

# Phylogenetic relationships among amphisbaenian reptiles based on complete mitochondrial genomic sequences

J. Robert Macey<sup>a,b,\*</sup>, Theodore J. Papenfuss<sup>b</sup>, Jennifer V. Kuehl<sup>a</sup>,  
H. Mathew Fourcade<sup>a</sup>, Jeffrey L. Boore<sup>a,c</sup>

<sup>a</sup> Department of Evolutionary Genomics, DOE Joint Genome Institute and Lawrence Berkeley National Laboratory, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

<sup>b</sup> Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720, USA

<sup>c</sup> Department of Integrative Biology, 3060 Valley Life Science Building, University of California, Berkeley, CA 94720, USA

Received 18 December 2003; revised 27 April 2004

Available online 6 July 2004

## Abstract

Complete mitochondrial genomic sequences are reported from 12 members in the four families of the reptile group Amphisbaenia. Analysis of 11,946 aligned nucleotide positions (5797 informative) produces a robust phylogenetic hypothesis. The family Rhineuridae is basal and Bipedidae is the sister taxon to the Amphisbaenidae plus Trogonophidae. Amphisbaenian reptiles are surprisingly old, predating the breakup of Pangaea 200 million years before present, because successive basal taxa (Rhineuridae and Bipedidae) are situated in tectonic regions of Laurasia and nested taxa (Amphisbaenidae and Trogonophidae) are found in Gondwanan regions. Thorough sampling within the Bipedidae shows that it is not tectonic movement of Baja California away from the Mexican mainland that is primary in isolating *Bipes* species, but rather that primary vicariance occurred between northern and southern groups. Amphisbaenian families show parallel reduction in number of limbs and *Bipes* species exhibit parallel reduction in number of digits. A measure is developed for comparing the phylogenetic information content of various genes. A synapomorphic trait defining the Bipedidae is a shift from the typical vertebrate mitochondrial gene arrangement to the derived state of *trnE* and *nad6*. In addition, a tandem duplication of *trnT* and *trnP* is observed in *Bipes biporus* with a pattern of pseudogene formation that varies among populations. The first case of convergent rearrangement of the mitochondrial genome among animals demonstrated by complete genomic sequences is reported. Relative to most vertebrates, the Rhineuridae has the block *nad6*, *trnE* switched in order with the block *cob*, *trnT*, *trnP*, as they are in birds.

Published by Elsevier Inc.

**Keywords:** Reptilia; Amphisbaenia; *Bipes*; Phylogeny; Mitochondrial DNA; Genomics; Limb evolution; Biogeography; Pangaea

## 1. Introduction

Amphisbaenians are fossorial reptiles that inhabit tropical and semitropical regions of the world (Fig. 1). The vast majority of species occur on land masses of Gondwanan origin in South America, Africa, and Arabia. However, two of the four families are restricted to regions associated with Laurasia, the Bipedidae, with

three species in Mexico (Fig. 2), and the Rhineuridae, with a single species in Florida. These taxa have pleurodont dentition (Gans, 1968) and were originally placed in the family Amphisbaenidae which now primarily includes taxa distributed in Gondwanan regions of South America and Africa (Gans, 1978). The fourth family of Amphisbaenia, the Trogonophidae, is the only one to have acrodont dentition and is restricted to Gondwanan plates of North Africa, Arabia, and the island of Socotra. By superimposing a phylogeny of these groups on their distribution relative to the breakup of these major

\* Corresponding author. Fax: 1-925-296-5666.  
E-mail address: [jrmacey@lbl.gov](mailto:jrmacey@lbl.gov) (J.R. Macey).

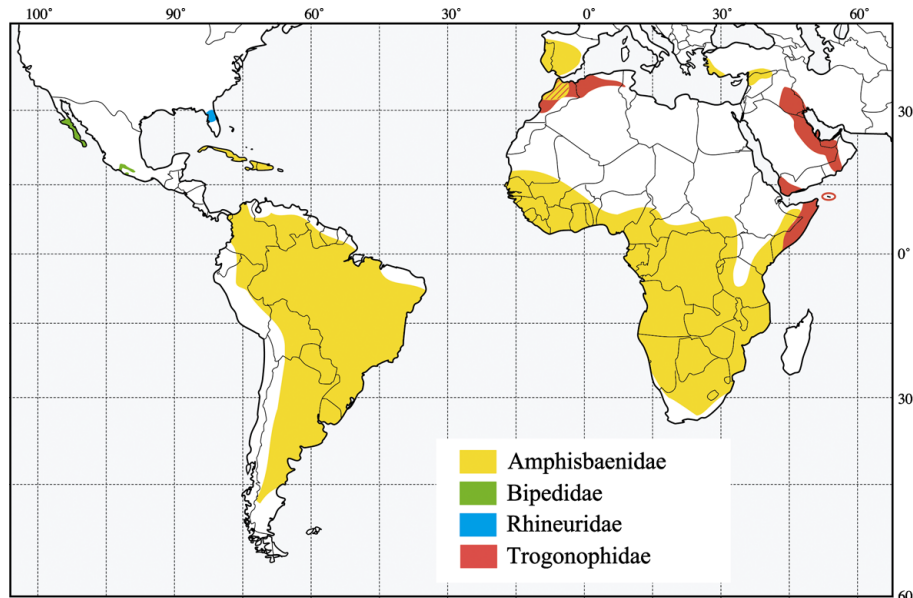


Fig. 1. Distribution of the four families of Amphisbaenia. Note that Rhineuridae is restricted to Florida and Bipedidae to the west coast of Mexico, both regions of Laurasian origin. The Trogonophidae is restricted to Gondwanan regions of North Africa, Arabia, and the Indian Ocean island of Socotra. The largest family, the Amphisbaenidae, is primarily distributed in Gondwanan regions of South America and Africa. The hatched area in Morocco represents overlapping distributions of Amphisbaenidae and Trogonophidae.

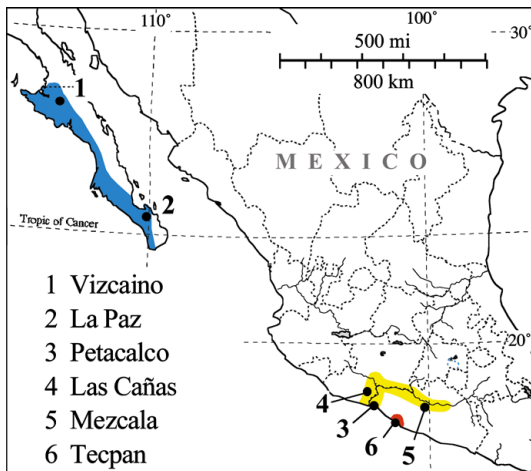


Fig. 2. Distribution of the three species in the Bipedidae. *Bipes biporus* (blue) is restricted to peninsular Baja California and is sampled from the northern and southern portions of its range (localities 1 and 2). Three samples of *B. canaliculatus* (yellow) are examined, two from along the Rio Balsas (Petacalco and Mezcala) and one from along the Rio Tepalcatepec (Las Cañas), (localities 3, 4, and 5). Three samples are examined of *B. tridactylus* (red) all from near the mouth of the Rio Tecpan (locality 6).

land masses, it is possible to test hypotheses about their origins. For example, monophyly of taxa found in Laurasia or Gondwana may reflect their isolation stemming from the breakup of Pangaea 200 million years before present (MYBP) (Figs. 3C and D).

A phylogeny of Amphisbaenia is also important for reconstructing patterns of morphological change.

Among the four families of Amphisbaenia, all taxa are limbless except the Bipedidae, which contains reptiles with two forelimbs. If gradual reduction in limbs were to have occurred from an ancestor with both hind- and forelimbs, then the Bipedidae, with its presumed intermediate state, would be thought to be the sister taxon to the remaining families as has been suggested by morphological data (Kearney, 2003; Fig. 3B).

Within the Bipedidae, species are variable for the number of digits on a limb (Papenfuss, 1982). The Baja California taxon, *Bipes biporus*, always has five digits on each limb. In mainland Mexico, the more northern taxon, *Bipes canaliculatus*, is variable, with either four or five digits, while the more southern species, *Bipes tridactylus*, always has three digits on each limb. If digits are being gradually lost from an ancestor with five digits then *B. biporus* is predicted to be the sister taxon to the other two species (Fig. 3E). This phylogenetic hypothesis is consistent with an initial splitting of *Bipes* by the separation of peninsular Baja California from the west coast of Mexico (12–14 MYBP; Ferrari, 1995), with the two mainland species later diverging. Alternatively, tectonic activity along the west coast of Mexico in the Cretaceous to Paleocene (Ferrari, 1995) may have isolated the more northern taxa, *B. biporus* and *B. canaliculatus*, from *B. tridactylus* to the south (Fig. 3F).

Mitochondrial gene order has often been found to be a robust phylogenetic character for reconstructing evolutionary relationships among animals (Boore, 1999; Boore and Brown, 1998; Macey et al., 1997a, 2000). Although no character can be asserted to be completely

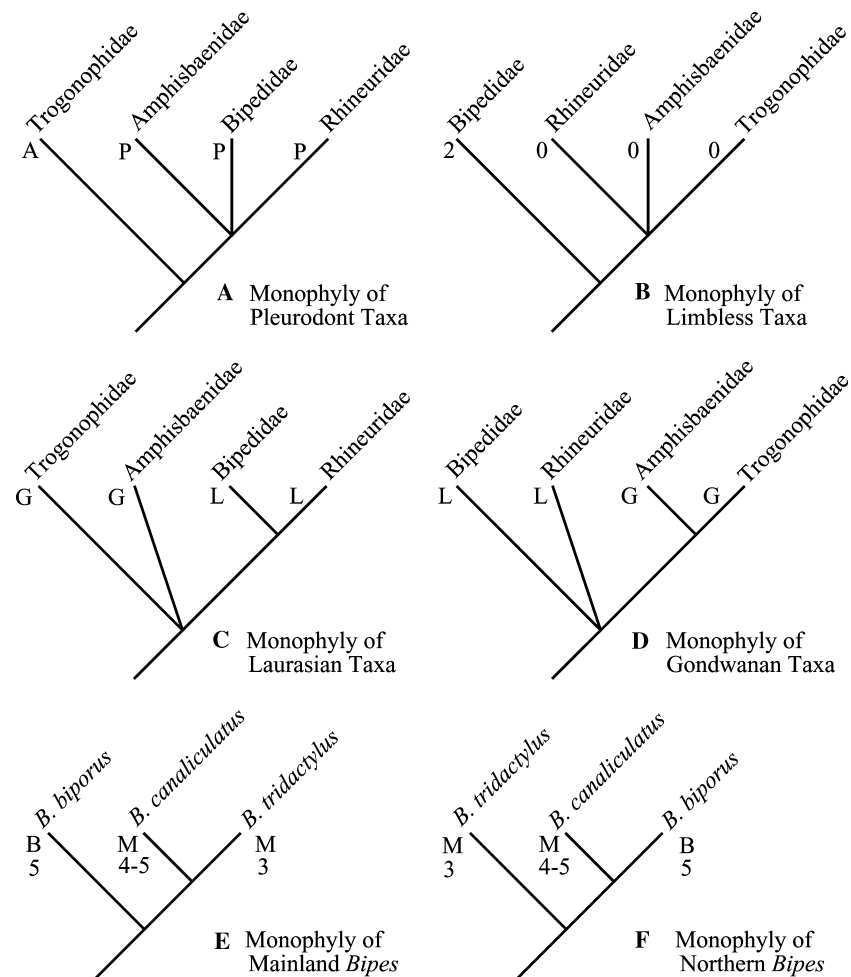


Fig. 3. Alternative phylogenetic hypotheses tested. Hypotheses (A) through (D) concern relationships among amphisbaenian families. (A) Monophyly of taxa with pleurodont “P” as apposed to acrodon “A” dentition (Gans, 1968). (B) Monophyly of limbless taxa (Kearney, 2003) “0” as apposed to having two forelimbs “2.” (C and D) Monophyly of Laurasian “L” vs. monophyly of Gondwanan “G” taxa. Note that hypothesis (A) is consistent with (C), and hypothesis (B) is consistent with (D). Hypotheses (E and F) concern relationships among the three species in the Bipedidae where on the trees “B” represents Baja California and “M” represents mainland Mexico. The number of digits in forelimbs are also plotted on the trees. (E) Monophyly of the two mainland *Bipes* taxa which have reduced numbers of digits (3 and usually 4; Papenfuss, 1982). (F) Monophyly of the more northern *Bipes* taxa, which suggests independent loss of digits in the two mainland species.

free from the possibility of homoplasious change, parallelisms or reversals of gene rearrangements seem, in principle, to be unlikely, and have not been observed to date in sampling approximately 300 complete mtDNA sequences. A common mode of mitochondrial genomic rearrangement appears to be a tandem duplication of genes with random loss of extra copies (Boore, 2000; Macey et al., 1997a; Moritz et al., 1987). Hence, the pattern of pseudogene formation in a tandemly duplicated segment is causative in establishing novel gene orders.

Amphisbaenian reptiles are known to have several novel mitochondrial genomic features. *Bipes biporus* has an unusual stem and loop structure atypical of vertebrates between *trnN* and *trnC* (Macey et al., 1997a) where light-strand replication usually initiates. This character has been linked to mitochondrial genomic re-

arrangement among vertebrates (Macey et al., 1997a). In this same taxon, *trnC* encodes a transfer RNA that lacks a dihydrouridine (D) arm and instead contains a D-arm replacement loop (Macey et al., 1997b). In addition, this taxon has a tandem duplication of *trnT* and *trnP* with a pattern of pseudogene formation that does not result in mitochondrial genomic rearrangement (Macey et al., 1998).

In addition, each complete mtDNA sequence includes 37 genes whose sequences can be compared (about 16kb in total). Recent studies have shown that comparisons of complete mtDNA sequences give much more robust phylogenetic reconstructions than when using smaller portions of the mtDNA (Ingman et al., 2001).

The complete mitochondrial genomic sequence is newly reported from 12 specimens representing all four

families of the Amphisbaenia (Fig. 1). The single species of Rhineuridae, *Rhineura floridana*, is sampled from Florida. In the Bipedidae, samples from several localities in each of the three species are sequenced (Fig. 2). *Diplometapon zarudnyi* is sampled from Iran on the Arabian Shield to represent the Trogonophidae. In the Amphisbaenidae, *Amphisbaena schmidti* from Puerto Rico is used as a representative of the New World radiation and *Geocalamus acutus* from Tanzania is sampled to represent the Old World. Two diverse lizard taxa are used as outgroups with data from the literature, an iguanian, *Iguana iguana* (AJ278511, Janke et al., 2001), and a sceleroglossan, *Eumeces egregius* (AB016606, Kumazawa and Nishida, 1999).

## 2. Materials and methods

The alignment of sequences is deposited at Elsevier electronic [supplementary material](#) to this article.

### 2.1. Specimen information

All voucher specimens are deposited at the Museum of Vertebrate Zoology, University of California at Berkeley (MVZ), or Universidad Nacional Autónoma de México (UNAM). Full locality data are listed in GenBank files. Species names, followed by corresponding museum voucher specimen numbers, then GenBank accession numbers are as follows: *Rhineura floridana*, Florida (MVZ 233342, AY605473); *Diplometapon zarudnyi*, Iran (MVZ 234273, AY605474); *Amphisbaena schmidti*, Puerto Rico (MVZ 232754, AY605475); *G. acutus*, Tanzania (MVZ 232838, AY605476); *Bipes tridactylus*, 13km south of Rio Tecpan, Guerrero, Mexico (MVZ 236261, AY605477); *B. tridactylus*, south bank Rio Tecpan, Guerrero, Mexico (MVZ 236262, AY605478); *B. tridactylus*, north bank Rio Tecpan, Guerrero, Mexico (MVZ 236305, AY605479); *B. biporus*, Vizcaino, northern Baja California, Mexico (MVZ 236257, AY605480); *B. biporus*, La Paz, southern Baja California, Mexico (MVZ 236258, AY605481); *B. canaliculatus*, Las Cañas, Michoacan, Mexico (MVZ 233341, AY605482); *B. canaliculatus*, Mezcala, Guerrero, Mexico (MVZ 240725, AY605483); and *B. canaliculatus*, Petacalco, Guerrero, Mexico (UNAM-TP27893, AY605484).

### 2.2. Laboratory protocols

Genomic DNA was extracted from liver or muscle using the Qiagen QIAamp tissue kit. Amplification of genomic DNA was conducted using rTth long PCR enzyme (Applied Biosystems) with a denaturation at 94 °C for 35s, annealing at 50 °C for 35s, and extension at 70 °C for 150s with 4s added to the extension per cycle, for 38 cycles. Negative controls were run on all amplifi-

cations to check for contamination. Initial amplifications were conducted using primers described in Macey et al. (1997a, 1998). Perfectly matching primers were then constructed for each taxon based on the DNA sequence of this fragment to complete the amplification of each mtDNA.

Amplification products were sheared randomly into fragments of approximately 1.5kb by repeated passage through a narrow aperture using a Hydrashear device. After end-repair, the sheared DNA was gel purified and ligated into pUC18 vector to construct a library of random fragments, and then transformed into bacterial cells. Automated colony pickers introduced single clones into bacterial broth in 384-well format. These plasmid clones were processed robotically through rolling circle amplification (Dean et al., 2001; Hawkins et al., 2002), sequencing reactions, and reaction cleanup using SPRI (Elkin et al., 2002). Sequences were determined using ABI3730xl DNA sequencers and then assembled to form deep contigs using Phrap or Sequencher.

### 2.3. Phylogenetic analysis

DNA sequences were aligned manually. Positions encoding proteins were translated to amino acids using MacClade 4.03 (Maddison and Maddison, 2001) for confirmation of alignment. Alignments of sequences encoding tRNAs were constructed based on secondary structural models (Kumazawa and Nishida, 1993; Macey and Verma, 1997). Secondary structures of tRNAs were inferred from primary structures of the corresponding tRNA genes using these models. Unalignable regions (see Section 3) were excluded from phylogenetic analyses. Non-coding and rRNA gene sequences were not included due to ambiguity of alignment.

Phylogenetic trees were inferred by parsimony using PAUP\* beta version 4.0b8 (Swofford, 2001) with the branch and bound search option, which guarantees an exact solution. Bootstrap resampling (Felsenstein, 1985a) was applied to assess support for individual nodes using 1000 branch and bound replicates. Decay indices (=“branch support” of Bremer, 1994) were calculated for all internal branches using branch and bound searches that retained suboptimal nodes.

Alternative phylogenetic hypotheses (Fig. 3) were tested using the most parsimonious phylogenetic topologies compatible with each. To find the most parsimonious tree(s) compatible with a particular phylogenetic hypothesis, topologies were first constructed using MacClade (Maddison and Maddison, 2001), providing input as constraint trees to PAUP\* (Swofford, 2001) for subsequent branch and bound searches. Wilcoxon-signified-ranks tests (Templeton, 1983) were used to examine statistical significance of the shortest tree relative to alternative hypotheses. This test determines whether the most parsimonious tree is significantly shorter than an

alternative or whether their differences in length are statistically indistinguishable. Wilcoxon signed-ranks tests were conducted as two-tailed tests (Felsenstein, 1985b). Tests were conducted using PAUP\* (Swofford, 2001), which incorporates a correction for tied ranks.

### 3. Results

#### 3.1. Alignment

Mitochondrial genomic sequences range in size from 16,196 to 17,423 nucleotides. Nearly all nucleotides coding for transfer RNAs or proteins are deemed alignable. In tRNA genes, the D-loop is excluded from *trnI*, *A*, *N*, *Y*, and *H*. The D-arm replacement loop and V(variable)-loop are excluded from *trnSI*. In *trnLI* the D- and V-loops are excluded. The D- and T-loops are excluded from *trnF*, *V*, *M*, *W*, *C*, *D*, *K*, *R*, *E*, *T*, and *P*. This also included the D-stem in *trnC* because some taxa contain a D-arm replacement loop. The D-, V-, and T-loops are excluded from *trnG*. This resulted in the exclusion of 271 positions among tRNA genes. In protein coding regions, a total of 258 amino acid positions (774 nucleotide sites) are excluded as follows with the number in parentheses representing the number of omitted aligned amino acid positions for each gene: *nad1* (21), *nad2* (12), *cox1* (7), *cox2* (3), *atp8* (18), *atp6* (2), *nad4* (21), *nad5* (60), *nad6* (101), and *cob* (12). Hence, a total of 1045 nucleotide positions are excluded from tRNA and protein coding regions, leaving an alignment of 11,946 positions.

#### 3.2. Phylogenetic information content

The genic regions included in the phylogenetic analysis have different levels of phylogenetic information

(Table 1). To compare regions that were sequenced for different numbers of bases, we standardized these values as the ratio of the number of informative sites over the number of bases sequenced in a region, then multiply by 1000. We term this the standardized information content (SIC). The SIC value for the total data set analyzed is 485. The tRNA genes combined, and each of the protein genes *nad1*, *cox1*, *cox2*, *cox3*, and *cob* have SIC values below that for the total data set, while *nad2*, *atp8*, *atp6*, *nad3*, *nad4L*, *nad4*, *nad5*, and *nad6* have SIC values higher than that for the total data set. The transfer RNA genes collectively have the lowest SIC value of 352 while contributing 8% of the total phylogenetic information. *Nad4L* has the highest SIC value of 621 while only contributing 3.1% of the total phylogenetic information. It is interesting that the protein encoding genes that are generally the most well conserved (e.g., *cox1* and *cob*), the most amenable to study by PCR amplification based on having well-conserved portions for designing primers, and that are most commonly employed in phylogenetic studies that use only one or a few genes, are among the lowest in SIC. Perhaps this is due to their having high among-site rate variation, i.e. a low fraction of their sites are free to change, so that undetected multiple substitutions occur at some sites, while many other sites remain constant, hence reducing phylogenetic signal.

#### 3.3. Phylogenetic relationships

The alignment of 11,946 nucleotide positions contains 5797 that are phylogenetically informative. A single most parsimonious tree is produced and all phylogenetic relationships are well supported (Fig. 4). Each branch of the tree has bootstrap support of 100% and very high decay indices with the exception of the monophyly of the Amphisbaenidae (*Amphisbaena* and *Geocalamus*) with a

Table 1  
Phylogenetic information by genic region in the common vertebrate gene order

	Number of characters	Non-informative variable sites	Phylogenetically informative sites	Percent of total informative sites (%)	Standardized information content (SIC) per kilobase <sup>a</sup>
<i>tRNAs</i>	1321	195	465	8.0	352
<i>nad1</i>	906	70	428	7.4	472
<i>nad2</i>	1002	111	585	10.1	584
<i>cox1</i>	1533	91	571	9.8	372
<i>cox2</i>	678	62	313	5.4	462
<i>atp8</i>	104	11	58	1.0	558
<i>atp6</i>	677	52	381	6.6	563
<i>cox3</i>	786	60	328	5.7	417
<i>nad3</i>	347	34	187	3.2	539
<i>nad4L</i>	290	32	180	3.1	621
<i>nad4</i>	1317	136	735	12.7	558
<i>nad5</i>	1644	169	943	16.3	574
<i>nad6</i>	219	32	119	2.0	543
<i>cob</i>	1122	119	504	8.7	449
Total	11,946	1174	5797	100	485

<sup>a</sup> Calculated as the ratio of phylogenetically informative positions and the total number of characters in a data partition, standardized per kilobase by multiplying by 1000.

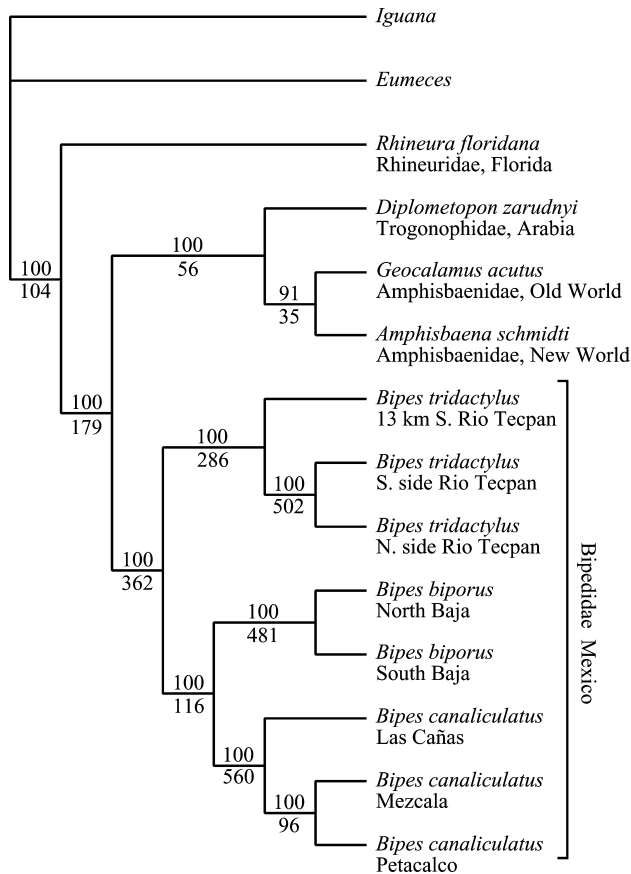


Fig. 4. A single most parsimonious tree is produced from analysis of the 11,946 aligned nucleotide positions containing 5797 phylogenetically informative sites from the 13 protein coding and 22 tRNA genes. The tree has a length of 20,104 steps. Bootstrap values appear above and decay indices below branches. Note that Gondwanan taxa, Trogonophidae (*Diplometapton*), and Amphisbaenidae (*Amphisbaena* and *Geocalamus*), are monophyletic (hypothesis D in Fig. 3). The two Laurasian families Rhineuridae (*Rhineura*) and Bipedidae (*Bipes*) represent successive, paraphyletic basal lineages indicating that amphisbaenian reptiles predate the breakup of Pangaea 200 MYBP.

bootstrap of 91% and a decay index of 35. In our sampling, the monophyly of Amphisbaenia receives further support from a decay index of 104. The Rhineuridae (*Rhineura*) is the sister taxon to the remaining taxa (decay index 179). The Trogonophidae (*Diplometapton*) and Amphisbaenidae are monophyletic (decay index of 56) and in a sister position to a monophyletic Bipedidae (*Bipes*, decay index of 362). Within the Bipedidae, *B. tridactylus* is monophyletic (decay index of 286) and in a sister position to a clade containing *B. canaliculatus* and *B. biporus* populations (decay index of 116). In *B. tridactylus*, the two populations on the north and south banks of the Rio Tecpan are monophyletic (decay index of 502). *B. biporus* is monophyletic (decay index of 481) as are *B. canaliculatus* populations (decay index of 560). Within *B. canaliculatus* the two populations from the Rio Balsas (Mezcala and Petacalco) are monophyletic (decay index of 96).

### 3.4. Statistical evaluation of alternative hypotheses

The phylogenetic results provide an area cladogram and a framework to examine the evolution of dentition and limbs in amphisbaenian reptiles. The Wilcoxon-signed-ranks test (Felsenstein, 1985b; Templeton, 1983) is applied to compare the most parsimonious tree from these nucleotide sequences with alternative hypotheses (Fig. 3).

- (1) Trogonophids have acrodont dentition, but the family is nested within taxa having pleurodont dentition (Amphisbaenidae, Bipedidae, and Rhineuridae), leading to the inference that the acrodont condition is derived for this lineage. The shortest alternative tree that unites taxa with pleurodont dentition (Fig. 3A) requires 312 extra steps and is rejected in favor of the unconstrained shortest tree ( $n = 1030$ ,  $Z = 9.7216$ ,  $P < 0.0001$ ).
- (2) Bipedids have forelimbs, but the family is nested within limbless taxa, leading to the inference that loss of limbs occurred multiple times, minimally twice, once in the Rhineuridae and once in the common ancestor of Trogonophidae and Amphisbaenidae. [The alternative, that limbs were lost once in the common ancestor of amphisbaenians, then regained in bipedids, seems very unlikely, but see Greene and Cundall (2000)]. The shortest alternative tree that unites limbless taxa as monophyletic (Fig. 3B) requires 187 extra steps and is rejected in favor of the unconstrained shortest tree ( $n = 523$ ,  $Z = 8.1769$ ,  $P < 0.0001$ ).
- (3) The Gondwanan Amphisbaenidae and Trogonophidae appear as monophyletic in the unconstrained shortest tree (Fig. 3D), with the Laurasian Rhineuridae and Bipedidae as paraphyletic. The shortest alternative tree that unites Laurasian taxa as monophyletic (Fig. 3C) requires 179 extra steps and is rejected in favor of the unconstrained shortest tree ( $n = 531$ ,  $Z = 7.7679$ ,  $P < 0.0001$ ).
- (4) The unconstrained shortest tree unites the northern *Bipes* species (*B. biporus* and *B. canaliculatus*) to the exclusion of *B. tridactylus* rather than phylogenetically separating the mainland Mexican species from those on the Baja peninsula (Fig. 3F). The shortest alternative tree that unites mainland Mexican *Bipes*, each of which has reduced numbers of digits, as monophyletic (Fig. 3E) requires 116 extra steps and is rejected in favor of the unconstrained shortest tree ( $n = 448$ ,  $Z = 5.4805$ ,  $P < 0.0001$ ).

### 3.5. Mitochondrial structural features

Amphisbaenian mtDNAs have several atypical structural features. Relative to the gene arrangement

commonly found for vertebrate mtDNAs, and most parsimoniously the ancestral condition for the Amphisbaenia, the Rhineuridae has the block *nad6*, *trnE* switched in order with the block *cob*, *trnT*, *trnP* (Fig. 5). This gene order is the same as that observed in all birds sampled to date, but not their sister taxon, the crocodylians, and so represents the first case of convergent gene order evolution found in comparisons of complete mtDNA sequences. [Some have suggested that there is a convergent gene rearrangement within the birds (Mindell et al., 1998), but only non-coding regions have moved; all studied bird mtDNAs actually have the same arrangement of all genes.]

All species of the Bipedidae share the synapomorphy of having *nad6* and *trnE* switched in order. There is a small region of non-coding sequence between *nad6* and *cob*, where *trnE* typically is found among vertebrate mtDNAs, perhaps as a vestige of the original *trnE* that formerly resided there. Interestingly, this sequence is conserved in length with *B. tridactylus* and *B. canaliculatus* populations having sequences of 72 bases while the northern *B. biporus* has 77 bases and the southern *B. biporus* has 79 bases. This is within the range typically observed for tRNA genes, but there is no obvious sequence similarity to *trnE*. In addition, *Bipes biporus*

has a tandem duplication of the block *trnT*, *trnP*, but with some appearing to be pseudogenes. The pattern of pseudogene formation is different between the two *B. biporus* populations with the southern sample retaining function of *trnT1* and *trnP2* and the northern sample retaining function of *trnT2* and *trnP2*. In all Amphisbaenia except the Rhineuridae, *trnC* encodes a transfer RNA that lacks a D-stem and instead contains a D-arm replacement loop (Macey et al., 1997b).

Species of the Rhineuridae, Amphisbaenidae, and Trogonophidae retain the potential for a strong stem-loop structure between *trnN* and *trnC* where light-strand replication is thought to usually initiate for vertebrate mtDNAs, whereas members of the Bipedidae have a short atypical stem-loop at this position. In particular, the 3'-GCC-5' heavy-strand template sequence identified as the point of light-strand elongation in mouse (Brennicke and Clayton, 1981) is not present in these structures for the Bipedidae. *B. tridactylus* has an eight base stem and the other two species have a seven base stem, which is less than that normally found among squamate reptiles and these stem regions show little base compositional similarity with that observed across other squamates (Macey et al., 1997a).

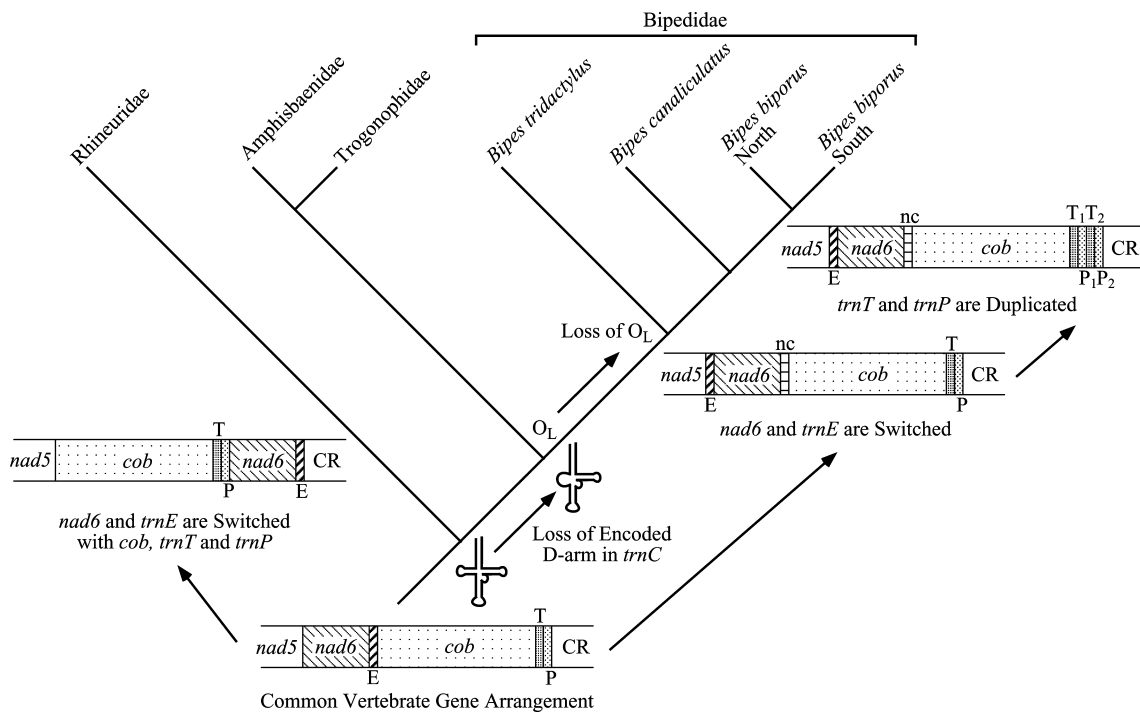


Fig. 5. Evolution of mtDNA structural features among amphisbaenian reptiles. At the base of the tree, the gene arrangement commonly observed between *nad5* and the Control Region (CR) among vertebrates is depicted. The Rhineuridae has the block *nad6*, *trnE* switched in order with *cob*, *trnT*, *trnP*, which is convergent with that observed in birds. The Bipedidae has the synapomorphic trait of having *nad6* and *trnE* switched in order. Also there is a small region of non-coding sequence (nc) between *nad6* and *cob* where *trnE* typically is found. *B. biporus* has a tandem duplication of the block *trnT*, *trnP*. The pattern of pseudogene formation is different in the two *B. biporus* populations (see text). In all amphisbaenians except the Rhineuridae, *trnC* encodes a transfer RNA that lacks a D-stem and instead contains a D-arm replacement loop. The Bipedidae has an atypical stem and loop structure between *trnN* and *trnC* where light-strand replication usually initiates among vertebrates.

## 4. Discussion

### 4.1. Historical biogeography

Our phylogenetic analysis shows that taxa found in Laurasian regions (Rhineuridae and Bipedidae) are basal and that the more broadly distributed Gondwanan groups (Amphisbaenidae and Trogonophidae) are nested within. This indicates that the most basal divergences of Amphisbaenia predate the geological split of Pangaea 200 MYBP. This event is likely responsible for the isolation of the ancestor of Amphisbaenidae and Trogonophidae in Gondwana.

Old and New World Amphisbaenidae, represented by *Geocalamus* and *Amphisbaena*, respectively, are monophyletic, so that this family dates back at least to the opening of the Atlantic Ocean 80 MYBP. Hence, the divergence between the sister taxa Amphisbaenidae and Trogonophidae in the Old World must have occurred even earlier.

The earliest divergence in the family Bipedidae from Mexico is between northern and southern species rather than between mainland and peninsular. The transversional distance between Old and New World Amphisbaenidae representing the opening of the Atlantic Ocean 80 MYBP is 12.6%. If transversions are accumulating substitutions in a linear fashion then the average transversional distance observed between *B. tridactylus* and the other two *Bipes* species of 10.9% is approximately equivalent to 69 MYBP. This is consistent with known Cretaceous to Paleocene tectonic activity along the west coast of Mexico (Ferrari, 1995) and inconsistent with the tectonic split of Baja California (12–14 MYBP; Ferrari, 1995). The transversional distance between the sister taxa *B. canaliculatus* from the Mexican mainland and *B. biporus* from peninsular Baja California is 8.6% which corresponds to a divergence 55 MYBP. The Cape Region of Baja California separated from Jalisco on the Mexican mainland 12–14 MYBP (Ferrari, 1995) far to the north of *B. canaliculatus*. Therefore, the divergence of *B. canaliculatus* and *B. biporus* is likely to be related to early tectonic activity as Baja California slid northward along the west coast of Mexico prior to oceanic separation.

### 4.2. Limb evolution

The Bipedidae, with two forelimbs, is nested within the remaining limbless taxa. Because the ancestor must have had four legs, the most parsimonious reconstruction requires independent loss of limbs in the Rhineuridae, and the lineage leading to the Amphisbaenidae and Trogonophidae. A less likely alternative is that the forelimbs in the Bipedidae have arisen secondarily, perhaps as an atavism mediated by genes regulating leg morphogenesis, such as those of the hox family (reviewed in Greene and Cundall, 2000).

Species within the Bipedidae vary in the number of digits on their forelimbs. *B. tridactylus*, with three digits, is the sister taxon to *B. canaliculatus* and *B. biporus*. While *B. biporus* always has five digits, *B. canaliculatus* is variable, usually having four, but sometimes five, digits (Papenfuss, 1982). It appears that there has been independent reduction of digits in *B. tridactylus* and *B. canaliculatus* from an ancestor with five digits.

### 4.3. Evolution of the amphisbaenian mitochondrial genome

Mitochondrial gene rearrangement and duplication has occurred in separate lineages of amphisbaenians, the Rhineuridae and Bipedidae (Fig. 5). Considering their distant placement on the phylogenetic tree, it appears that these two lineages acquired novel mitochondrial gene orders independently via different mechanisms, presumably involving replicational errors. However, surprisingly, these events occurred in the same genomic region.

Mitochondrial gene rearrangements seem to commonly evolve through tandem duplication of genes followed by loss of extra copies (Boore, 2000; Moritz et al., 1987). For faithful replication of a circular genome, termination must be made to exactly coincide with the leading strand's initiation point, whereas duplications can result from that strand's extension beyond an alternative termination (see Boore, 2000; Macey et al., 1997a). In the mtDNAs of vertebrates, the origin of replication for the heavy-strand ( $O_H$ ) is thought to consistently occur in the Control Region, and the origin of replication for the light-strand ( $O_L$ ) occurs in a stem-loop structure between *trnN* and *trnC*, formed once the growing, nascent heavy-strand reaches that region. Gene rearrangements stemming from errors at either origin are then expected to be limited to those genes immediately downstream of the Control Region or the  $O_L$ . Previous work has cautioned that rearrangements of genes flanking the replication origins might be prone to homoplasy due to this simple predisposing mechanism (Boore and Brown, 1998).

There are few examples to date of rearrangements among vertebrate mitochondrial genomes, and the striking observation has been made (Macey et al., 1997a) that disparate lineages that have gene rearrangements are lacking a recognizable stem-loop for  $O_L$ . This correlation holds for the mtDNAs of sea lamprey (*Petromyzon marinus*; Lee and Kocher, 1995), birds (Desjardins and Morais, 1990), crocodilians (Kumazawa and Nishida, 1995), tuatara (*Sphenodon punctatus*; Rest et al., 2003), acrodont lizards (reviewed in Macey et al., 2000), and the Texas blind snake (*Leptotyphlops dulcis*; Kumazawa and Nishida, 1995). As an alternative to the stem-loop at  $O_L$ , other structures formed in the displaced strand may potentially serve this function (Macey et al., 1997a),



including those of tRNA genes (Clayton, 1982). It may then be possible that having alternative origins, perhaps with suboptimal regulatory signals, increases the likelihood of gene duplications followed by gene rearrangements. These rearrangements could potentially occur at various sites in the mtDNA depending on the location of the alternate  $O_L$ . It can be reasoned that the position of  $O_L$  would be much more labile than that of  $O_H$  since the latter regulates replication initiation, controls D-loop formation, and serves a dual role as a transcription initiation site (Clayton, 1991). In addition, light-strand synthesis requires heavy-strand replication to produce a single stranded stem-loop template typically 11 kb from  $O_H$ , therefore it could potentially initiate at alternative sites along this strand throughout the mitochondrial genome.

The mitochondrial gene order observed in *Rhineura* is identical to that of birds, where the block *nad6*, *trnE* is switched in order with *cob*, *trnT*, *trnP* relative to the arrangement commonly found for vertebrates. This is the first parallelism of gene rearrangement fully documented by complete mitochondrial genome sequences. To our knowledge, there have been only two cases (Dowton and Austin, 1999; Flook et al., 1995) of purported parallelism suggested even from partial sequences. Unlike all other amniotes that have rearranged mitochondrial gene orders, including birds, *Rhineura* has an apparently functional  $O_L$  in the typical location. Considering this, and that the rearranged genes flank the Control Region, the gene rearrangement observed in *Rhineura* appears to be derived from errors of heavy—rather than light-strand replication.

Surprisingly, the gene rearrangement in *Bipes* involves the same gene region. However, we suggest that this gene rearrangement, where *trnE* and *nad6* are switched in order, is produced by error at a light-strand replicational origin. This taxon lacks a recognizable stem-loop at the typical location of  $O_L$ , so some other region(s) of the mtDNA must perform this function. *Bipes* appears to be yet another case having the loss of  $O_L$  for its typical location accompanying gene rearrangement.

The typical  $O_L$  has functionally important, overlapping regions with the adjacent gene encoding tRNA(C) (reviewed in Macey et al., 1997a). All sampled amphibaenians except *Rhineura* have a tRNA(C) lacking a D-stem and instead containing a D-arm replacement loop (Macey et al., 1997b). Macey et al. (1997c) proposed a cascade model suggesting that changes to the secondary structure of *trnC* may destroy function of  $O_L$  in its typical vertebrate location which, in turn, may facilitate shifts in gene order by enabling the utility of alternative sites for initiation and termination of light-strand synthesis (see above). It seems to be reasonable that *Bipes* constitutes an additional example in support of this model.

Additionally, *B. bipes* mtDNA contains a tandem duplication of *trnT* and *trnP*. We cannot confidently spec-

ulate whether this has been induced by light-strand or heavy-strand replicational error because, on the one hand, all *Bipes* have an atypical stem-loop structure where  $O_L$  commonly is located, but on the other hand, this duplication is adjacent to the Control Region. Interestingly, although the two *B. bipes* sampled retain function of different copies of *trnT*, no gene rearrangement will occur in either population because both taxa retain function of the second copy of *trnP* which is preceded by both copies of *trnT*. In order for a gene rearrangement to result from a tandem duplication of two genes, non-adjacent copies of both genes must lose function (Macey et al., 1997a), which, in this case, is the first copy of *trnT* and the second copy of *trnP* (Macey et al., 1998).

While heavy-strand induced gene rearrangements are restricted to the vicinity of the Control Region, as observed for large, tandem duplications among lizards (Moritz and Brown, 1986, 1987), light-strand induced rearrangements can occur throughout the mitochondrial genome. Hence, mitochondrial genomic rearrangements caused by heavy-strand replication are less reliable phylogenetic characters than rearrangements produced by light-strand replication, which are known to be located throughout the genome and identified by loss of a recognizable  $O_L$  in the typical vertebrate location (Macey et al., 1997a).

## Acknowledgments

Gabriela Parra Olea and SEMARNAP provided tissue samples and permits in Mexico. We also thank Paul Moler and Richard Thomas for providing tissue samples and Jonathan Fong for assistance with GenBank submissions. Karen Klitz prepared Figs. 1 and 2. This work is LBNL-53616 and was performed under the auspices of the US Department of Energy, Office of Biological and Environmental Research, under Contract No. DE-AC03-76SF00098 with the University of California, Lawrence Berkeley National Laboratory.

## Appendix A. Supplementary material

Supplementary data associated with this article containing the aligned DNA sequences as a Nexus file, can be found, in the online version, at [doi:10.1016/j.ympcv.2004.05.003](https://doi.org/10.1016/j.ympcv.2004.05.003).

## References

- Boore, J.L., 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767–1780.
- Boore, J.L., 2000. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In: Sankoff, D., Nadeau, J. (Eds.), *Comparative*

- Genomics, Computational Biology Series, vol. 1. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 133–147.
- Boore, J.L., Brown, W.M., 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Curr. Opin. Genet. Dev.* 8, 668–674.
- Bremer, K., 1994. Branch support and tree stability. *Cladistics* 10, 295–304.
- Brennicke, A., Clayton, D.A., 1981. Nucleotide assignment of alkali-sensitive sites in mouse mitochondrial DNA. *J. Biol. Chem.* 256, 10613–10617.
- Clayton, D.A., 1982. Replication of animal mitochondrial DNA. *Cell* 28, 693–705.
- Clayton, D.A., 1991. Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol.* 7, 453–478.
- Dean, F.B., Nelson, J.R., Giesler, T.L., Lasken, R.S., 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle. *Genome Res.* 11, 1095–1099.
- Desjardins, P., Morais, R., 1990. Sequence and gene organization of the chicken mitochondrial genome. *J. Mol. Biol.* 212, 599–634.
- Dowton, M., Austin, A.D., 1999. Evolutionary dynamics of a mitochondrial rearrangement “hot spot” in the Hymenoptera. *Mol. Biol. Evol.* 16, 298–309.
- Elkin, C., Kapur, H., Smith, T., Humphries, D., Pollard, M., Hammon, N., Hawkins, T., 2002. Magnetic bead purification of labeled DNA fragments for high-throughput capillary electrophoresis sequencing. *Biotechniques* 32, 1296–1302.
- Felsenstein, J., 1985a. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Felsenstein, J., 1985b. Confidence limits on phylogenies with a molecular clock. *Syst. Zool.* 34, 152–161.
- Ferrari, L., 1995. Miocene shearing along the northern boundary of the Jalisco block and the opening of the southern Gulf of California. *Geology* 23, 751–754.
- Flook, P., Rowell, H., Gellissen, G., 1995. Homoplastic rearrangements of insect mitochondrial tRNA genes. *Naturwissenschaften* 82, 336–337.
- Gans, C., 1968. Relative success of divergent pathways in amphisbaenian specialization. *Am. Nat.* 102, 345–362.
- Gans, C., 1978. The characteristics and affinities of the Amphisbaenia. *Trans. Zool. Soc. Lond.* 34, 347–416.
- Greene, H.W., Cundall, D., 2000. Limbless tetrapods and snakes with legs. *Science* 287, 1939–1941.
- Hawkins, T.L., Detter, J.C., Richardson, P.M., 2002. Whole genome amplification-applications and advances. *Curr. Opin. Biotechnol.* 13, 65–67.
- Ingman, M., Kaessmann, H., Paabo, S., Gyllensten, U., 2001. Mitochondrial genome variation and the origin of modern humans. *Nature* 408, 708–713.
- 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. *Proc. R. Soc. Lond. B Biol. Sci.* 268, 623–631.
- Kearney, M., 2003. Systematics of the Amphisbaenia (Lepidosauria: Squamata) based on morphological evidence from recent and fossil forms. *Herpetol. Monogr.* 17, 1–74.
- Kumazawa, Y., Nishida, M., 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* 37, 380–398.
- Kumazawa, Y., Nishida, M., 1995. Variations in mitochondrial tRNA gene organization of reptiles as phylogenetic markers. *Mol. Biol. Evol.* 12, 759–772.
- Kumazawa, Y., Nishida, M., 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Mol. Biol. Evol.* 16, 784–792.
- Lee, W.-J., Kocher, T.D., 1995. Complete sequence of a Sea Lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. *Genetics* 139, 873–887.
- Macey, J.R., Larson, A., Ananjeva, N.B., Papenfuss, T.J., 1997b. Replication slippage may cause parallel evolution in the secondary structures of mitochondrial transfer RNAs. *Mol. Biol. Evol.* 14, 30–39.
- Macey, J.R., Larson, A., Ananjeva, N.B., Papenfuss, T.J., 1997c. Evolutionary shifts of three major structural features in the mitochondrial genome among iguanian lizards. *J. Mol. Evol.* 44, 660–674.
- Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., Papenfuss, T.J., 1997a. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Mol. Biol. Evol.* 14, 91–104.
- Macey, J.R., Schulte II, J.A., Larson, A., 2000. Evolution and phylogenetic information content of mitochondrial genomic structural features illustrated with acrodont lizards. *Syst. Biol.* 49, 257–277.
- Macey, J.R., Schulte II, J.A., Larson, A., Papenfuss, T.J., 1998. Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. *Mol. Biol. Evol.* 15, 71–75.
- Macey, J.R., Verma, A., 1997. Homology in phylogenetic analysis: alignment of transfer RNA genes and the phylogenetic position of snakes. *Mol. Phylogenet. Evol.* 7, 272–279.
- Maddison, W.P., Maddison, D.R., 2001. *MacClade, Analysis of Phylogeny and Character Evolution, Version 4.03*. Sinauer, Sunderland, MA.
- Mindell, D.P., Sorenson, M.D., Dimcheff, D.E., 1998. Multiple independent origins of mitochondrial gene order in birds. *Proc. Natl. Acad. Sci. USA* 95, 10693–10697.
- Moritz, C., Brown, W.M., 1986. Tandem duplications of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. *Science* 233, 1425–1427.
- Moritz, C., Brown, W.M., 1987. Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. *Proc. Natl. Acad. Sci. USA* 84, 7183–7187.
- Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* 18, 269–292.
- Papenfuss, T.J., 1982. The ecology and systematics of the amphisbaenian genus *Bipes*. *Occ. Pap. Calif. Acad. Sci.* 136, 1–42.
- Rest, J.S., Ast, J.C., Austin, C.C., Waddell, P.J., Tibbetts, E.A., Hay, J.M., Mindell, D.P., 2003. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Mol. Phylogenet. Evol.* 29, 289–297.
- Swofford, D.L., 2001. *PAUP\* Phylogenetic Analysis Using Parsimony (\* and Other Methods)*, Beta Version 4.0b8. Sinauer, Sunderland, MA.
- Templeton, A.R., 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37, 221–244.