudus	NC_011390	Chen and Zhao (2009)
L. dulcis	NC_005961	Kumazawa et al. (2004)
Naja atra	NC_011389	Chen and Zhao (2009)
Naja naja	NC_010225	Yan et al. (2008)
O. hannah	NC_011394	Chen and Zhao (2009)
O. okinavensis	NC_007397	Dong and Kumazawa (2005)
P. guttatus guttatus	AM236349	Douglas et al. (2006)
P. slowinskii	NC_009769	Jiang et al. (2007)
Python regius	NC_007399	Dong and Kumazawa (2005)
R. australis	AM236345	Douglas et al. (2006)
R. braminus	NC_010196	Yan et al. (2008)
T. zhaoermii	GQ166168	Present study
T. reticulatus	NC_010971	Castoe et al. (2008)
T. mirus	AM236345	Douglas et al. (2006)
Xenopeltis unicolor	NC_007402	Dong and Kumazawa (2005)
Iguana iguana (out-group)	NC_002793	Janke et al. (2001)

Sequence analysis

DNA sequences were analyzed using the software MEGA 4 (Tamura et al. 2007). The locations of protein-coding and rRNA genes were identified by comparison with known sequences as shown in Table I. The transfer RNA (tRNA) genes were identified by tRNAscan-SE 1.21 (Lowe and Eddy 1997). The cloverleaf secondary structure and anticodon sequences were determined using RNAstructure 4.6 (http://rna.urmc.rochester.edu/rnastructure.html).

Phylogenetic analysis

Twelve H-strand protein-coding gene sequences were aligned using Clustal X 1.83 (Thompson et al. 1997) with default settings. The *ND6* gene, which is encoded by the L-strand, was excluded from the phylogenetic analysis because of its strand-specific base composition bias (Yan et al. 2008). Aligned nucleotide sequences were translated into amino acid sequences in order to search for premature stop codons and to

Table II. Primers used for sequencing the *T. zhaoermii* mitochondrial genome.

Primer name (forward/reverse)	Forward primer sequence (5′ → 3′)	Reverse primer sequence (5′ → 3′)	Length (bp)
TB01F/TB01R	TAATACCCGTGCCAGCGAC	CTTCACAGGGTCTTCTGGTCTT	1757
TZ02F/TZ02R	CACTATAAAAGGAACTCGGCAACC	GATTAGGRCCTTTTCGGTGTTGTAT	870
16Sb/H3518	ACGTGATCTGAGTTCAGACCGG	CCGTGTCTACTCTATCAAGGTAGTCC	1183
TB_nd1_f/Met_R	GGCACAACGAACCCACAACTATT	CGTGTGGCATTTTCGGGGTAT	1561
TZ_leu2_f/TZ_nd2_R	AAGGCTTAAAACCTCAACACAGATG	GAATAAGTCATTTGGGTATAAAGCCTG	979
TB04F/TB04R	CCCAACTACGAAAACTAATAGCCTTYTCATC	GGATGATTCCAGTTATTCCGCCGAC	1949
TB05F/TB05R	TTATTCCAACACCTATTCTGATTTTTCG	ATTGCTCAGGAGTGGAGGACATCTTCTGC	1535
TZ06F/TZ06R	CCGCAACAATAGCATCCCTAACCC	GACGATAGTGGTTATTAGGGCATTGGC	1958
TZ07F/TZ07R	CAGAGGTCCCCCTACTCAACACAGC	TGGGGGCTTCTACGTGGGCTTT	1810
TB08F/TB08R	CTAAAYAACACATCAAACACCCCAAC	GAGTGKARTAGGGCTGATACWGGTGT	1775
TB09F/TB09R	TCTCCACCTGATACATAAGCACCGACCC	ATGGTTCAACCGTAGGGGACATCTCG	2247
L14910/H16064	GACCTGTGATMTGAAAAACCAYCGTTGT	CTTTGGTTTACAAGAACAATGCTTTA	1203
TB_cytb/All_12s	TATCCATCACTATTCTACTCACAGCAC	GGCTAGGCATAGTAAGGTATCTAAT	1947

confirm the absence of pseudo-genes from our data. Gaps in alignments were treated as missing data in the subsequent phylogenetic analyses.

The best-fitting model of sequence evolution for the Bayesian inference (BI) was obtained using MrModeltest 2.2 (Nylander 2004) under the Akaike information criterion (Akaike 1974) and was the general time-reversible model (GTR, Lanave et al. 1984) with a proportion of invariant sites (I) and among-site rate heterogeneity modeled by the Gamma distribution (G) (Yang 1994).

BI was carried out in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Posterior distributions of parameters and topologies were obtained using the Markov chain Monte Carlo (MCMC) sampling method with one cold and three heated chains for 2,000,000 generations, with every 100th sample being retained. The first 10,000 trees were discarded as burn-in and the remaining samples were used to generate a 50% majority-rule consensus tree. All MCMC runs were repeated twice to confirm consistent approximation of the posterior parameter distributions.

Maximum likelihood (ML) analysis was performed using RAxML as implemented on the CIPRES portal 1.14 (http://www.phylo.org) (Stamatakis et al. 2008). Node support was evaluated with 1000 rapid bootstrap pseudo-replicates with default parameters.

Estimation of divergence times

We used the BEAST 1.4.8 (Drummond and Rambaut 2007) to infer the temporal origin of each lineage under a relaxed molecular clock assumption with an uncorrelated log-normal model. We used a Yule tree prior, which assumes a constant speciation rate per lineage for species-level phylogenies. We constrained three nodes by extracting information from the fossil record to calibrate the phylogenetic tree: (1) "Colubrinae," as represented by the fossil *Texasophis galbreathi* which appeared around 28.5 million years ago (Mya) (Holman 2000), we used a log-normal prior of 28.5 Mya as zero offset, the default log-normal mean and SD of 0 and 1, respectively, to constrain this node; (2) "Viperinae", suggested to have emerged around 23.8 Mya (see Viperinae indet. in Szyndlar and Rage 2002), we used a log-normal prior of 23.8 Mya as zero offset and a log-normal mean and SD of 1 for this node; and (3) "Colubroidea," following Wüster et al.'s (2008) suggestion of a log-normal prior with a zero offset of 40 Mya, a log-normal mean and SD of 2 and 1.2, respectively, corresponding to a 95% CI of 40–95 Mya, thus covering the potential divergence dates suggested by Rage and Werner (1999), Rage et al. (2003), and Head et al. (2005). The GTR + I + G model was used for 5,000,000 MCMC iterations with every 1000th sample being retained. The first convergence of the chains to the stationary distribution was

checked by eye using a Tracer 1.4 (Rambaut and Drummond 2007) and the first 1000 steps were discarded as burn-in.

Results and discussion

Structure of the mitochondrial genome

The complete mitochondrial genome sequence of *T. zhaoermii* is reported first in the present study. The total length of the *T. zhaoermii* mtDNA sequence is 17,322 bp. It consists of 13 protein-coding genes, 2 rRNA genes (12S and 16S rRNA), 22 tRNA genes, and 2 control regions (CR I and CR II) (Table III), which is similar to other reported alethinophidian snakes (Figure 1). Most *T. zhaoermii* mitochondrial genes are encoded on the H-strand, except for the *ND6* gene and eight tRNA genes ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser(UCN)}$, $tRNA^{Glu}$, and $tRNA^{Pro}$), which are encoded on the L-strand. CR I is located between a $tRNA^{pro}$ gene and a $tRNA^{Phe}$ gene, while CR II is found between a pseudo-$tRNA^{Pro}$ and a $tRNA^{Leu\ (UUR)}$.

Sequence features of protein-coding genes

Detailed information on the 13 protein-coding genes is presented in Table III. There are no introns in the protein-coding genes—as seen in other vertebrates. ATG acts mainly as the initiation codon in nine protein-coding genes (*COXII*, *ATP8*, *ATP6*, *COXIII*, *ND4L*, *ND4*, *ND5*, *ND6*, and *CYTB*), while *ND1*, *ND2*, and *ND3* use ATA and *COXI* uses GTG as initiation codons. As for stop codons, seven protein-coding genes end with the complete stop codon TAA (*ATP8*, *ATP6*, *ND4L*, and *ND5*), AGG (*COXI* and

Table III. Detailed genome annotation of *T. zhaoermii*.

Gene	Start position	Stop position	Length (bp)	Anticodon	Start codon	Stop codon	Strand
Phe	1	62	62	GAA			H
12S rRNA	63	987	925				–
Val	988	1052	65	TAC			H
16S rRNA	1053	2531	1479				–
ND1	2532	3495	964		ATA	T++	H
Ile	3496	3560	65	GAT			H
Pseudo-Pro	3601	3626	26				–
CRII	3627	4649	1023				–
Leu (UUR)	4650	4722	73	TAA			H
Gln	4723	4794	73	TTG			L
Met	4794	4857	64	CAT			H
ND2	4858	5887	1030		ATA	T++	H
Trp	5888	5952	65	TCA			H
Ala	5955	6019	65	TGC			L
Asn	6020	6093	74	GTT			L
O_L	6095	6131	37				–
Cys	6129	6187	59	GCA			L
Tyr	6188	6251	64	GTA			L
COXI	6252	7853	1602		GTG	AGG	H
Ser (UCN)	7844	7910	67	TGA			L
Asp	7911	7976	66	GTC			H
COXII	7977	8661	685		ATG	T++	H
Lys	8662	8724	63	TTT			H
ATP8	8726	8890	165		ATG	TAA	H
ATP6	8881	9561	681		ATG	TAA	H
COXIII	9561	10,344	784		ATG	T++	H
Gly	10,345	10,405	61	TCC			H
ND3	10,406	10,748	343		ATA	T++	H
Arg	10,749	10,813	65	TCG			H
ND4L	10,814	11,104	291		ATG	TAA	H
ND4	11,104	12,441	1338		ATG	AGA	H
His	12,443	12,508	66	GTG			H
Ser (AGY)	12,510	12,566	57	GCT			H
Leu (CUN)	12,564	12,634	71	TAG			H
ND5	12,637	14,412	1776		ATG	TAA	H
ND6	14,408	14,923	516		ATG	AGG	L
Glu	14,924	14,986	63	TTC			H
CYTB	14,985	16,101	1117		ATG	T++	H
Thr	16,102	16,166	65	TGT			H
Pro	16,167	16,229	63	TGG			L
CRI	16,230	17,322	1093				–

Locus names in roman are tRNAs.

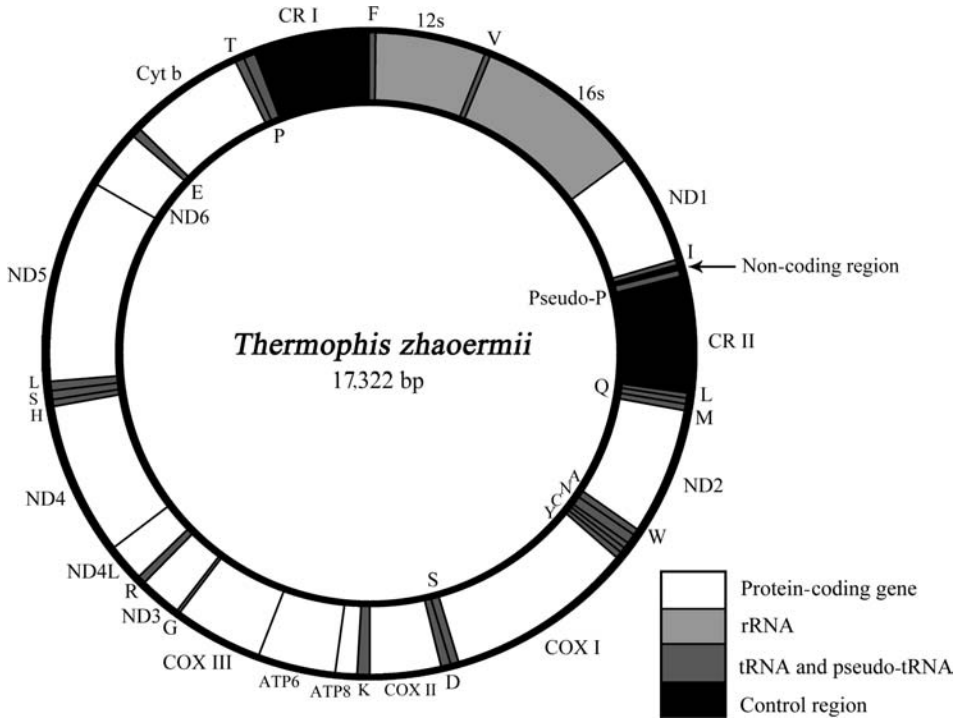

Figure 1. *T. zhaoermii* mitochondrial genome structure.

ND6), and AGA (*ND4*), whereas six protein-coding genes (*ND1*, *ND2*, *COXII*, *COXIII*, *ND3*, and *CYTB*) use the incomplete stop codon T. The *ATP8* and *ATP6* genes overlap for 10 bp, *ATP6* overlaps for 1 bp with *COXIII*, and *ND4L* and *ND4* overlap for 1 bp. Additionally, a frameshift is found in the *Enhydris plumbea ND1* gene between positions 3315 and 3431 (Appendix 1).

Sequence features of rRNA, tRNA, and pseudo-tRNA^{Pro} genes

The sequence length of the 12S and 16S rRNA is 925 and 1479 bp, respectively. They are located between the tRNA^{Phe} and *ND1* genes and are separated by the tRNA^{Val} gene, as observed in other snakes. The base composition of the 12S rRNA gene is 18.6% T, 28% C, 34.5% A, and 18.9% G, while the 16S rRNA composition is 20.2% T, 25.6% C, 38.1% A, and 16.2% G. The total length of rRNAs is 2404 bp, which is similar to that of other colubroid (Jiang et al. 2007) and *Achalinus* (Wang et al. 2009) snakes but shorter than that of members of Scolecophidia, Henophidia, and *Acrochordus* (Jiang et al. 2007).

The mitochondrial genome of *T. zhaoermii* contains 22 tRNA genes, in which tRNA$^{Ser\ (AGY)}$ is the shortest (57 bp) and tRNA^{Asn} is the longest (74 bp). Detailed information on the 22 tRNA genes is also presented in Table III. Furthermore, the total length of tRNA is 1436 bp, which is similar to other snakes (Jiang et al. 2007). The length reduction of tRNAs in snakes is due to the truncation of the TΨC arm (Kumazawa et al. 1996, 1998) and "DHU" arm in some tRNAs (Jiang et al. 2007).

A 26 bp pseudo-tRNA^{Pro} gene also appears in the mitochondrial genome of *T. zhaoermii* in front of CR II, which is similar to other snakes in Colubridae and Homalopsidae (Jiang et al. 2007; Yan et al. 2008). However, before the pseudo-tRNA^{Pro} gene, there is a 40 bp non-coding region containing a 16 bp C-rich segment (CCCCCCCTACCCCCCC). A similar feature has been observed in the mitochondrial genome of Russell's viper (*Daboia russellii*), which also possesses a 92 bp non-coding region between tRNA^{Ile} and tRNA^{Pro}, containing an 18 bp C-rich segment (CCCCCCCCCTACCCCCCCC). In the rice paddy snake (*E. plumbea*) mtDNA, there is a 28 bp non-coding region upstream of the 12 bp pseudo-tRNA^{Pro} gene. The appearance of the pseudo-tRNA^{Pro} gene in *T. zhaoermii*, which could be characteristic of Xenodontinae, may suggest that after the duplication of the pseudo-tRNA^{Pro} gene in Colubroidea (except for *Achalinus*), Homalopsidae, and Colubridae, the tRNA^{Pro} duplicate adjacent to CR II started to degrade, while in Viperidae the tRNA^{Pro} duplicate adjacent to CR I started to degrade.

Control regions

Like all vertebrates, all snakes sampled in the current study have a control region (CR I) adjacent to the 5′ end of the 12S rRNA gene (Kumazawa et al. 1996, 1998; Dong and Kumazawa 2005; Jiang et al. 2007; Yan et al. 2008). Except for scolecophidian snakes

(*Leptotyphlops dulcis*, *Ramphotyphlops australis*, *Ramphotyphlops braminus*, *Typhlops reticulatus*, and *Typhlops mirus*), the other sampled snakes have a duplicated control region (CR II) between *ND1* and *ND2*. In this study, the lengths of CR I and CR II in *T. zhaoermii* are 1093 and 1023 bp, respectively. The base composition of CR I and CR II is very similar. CR I contains 31.6% T, 28% C, 26.8% A, and 13.6% G, while CR II contains 31.9% T, 28.2% C, 26.6% A, and 13.4% G. The structure of the 40 bp non-coding region adjacent to the pseudo-tRNA^{Pro} gene is found to be highly homologous to the initial 5–46 bp of CR II and the beginning 6–30 bp and 32–73 bp of CR I with some substitutions and indels (Appendix 2).

Phylogenetic results

The aligned 12 heavy-strand encoded protein-coding genes totaling 10,668 bp in length were used to reconstruct the snake phylogeny using the BI and ML methods. Both phylogenetic methods provide identical and well-supported tree topologies; most branches were strongly supported with over 90% bootstrap support values (BS) and 1.0 Bayesian posterior probabilities (PP) (Figure 2). The BI tree is shown in Figure 2 and the BS/PP are shown on the corresponding branches.

The phylogenetic results indicate that Scolecophidia do not form a monophyletic group. *R. australis*, *R. braminus*, *T. reticulatus*, and *T. mirus* form a solid monophyletic group (BS = 100 and PP = 1.00), and *L. dulcis* is strongly suggested (BS = 100 and PP = 1.00) to be the sister group to the monophyletic Alethinophidia (BS = 100 and PP = 1.00) (Henophidia + Caenophidia). This result supports the non-monophyly of Scolecophidia mentioned by Yan et al. (2008), which involved only two species (*R. braminus* and *L. dulcis*) to illustrate Scolecophidia based on the mitochondrial genome. Furthermore, these results corroborate the findings of

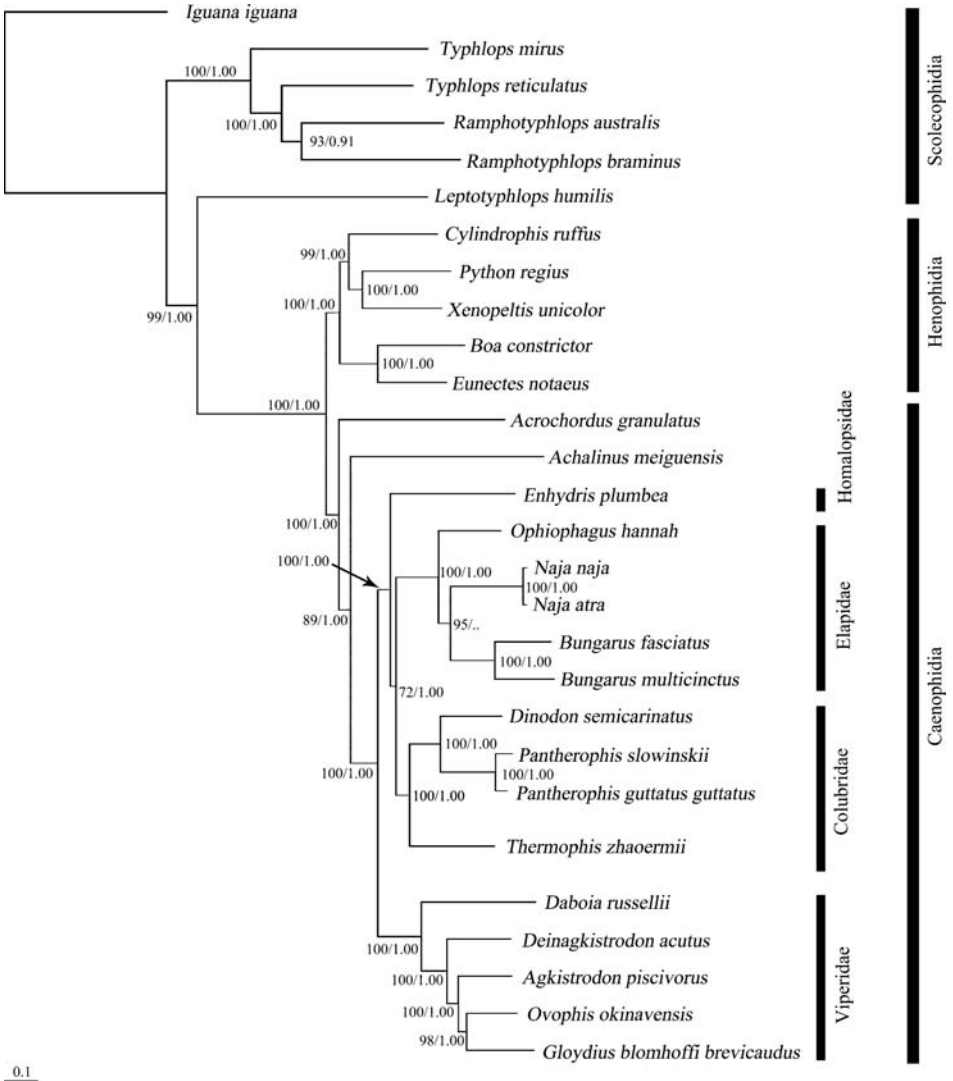

Figure 2. Bayesian tree estimated using the combined 12 protein-coding gene alignment. Values on the nodes indicate bootstrap and posterior probability support (BS/PP).

Heise et al. (1995) and Wiens et al. (2008), concurring on the non-monophyly of Scolecophidia. However, these results disagree with those of Slowinski and Lawson (2002), Vidal and Hedges (2002), and Vidal et al. (2009). Furthermore, the genus *Typhlops* in the current study does not form a monophyletic group. *T. mirus* becomes the sister lineage to the three other species, whereas *T. reticulatus* acts as the sister group to the monophyletic genus *Ramphotyphlops*.

Within Alethinophidia, the monophyly of both Henophidia (BS = 100 and PP = 1.00) and Caenophidia (BS = 100 and PP = 1.00) is also shown. At the same time, in Caenophidia, *Acrochordus granulatus* is the basal taxon, while *Achalinus meiguensis* is the sister group to the remaining colubroid snakes (BS = 89 and PP = 1.00). As a result, the "Acro-Heno" clade, which is suggested to be comprised of *A. granulatus* and the henophidian snakes (Jiang et al. 2007), is not supported. The phylogenetic position of "Xenodermatinae" snakes as represented by *Achalinus* (*A. meiguensis*) in the current study was identical to the topology structured by *Acrochordus*, *Achalinus,* and other colubroids presented by Vidal and Hedges (2002). In addition to that, Wiens et al.'s (2008) nuclear gene analysis suggested that the Xenodermatinae, which was represented by *Xenodermus*, should be placed at the root of Colubroidea—a finding corroborated by our results. Following our complete mtDNA data analysis, we support the results of Vidal et al. (2007), suggesting that the subfamily Xenodermatinae should not be placed within the Elapidae (Lawson et al. 2005), but should rather be upgraded to a family rank.

Three monophyletic families of advanced snakes are presented: Colubridae (BS = 100 and PP = 1.00), Elapidae (BS = 100 and PP = 1.00), and Viperidae (BS = 100 and PP = 1.00). The monophyletic Viperidae clade is supported as the oldest diverging lineage in Colubroidea in which the Viperinae clade, as represented by *D. russellii*, is sister to the monophyletic Crotalinae (*Agkistrodon piscivorus*, *Deinagkistrodon acutus*, *Gloydius blomhoffi brevicaudus*, and *Ovophis okinavensis*). This is well supported by numerous previous molecular phylogenetic studies (Parkinson 1999; Lawson et al. 2005; Castoe and Parkinson 2006; Wiens et al. 2008; Wüster et al. 2008).

Homalopsidae, represented in our study by only one member (*E. plumbea*), is well supported as a sister group (BS = 100 and PP = 1.00) to the monophyletic (BS = 72 and PP = 1.00) Colubridae + Elapidae grouping. In the monophyletic Elapidae (BS = 100 and PP = 1.00), the two genera *Bungarus* and *Naja* received unconditional monophyly; however, the relationship between *Ophiophagus hannah*, *Bungarus,* and *Naja* remains unclear. Accordingly, the familial rank of Homalopsidae is reconfirmed in agreement with previous studies (Vidal and Hedges 2002; Kelly et al. 2003; Lawson et al. 2005; Wiens et al. 2008).

T. zhaoermii, which could represent Xenodontinae, is solidly placed in the Colubridae as a sister group to the monophyletic Colubrinae (*D. semicarinatus*, *Pantherophis guttatus guttatus*, and *Pantherophis slowinskii*). Due to our limited taxonomic sampling, we could only conclude that the genus *Thermophis* should be firmly placed in the Colubridae in accordance with morphological (Wall 1907; Malnate 1953; Zaher 1999; Guo et al. 2008) and molecular analyses (He et al. 2009; Huang et al. 2009).

Divergence time estimates

The relaxed molecular clock analyses (Figure 3) suggest that Alethinophidia began to diverge from the paraphyletic Scolecophidia approximately 130 Mya in the early Cretaceous, which is consistent with the results of Burbrink and Pyron (2008) even with a relatively large 95% highest posterior density (74–130 Mya). The estimates also suggest that the divergence of living alethinophidian snakes, the radiation of the Caenophidia, and the separation between *Acrochordus* and the Colubroidea (Colubroidea + *A. meiguensis*) have taken place approximately 70, 65, and 60 Mya, respectively. These are consistent with the conclusions of Wüster et al. (2008) and Burbrink and Pyron (2008), but younger than Vidal et al.'s (2009) estimates. The above events occurred around the Cretaceous–Tertiary boundary, so those divergences might have been caused by the K/T extinction event that erased nearly 76% of life on Earth (Pope et al. 1998). During that period, the duplication of the CR and transposition of $tRNA^{Leu}$ in Alethinophidia, the length increase of the rRNA genes in Henophidia and *Acrochordus*, and the duplication of $tRNA^{Pro}$, and length reduction in the tRNA and rRNA genes in Colubroidea were rapidly taking place in mitochondrial genomes (Jiang et al. 2007).

The origin of Colubridae was estimated at 40 Mya while the first viper diverged approximately 35 Mya. This result is consistent with Rage's (1987) and Rage et al.'s (1992) suggestion that Asia is the center of origin for both the Colubridae and Viperidae based on the fact that the oldest known colubrid fossil was found in Thailand and is from the late Eocene (Rage et al. 1992). Finally, Cadle (1987) and Keogh (1998) pointed out that Asia could be the most probable center of origin for elapids based on morphological characters and molecular phylogeny, and the origin time of Elapidae (about 35 Mya) is somewhat coincident with the origin time of Viperidae and close to Colubridae. We are unaware of any major geological or climatic events occurring during that time, leading to the tremendous divergence of Colubroidea. As the rapid uplift of southern Tibet that took place about 20 Mya (Harrison et al. 1992) may have lead to the divergence of *Thermophis* from

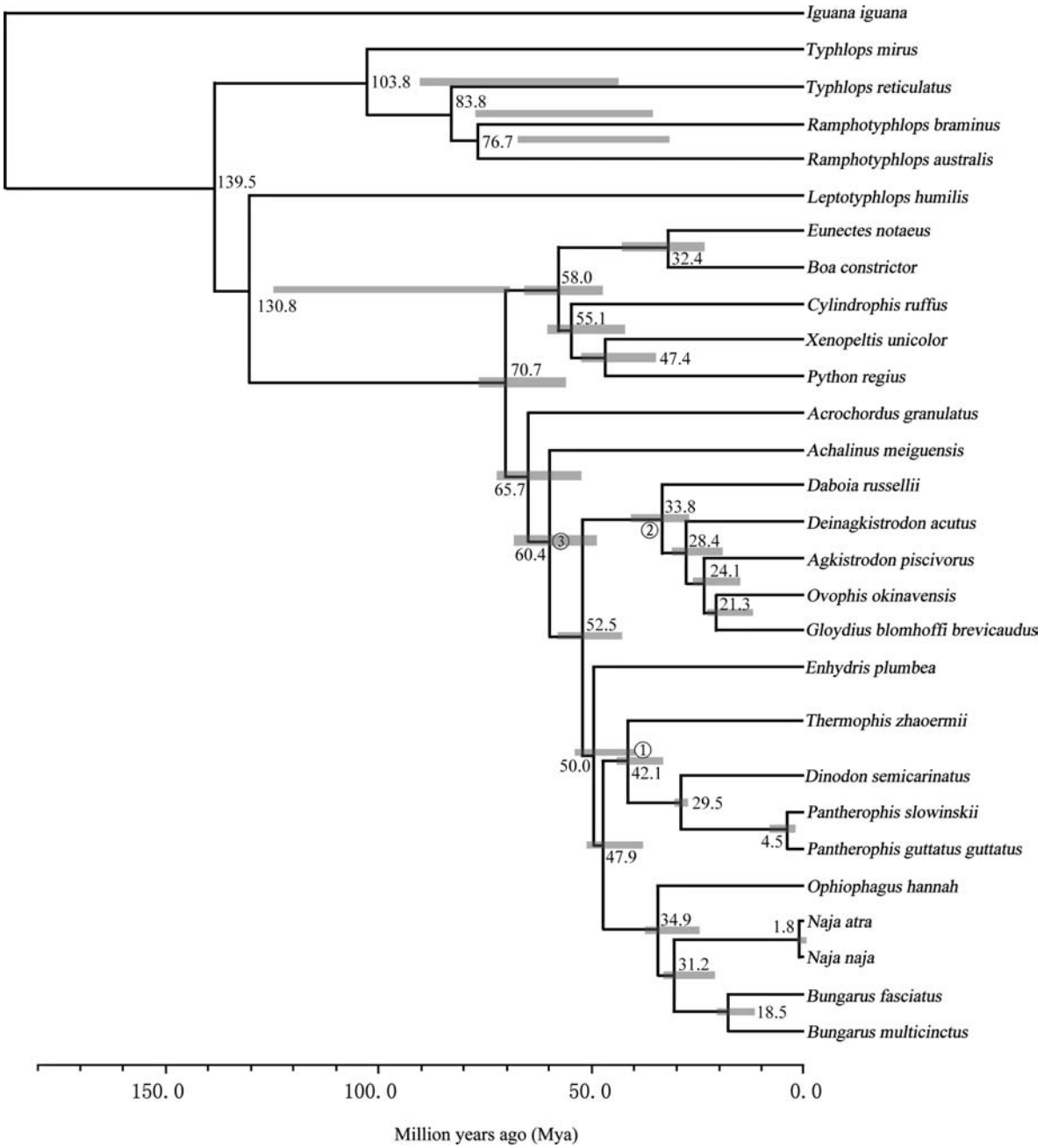

Figure 3. BEAST maximum credibility ultrametric tree. Calibration points: (1) "Colubrinae", (2) "Viperinae", and (3) "Colubroidea". Mean divergence times (Mya) are shown on the nodes. Gray node bars indicate the 95% highest posterior density of the divergence times.

the ancestor of xenodontine snakes distributed in the New World (He et al. 2009; Huang et al. 2009), we speculate that the northward convergence of India into Asia occurring about 40–50 Mya (Harrison et al. 1992) probably promoted Asia to be the divergence center of Colubroidea.

In conclusion, the complete mitochondrial genome of *T. zhaoermii* is first reported here. These results not only shed light on the evolution of this mysterious snake species, but also help setting the stage for future studies on the origin of New World xenodontine snakes and the molecular evolution of the mitochondrial genome of snakes.

Acknowledgments

The authors are grateful to Jiatang Li for collecting specimens and to Bo Zeng, Kun Wei, Emily King, and the anonymous reviewers for kindly commenting on the manuscript. This research was funded in part by "Project 985" of The Central University for Nationalities (CUN985-3-3).

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 identification model. IEEE Trans Autom Contr 19:716–723.

Burbrink FT, Pyron RA. 2008. The taming of the skew: Estimating proper confidence intervals for divergence dates. Syst Biol 57: 317–328.

Cadle JE. 1987. Geographic distribution: Problems in phylogeny and zoogeography. In: Seigel RA, Collins JT, Novak SS, editors. Snakes: Ecology and evolutionary biology. New York: McGraw-Hill. p 77–105.

Castoe TA, Parkinson CL. 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol Phylogenet Evol 39:91–110.

Castoe TA, Jiang ZJ, Gu WJ, Wang ZYO, Pollock DD. 2008. Adaptive evolution and functional redesign of core metabolic proteins in snakes. PLoS ONE 3(5):e2201.

Chen N, Zhao SJ. 2009. New progress in snake mitochondrial gene rearrangement. Mitochondrial DNA 20(4):69–71.

Dong S, Kumazawa Y. 2005. Complete mitochondrial DNA sequences of six snakes: Phylogenetic relationships and molecular evolution of genomic features. J Mol Evol 61(1):12–22.

Douglas D, Janke A, Arnason U. 2006. A mitogenomic study on the phylogenetic position of snakes. Zool Scr 35(6):545–558.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214.

Guo P, Liu SY, Feng JC, He M. 2008. The description of a new species of *Thermophis* (Serpentes; Colubridae). Sichuan J Zool 27:321.

Harrison TM, Copeland P, Kidd WSF, Yin A. 1992. Raising Tibet. Science 255:1663–1670.

He M, Feng JC, Liu SY, Guo P, Zhao EM. 2009. The phylogenetic position of *Thermophis* (Serpentes: Colubridae), an endemic snake from the Qinghai-Xizang Plateau, China. J Nat Hist 43: 479–488.

Head JJ, Holroyd PA, Hutchison JH, Ciochon RL. 2005. First report of snakes (Serpentes) from the Late Middle Eocene of the Ponduang Formation, Myanmar. J Vert Paleontol 25:246–250.

Heise PJ, Maxson LR, Dowling HG, Hedges SB. 1995. Higherlevel snake phylogeny inferred from mitochondrial DNA sequences of 12S rRNA and 16S rRNA genes. Mol Biol Evol 12:259–265.

Holman JA. 2000. Fossil snakes of North America: Origin, evolution, distribution, paleoecology. Bloomington: Indiana University Press.

Huang S, Liu SY, Guo P, Zhang YP, Zhao EM. 2009. What are the closest relatives of the hot-spring snakes (Colubridae, *Thermophis*), the relict species endemic to the Tibetan Plateau? Mol Phylogenet Evol 51:438–446.

Janke A, Erpenbeck D, Nilsson M, Arnason U. 2001. The mitochondrial genomes of the iguana (Iguana iguana) and the caiman (Caiman crocodylus): Implications for amniote phylogeny. Proceedings of the Royal Society of London Series B-Biological Sciences 268(1467):623–631.

Jiang ZJ, Castoe TA, Austin CC, Burbrink FT, Herron MD, McGuire JA, Parkinson CL, Pollock DD. 2007. Comparative mitochondrial genomics of snakes: Extraordinary substitution rate dynamics and functionality of the duplicate control region. BMC Evol Biol 7(1):123.

Kelly CMR, Barker NP, Villet MH. 2003. Phylogenetics of advanced snakes (Caenophidia) based on four mitochondrial genes. Syst Biol 52:439–459.

Keogh JS. 1998. Molecular phylogeny of elapid snakes and a consideration of their biogeographic history. Biol J Linn Soc 63: 177–203.

Kumazawa Y, Ota H, Nishida M, Ozawa T. 1996. Gene rearrangements in snake mitochondrial genomes: Highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Mol Biol Evol 13(9): 1242–1254.

Kumazawa Y, Ota H, Nishida M, Ozawa T. 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. Genetics 150(1): 313–329.

Kumazawa Y. 2004. Mitochondrial DNA sequences of five squamates: Phylogenetic affiliation of snakes. DNA Res 11(2): 137–144.

Lanave C, Preparata G, Saccone C, Serio G. 1984. A new method for calculating evolutionary substitution rates. J Mol Evol 20: 86–93.

Lawson R, Slowinski JB, Crother BI, Burbrink FT. 2005. Phylogeny of the Colubroidea (Serpentes): New evidence from mitochondrial and nuclear genes. Mol Phylogenet Evol 37:581–601.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 25:955–964.

Malnate EV. 1953. The taxonomic status of the Tibetan colubrid snake *Natrix baileyi*. Copeia 1953:92–96.

Nylander JAA. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Parkinson CL. 1999. Molecular systematics and biogeographical history of pitvipers as determined by mitochondrial ribosomal DNA sequences. Copeia 1999:576–586.

Pope KO, D'Hondt SL, Marshall CR. 1998. Meteorite impact and the mass extinction of species at the Cretaceous/Tertiary boundary. Proc Natl Acad Sci USA 95:11028–11029.

Rage JC. 1987. Fossil record. In: Seigel RA, Collins JT, Novak SS, editors. Snakes: Ecology and evolutionary biology. New York: McGraw-Hill Publishing. p 51–76.

Rage JC, Werner C. 1999. Mid-Cretaceous (Cenomanian) snakes from Wadi Abu Hashim, Sudan: The earliest snake assemblage. Palaeontol Afr 35:85–110.

Rage JC, Buffetaut E, BuVetaut-Tong H, Chaimanee Y, Ducrocq S, Jaeger JJ, Suteethorn V. 1992. A colubrid snake in the late Eocene of Thailand: The oldest known Colubridae (Reptilia, Serpentes). C R Biol 2314:1085–1089.

Rage JC, Bajpai S, Thewissen JGM, Tiwari BN. 2003. Early Eocene snakes from Kutch, Western India, with a review of the Palaeophiidae. Geodiversitas 25:695–716.

Rambaut A, Drummond AJ. 2007. Tracer v1.4. Available online at: http://beast.bio.ed.ac.uk/Tracer. Accessed: Dec. 28, 2007

Ronquist FR, Huelsenbeck JP. 2003. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 19:1572–1574.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: A laboratory manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press.

Slowinski JB, Lawson R. 2002. Snake phylogeny: Evidence from nuclear and mitochondrial genes. Mol Phylogenet Evol 23: 194–202.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RaxML web-servers. Syst Biol 75(5):758–771.

Szyndlar Z, Rage JC. 1990. West Palearctic cobras of the genus *Naja* (Serpentes: Elapidae): Interrelationships among extinct and extant species. Amphibia-Reptilia 11:385–400.

Szyndlar Z, Rage JC. 2002. Fossil record of the true vipers. In: Schuett GW, Höggren M, Douglas ME, Greene HW, editors. Biology of the vipers. Eagle Mountain, Colorado: Eagle Mountain Press. p 419–444.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins D. 1997. The Clustal X windows interface: Flexible strategies for

multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 25:4876–4882.

Vidal N, Hedges SB. 2002. Higher-level relationships of caenophidian snakes inferred from four nuclear and mitochondrial genes. C R Biol 325:987–995.

Vidal N, Delmas A, David P, Cruaud C, Couloux A, Hedges SB. 2007. The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes. C R Biol 330: 182–187.

Vidal N, Rage J-C, Couloux A, Hedges SB. 2009. Snakes (Serpentes). In: Hedges SB, Kumar S, editors. The timetree of life. Oxford: Oxford University Press. p 390–397.

Wall F. 1907. Some new Asian snakes. J Bombay Nat Hist Soc 17: 612–618.

Wang GL, He SP, Huang S, He M, Zhao EM. 2009. The complete mitochondrial DNA sequence and the phylogenetic position of *Achalinus meiguensis* (Reptilia: Squamata). Chin Sci Bull 54(10): 1713–1724.

Wiens JJ, Kuczynski CA, Smith SA, Mulcahy DG, Sites JW Jr, Townsend TM, Reeder TW. 2008. Branch lengths, support, and congruence: Testing the phylogenomic approach with 20 nuclear loci in snakes. Syst Biol 57(3):420–431.

Wüster W, Peppin L, Pook CE, Walker DE. 2008. A nesting of vipers: Phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes). Mol Phylogenet Evol 49:445–459.

Yan J, Li HD, Zhou KY. 2008. Evolution of the mitochondrial genome in snakes: Gene rearrangements and phylogenetic relationships. BMC Genomics 9:569.

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J Mol Evol 39:306–314.

Zaher H. 1999. Hemipenial morphology of the South American xenodontine snakes, with a proposal for a monophyletic Xenodontinae and a reappraisal of colubroid hemipenes. Bull Am Mus Nat Hist 240:1–168.

Appendix 1. Partial aligned sequence of the *ND1* gene

A frameshift is found in *E. plumbea*.

Appendix 2. Aligned sequence of the non-coding region and beginning part of control regions I and II (CR I and II)

The non-coding region adjacent to pseudo-P is highly homologous to the initial 5–46 bp of CR II and the beginning 6–30 bp and 32–73 bp of CR I.