# Multiple colonization of Madagascar and Socotra by colubrid snakes: evidence from nuclear and mitochondrial gene phylogenies

**Zoltán Tamás Nagy**[1*]**, Ulrich Joger**[2]**, Michael Wink**[1]**, Frank Glaw**[3]
**and Miguel Vences**[4]

[1]*Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany*
[2]*State Natural History Museum Braunschweig, Pockelsstrasse 10, 38106 Braunschweig, Germany*
[3]*Zoologische Staatssammlung, Münchhausenstrasse 21, 81247 München, Germany*
[4]*Institute for Biodiversity and Ecosystem Dynamics, Zoological Museum, University of Amsterdam, PO Box 94766, 1090 GT Amsterdam, The Netherlands*

Colubrid snakes form a speciose group of unclarified phylogeny. Their almost cosmopolitan distribution could be interpreted as a product of plate-tectonic vicariance. We used sequences of the nuclear *c-mos*, the mitochondrial cytochrome *b* and the 16S rRNA genes in 41 taxa to elucidate the relationships between the endemic colubrid genera found in Madagascar and in the Socotra archipelago. The well-resolved trees indicate multiple origins of both the Malagasy and the Socotran taxa. The Malagasy genus *Mimophis* was nested within the Psammophiinae, and the Socotran *Hemerophis* was closely related to Old World representatives of the former genus *Coluber*. The remaining 14 genera of Malagasy colubrids formed a monophyletic sister group of the Socotran *Ditypophis* (together forming the Pseudoxyrhophiinae). Molecular-clock estimates place the divergence of Malagasy and Socotran colubrids from their non-insular sister groups into a time-frame between the Eocene and Miocene. Over-seas rafting is the most likely hypothesis for the origin of at least the Malagasy taxa. The discovery of a large monophyletic clade of colubrids endemic to Madagascar indicates a need for taxonomic changes. The relationship of this radiation to the Socotran *Ditypophis* highlights the potential of the Indian Ocean islands to act as an evolutionary reservoir for lineages that have become extinct in Africa and Asia.

**Keywords:** Madagascar; Socotra; colonization; molecular phylogeny; Serpentes: Colubridae; Pseudoxyrhophiinae

## 1. INTRODUCTION

The origin of the endemic terrestrial vertebrate faunas of non-oceanic islands that at some point in geological history had a connection to other continental land masses is a subject of continuing debate. Different opinions exist as to whether faunas such as those of the Caribbean archipelagos have mainly been shaped by vicariance or dispersal (e.g. Crother & Guyer 1996; Hedges 1996). A particularly enigmatic area in this context is the Malagasy region. It comprises the fourth largest island in the world, Madagascar, as well as the continental Seychelles and the volcanic (and thus oceanic) Comoros, Aldabra, Mascarenes and several other small islands. Located off the southeastern African coast, and separated from the continent by sea depths of over 4000 m, Madagascar is characterized by diverse and highly endemic amphibian, reptilian and mammalian faunas, and by a freshwater fish assemblage that is depauperate in major lineages. Madagascar was part of the Mesozoic supercontinent Gondwana and separated from Africa 165–121 Myr ago and from India and the Seychelles 88–63 Myr ago (Rabinowitz *et al.* 1983; Storey *et al.* 1995).

A further archipelago in the Indian Ocean is Socotra, located off the East African coast and separated from it by ocean depths of only a few hundred metres. This continental fragment, politically belonging to Yemen, has long been neglected in attempts to reconstruct the biogeographical history of the Indian Ocean, and the affinities of its considerable endemic reptile fauna remained virtually unstudied until the end of the twentieth century (Joger 2000). Its separation age is under debate. Kopp (1999) considers isolation from Africa for at least 70 Myr, but minimal estimates are *ca.* 10 Myr (Laughton *et al.* 1970). Krupp *et al.* (2002) hypothesize that the archipelago was separated from Africa 'at about the same time as India and Madagascar'.

An origin by ancient vicariance has been invoked for Madagascar's freshwater fishes and frogs (Richards & Moore 1996; Murphy & Collier 1997; Farias *et al.* 2001; but see Vences *et al.* 2001). By contrast, dispersalist explanations prevail for mammals (e.g. Jansa *et al.* 1999) and reptiles (e.g. Caccone *et al.* 1999; Mausfeld *et al.* 2000; Palkovacs *et al.* 2002; Raxworthy *et al.* 2002; see also Arnold 2000). Modern snakes (Colubroidea) are not known in the world's fossil record prior to the Eocene (Rage 1987, 1988). However, their record is mainly based on holarctic sites; a noteworthy palaeontological record is absent from tropical Africa and Madagascar (Krause *et al.*

*Proc. R. Soc. Lond.* B (2003) **270**, 2613–2621
DOI 10.1098/rspb.2003.2547

2613

© 2003 The Royal Society

1997). However, considering their almost cosmopolitan distribution, a much more ancient origin prior to the major continental breakups in the Mesozoic may be hypothesized. Because of the complete lack of colubroid fossils from Madagascar, hypotheses on the origins of its colubrid fauna must rely on phylogenetic reconstructions of recent taxa.

Colubrid snakes are a highly diverse group of currently about 1800 species (Zaher 2000), thus comprising the vast majority of snakes. Their subfamilial taxonomy is under debate and will finally depend on molecular-based phylogenetic reconstructions, which are likely to render the traditional family Colubridae a paraphyletic assemblage (Kelly *et al.* 2003). Fossils of the Colubridae are known since the Oligocene (Carroll 1988), but not from Africa, Socotra or Madagascar. A rich extant colubrid fauna with 18 genera and about 75 nominal species has been described from Madagascar, but only two genera with two endemic species, closely related to Malagasy taxa, are found in the Comoros archipelago (Glaw & Vences 1994). Socotra harbours only two species of colubrid that both belong to endemic monotypic genera: *Hemerophis* (formerly *Coluber*) *socotrae* and *Ditypophis vivax* (Joger 2000). On the Seychelles two endemic species exist, *Lamprophis geometricus* and *Lycognathophis seychellensis* (Nussbaum 1984), the latter also belonging to a monotypic genus. No endemic colubrids are known in the Mascarenes, Aldabra or any of the small oceanic islands of the region (Arnold 2000). In a pan-African perspective, colubrids are absent from some truly oceanic (volcanic) islands, such as the Canary and Cabo Verde archipelagos, which could be interpreted as indicative of a limited capacity for oceanic dispersal (Nussbaum 1984).

In this study, we obtained DNA sequences of one nuclear and two mitochondrial genes of the two Socotran genera and of 15 Malagasy genera to address four questions.

(i) Are the Malagasy and Socotran colubrids monophyletic lineages?
(ii) Which are their closest living relatives on the African or Asian mainland?
(iii) Are there any conspicuous relationships between Malagasy and Socotran taxa?
(iv) Which biogeographical scenario is the most likely to explain the patterns observed?

## 2. MATERIAL AND METHODS

### (a) *Taxon sampling*

The wealth of worldwide colubrid diversity poses difficulties when deciding which species to include in a phylogenetic analysis. Previous studies have shown no conspicuous affinities between Malagasy colubrids and South American lineages (Dowling *et al.* 1983; Heise *et al.* 1995; Kraus & Brown 1998; Vidal *et al.* 2000; Slowinski & Lawson 2002) and indicated monophyly of an African clade that includes Malagasy colubrids (Vidal & Hedges 2002), similar to the situation in other reptile groups (e.g. Caccone *et al.* 1999; Mausfeld *et al.* 2000; Raxworthy *et al.* 2002). We therefore restricted our study to African and Asian species of the Colubridae, Elapidae and Atractaspididae (see electronic Appendix A, available on The Royal Society's

Publications Web site). *Boa constrictor* (Boidae) and *Bitis nasicornis* (Viperidae) were used as hierarchical outgroups. For current classifications of taxa studied, voucher specimens and GenBank accession numbers of sequences, see electronic Appendix B.

### (b) *Molecular techniques*

Three target genes known to show different evolutionary characteristics were selected for molecular phylogenetic analyses:

(i) the complete sequence of the mitochondrial protein-encoding cytochrome *b* gene (1111–1117 bp depending on indel sites and the inclusion of a terminal T);
(ii) an approximately 510 bp segment of the non-protein-coding mitochondrial 16S rRNA;
(iii) a fragment of 567 bp (colubrines, *Prosymna* and *Bitis*) or 570 bp (other snakes) of the nuclear genomic proto-oncogene *c-mos*.

Isolation of total genomic DNA from preserved blood or tail tissues followed standard protocols involving phenol–chloroform, chloroform–isoamylalcohol and ethanol washing. The fragments of interest were amplified by PCR under standard conditions. We used the PCR primers (i) L14910, L14919 and H16064 (Burbrink *et al.* 2000, with modification as described by de Queiroz *et al.* (2002)), (ii) 16SA and 16Sb (Palumbi *et al.* 1991) and (iii) S77 and S78 (R. Lawson, unpublished data) for amplification and direct sequencing of the three genes. New primers for cytochrome *b* sequencing were designed: L-410 (5′-TCNTTCTGAGCAGCAACAGT-3′), H-391 (complement and reverse of L-410) and H-739 (5′-GGGTTRGCTTTTGAGAARTT-3′). PCR products were taken as templates in cycle sequencing reactions, and then sequenced in ABI Prism 3100 (Applied Biosystems, Foster City, CA, USA) and ALFExpress II (Amersham Pharmacia Biotech, Uppsala, Sweden) automatic sequencers. Cleaning up of PCR products, cycle sequencing reactions and sequencing were performed according to the manufacturer's instructions.

In some taxa, it was necessary to re-amplify the cytochrome *b* fragments from initial amplificates. Alternatively, PCR products of the cytochrome *b* gene were cloned into *E. coli* using a TransformAid Bacterial Transformation Kit (MBI Fermentas) and pGEM-T Vector System I (Promega), and then sequenced with plasmid primers. For *Dromicodryas*, slightly different sequences were identified. We verified a mitochondrial origin for one sequence in each of the two species by (i) comparison with the results from direct sequencing of other snake taxa (e.g. constant sites should remain steady) and (ii) searching for stop codons.

### (c) *Data analysis*

The obtained sequences were aligned manually (*c-mos* and cytochrome *b*) or by using CLUSTALX 1.81 (16S rRNA; Thompson *et al.* 1997). A total of 39 bp of two hypervariable regions of the 16S rRNA gene could not be reliably aligned and were excluded from further analysis. An incongruence length difference (ILD) test applied to the combined sequences, using PAUP* 4b10 software (Swofford 2002), did not contradict the congruence of the three genes ($p = 0.914$), and we therefore entered all data in combined phylogenetic analyses.

Prior to phylogenetic reconstruction, we used hierarchical likelihood ratio tests and the Akaike information criterion as implemented in the software MODELTEST 3.06 (Posada & Crandall 1998) to discover which substitution model best fitted our data. The results of both methods indicated the GTR+I+G substitution model for the combined dataset. In most separate analyses of gene fragments, the GTR+I+G model was suggested as well.

Phylogenetic analysis was carried out with PAUP* 4b10. We calculated maximum-likelihood (ML) trees under substitution models suggested by MODELTEST, using heuristic searches with tree bisection–reconnection (TBR) branch swapping. Further trees were calculated by unweighted maximum-parsimony (MP), also using heuristic searches with TBR branch swapping, with 100 random addition sequence replicates. Gaps were coded as a fifth state in MP analyses. Bootstrapping was used to test the topology of the trees obtained. We evaluated 1000 replicates under MP and 100 replicates under ML. Furthermore, we used Bayesian inference in the program MRBAYES 2.01 (Huelsenbeck & Ronquist 2001). We ran four simultaneous Metropolis-coupled Monte Carlo Markov chains for 500 000 generations, sampling a tree every 10 generations. The number of generations needed before convergence on stable likelihood values ('burn-in') was set at 50 000 (10%), based on empirical evaluation. Two additional full repetitions of the Bayesian analysis were in excellent agreement with the first analysis (see electronic Appendix C).

Under ML models of sequence evolution, two types of statistical tests of competing phylogenetic-tree hypotheses are well established (Whelan *et al.* 2001): the SH test (Shimodaira & Hasegawa 1999) and parametric bootstrapping (Huelsenbeck *et al.* 1996). We used both approaches to evaluate the significance of non-monophyly of the Socotran and Malagasy taxa in our phylogeny. The SH test is heavily dependent on the number of topologies included in the analysis, and any selection of trees to compare with the preferred ML tree (Shimodaira & Hasegawa 1999) brings a risk of bias. For this reason, we applied SH tests to a set of only seven taxa, which is the highest number for which all possible unrooted trees ($n = 954$) can be computed and compared in a reasonable time-frame. Taxa were selected in a three-step procedure based on the ML phylogram of the complete dataset:

(i) we included representatives of each Malagasy and Socotran clade, and of their respective sister groups;

(ii) if clades consisted of more than one species, we chose the one with the shortest terminal branch, assuming that it had accumulated fewer autapomorphies that could mask its phylogenetic affinities;

(iii) *Boa constrictor* was included as the outgroup.

We used exhaustive searches under MP to find all possible unrooted trees for this set of taxa, and compared all 954 trees simultaneously with SH tests under resampling estimated log-likelihood optimization with 1000 bootstrap replicates. Parametric bootstrapping is a method used to test the strength of a computed phylogenetic arrangement (real data tree) by generating an alternative null hypothesis (constrained tree) and comparing both with a number of simulated datasets, resulting in a number of constrained and unconstrained trees. We followed the MP approximation of a likelihood evaluation suggested by Ruedi *et al.* (1998). ML searches were performed with the constraint of monophyly of Malagasy or Socotran taxa, or both, and the trees obtained were used to generate 100 sequence matrices

using SEQ-GEN (Rambaut & Grassly 1997). Heuristic searches under MP were carried out with and without constraint on these matrices, and the step difference between the most parsimonious constrained tree and the most parsimonious unconstrained tree was computed for each of these simulated datasets (constrained – unconstrained). These differences were plotted and compared with those from the real dataset. The null hypothesis can be rejected if the step differences between unconstrained and constrained trees in the real dataset fall well outside the range of step differences between unconstrained and constrained trees in the simulated dataset.

To estimate the age of divergence of the most relevant clades, we added a number of taxa that could be used for calibration of the molecular clock. A mammal (*Sus*) and a bird (*Gallus*) were used as further outgroups, with the root of the tree corresponding to the divergence of synapsids and diapsids at 310 Myr ago (Kumar & Hedges 1998). Two species each of the eastern and western subgroups of the colubrid genus *Hemorrhois* (*H. ravergieri* and *H. nummifer* versus *H. algirus* and *H. hippocrepis*) were added; the time of the earliest land connection between Asia and Africa across the former Tethys Sea at 16–18 Myr ago is a maximum estimate of their divergence (Joger 1987). Minimum constraints used from the fossil record were the radiations of colubrids 34 Myr ago (Rage 1987, 1988; Ivanov *et al.* 2000) and of boids 56 Myr ago (Rage 2001). Molecular-clock estimates of divergence times must take saturation effects and rate inequality among taxa into account. To at least partly overcome these problems, we first tested for saturation of different data subsets (the three genes, and first, second and third positions of cytochrome *b* and *c-mos*) by plotting p-distances against Kimura-2-parameter distances. This model of distance correction was chosen as a compromise because in more complex models we would have been unable to compute distances between all taxa. On the basis of the results of the saturation analysis (see electronic Appendix D) we excluded third positions of *c-mos* and cytochrome *b* from the analysis; for these data partitions, strong indications of progressive saturation (multiple hits at identical mutation sites) were found. Additionally, we also excluded first positions of *c-mos* from analysis to avoid problems of codon bias which have been described for this gene (Harris 2003). For the remaining data subsets (*c-mos* second positions, cytochrome *b* first and second positions, 16S) and a combined dataset of these we estimated the substitution models using MODELTEST. We then computed trees by enforcing the topological constraint of the preferred ML tree found in phylogenetic analysis, and tested for their clock-like behaviour using TREE-PUZZLE 5.0 (Schmidt *et al.* 2002). Subsequently we applied non-parametric rate smoothing (Sanderson 1997) as implemented in the r8s program to account for the rate inequality observed, and computed 95% confidence intervals of the divergence-time estimates by the cross-validation procedure (Sanderson 2002).

## 3. RESULTS

### (a) *Phylogenetic reconstruction*

Out of 2176 characters included in the combined dataset, 1022 were variable, and 795 of these were parsimony informative. MP analysis yielded six equally parsimonious trees of 6295 steps (consistency index of 0.266; retention index of 0.337), which differed in the relative positions of three taxa in a terminal clade (*Ithycyphus*, *Micropisthodon*, *Langaha*). The strict consensus of these (not shown) agreed with the ML tree (figure 1; $-\ln L = 27\ 122.4$) in
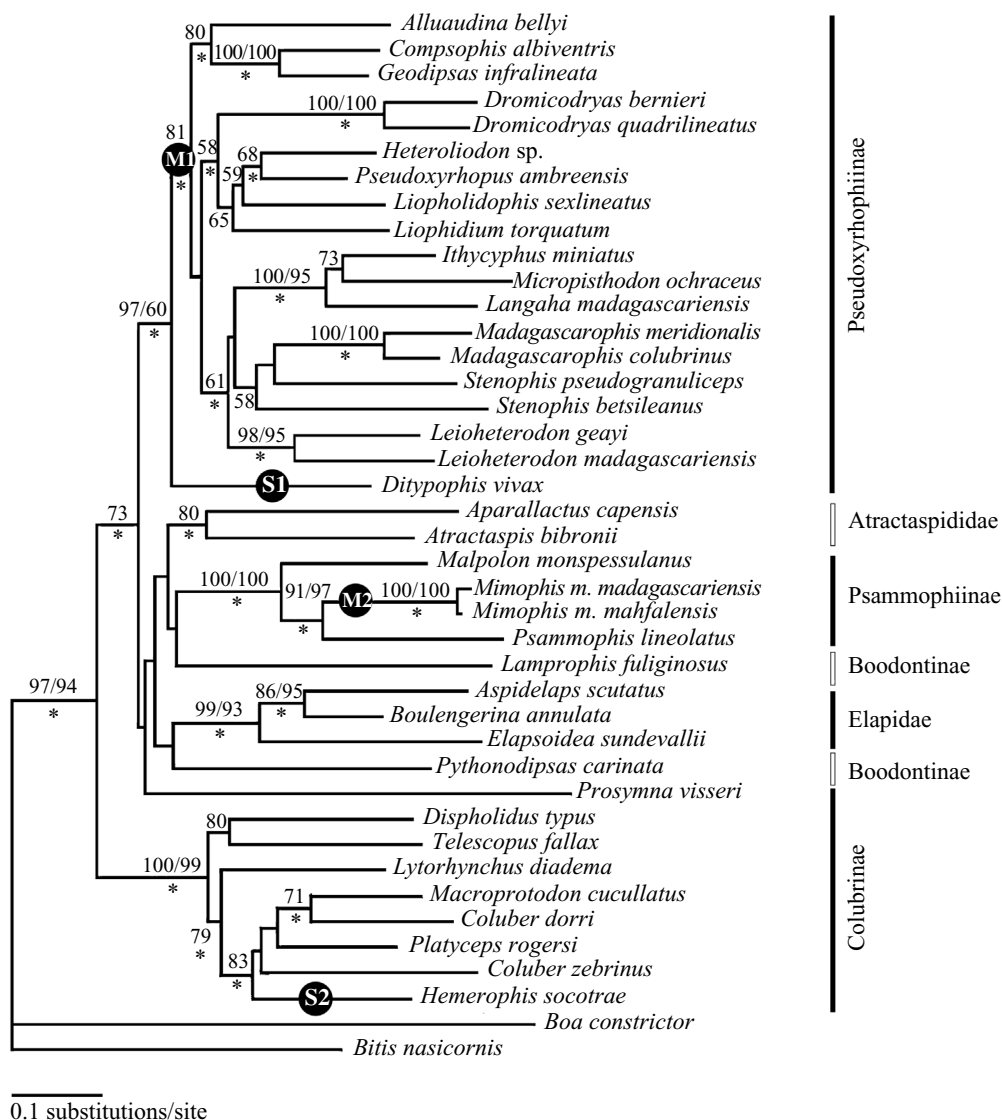
Figure 1. Maximum-likelihood phylogram (GTR+I+G model) based on combined analysis of 2176 bp of the nuclear *c-mos* and the mitochondrial cytochrome *b* and 16S rRNA genes in 40 colubroid snakes from the Viperidae, Elapidae, Atractaspididae and four colubrid subfamilies. *Boa constrictor* and *Bitis nasicornis* were used as outgroups. Numbers are bootstrap values in percentages from 100 full heuristic replicates under ML. At nodes that received bootstrap support of more than 50% in MP bootstrap analysis (1000 replicates), the second value is also given. Asterisks mark nodes with Bayesian posterior probabilities of 99–100%. Black circles mark the clades exclusive to Madagascar (M1, M2) and Socotra (S1, S2).

general aspects. Considering the results of MP and ML bootstrapping (figure 1), four major clades (i.e. clades that represent relevant taxonomic units at a subfamilial or familial level) consistently received strong support: (i) all Malagasy taxa except *Mimophis*, but with the Socotran *Ditypophis* (the Pseudoxyrhophiinae, including *Dromicodryas*, which was previously placed in the Boodontinae); (ii) the genera *Coluber*, *Dispholidus*, *Hemerophis*, *Lytorhynchus*, *Macroprotodon*, *Platyceps* and *Telescopus* (the Colubrinae, including *Macroprotodon*, which was previously placed in the Boodontinae); (iii) the psammophiines (*Malpolon*, *Psammophis* and *Mimophis*); and (iv) the elapids (*Aspidelaps*, *Boulengerina* and *Elapsoidea*). Minor clades with strong support include: within Pseudoxyrhophiinae, *Compsophis* with *Geodipsas*, as well as *Ithycyphus* with *Micropisthodon* and *Langaha*; within Psammophiinae, *Mimophis* with *Psammophis*; and within Elapidae, *Aspidelaps* with *Boulengerina*. Major disagreements between the

two analyses were found as regards the position of *Ditypophis* (sister group of the Malagasy clade except *Mimophis* in the ML analysis, nested within that clade in MP), the position of the elapids (basal to colubrids in MP, nested within colubrids in ML) and the position of *Aparallactus* (sister to *Lamprophis* in MP, sister to *Atractaspis* in ML). Bayesian posterior probabilities (figure 1 and electronic Appendix C) further supported the ML results.

Analyses of single-gene fragments were largely in accordance with the above results. The four major clades defined above were largely supported. Elapid monophyly was not recovered by ML analysis of 16S rRNA sequences. MP analysis of cytochrome *b* and ML analysis of *c-mos* placed *Ditypophis* among Malagasy pseudoxyrhophiines. The deviant topologies were not supported by relevant bootstrap values (table 1). Hence, the three datasets were congruent in suggesting double origins of both the Socotran and the Malagasy colubrid fauna.

Table 1. Results of separate ML and MP analyses of the three gene fragments, with respect to the most relevant results of the combined analysis (figure 1).

(Yes/no indicates whether a clade was recovered by the analysis; dashes indicate ambiguous results. Numbers in parentheses refer to percentage bootstrap support (MP: 1000 replicates; ML: 100 replicates, by fast stepwise addition without branch swapping; given only when greater than 50%).)

| clade | c-mos MP | 16S rRNA MP | cytochrome b MP | c-mos ML | 16S rRNA ML | cytochrome b ML |
|---|---|---|---|---|---|---|
| Pseudoxyrhophiinae[a] | yes (84%) | — | yes | yes (79%) | — | yes |
| Malagasy pseudoxyrhophiines | yes (56%) | — | no[b] | no[b] | — | yes |
| Psammophiinae including *Mimophis* | yes (71%) | yes (71%) | yes (97%) | yes (66%) | yes (64%) | yes (95%) |
| colubrine clade[c] including *Hemerophis* | yes (98%) | yes (57%) | yes (79%) | yes (99%) | yes (51%) | yes (79%) |
| Atractaspididae | yes (62%) | no | no | yes (61%) | no | yes |
| Elapidae | yes (98%) | no | yes (62%) | yes (95%) | no | yes (77%) |
| elapids nested in Colubridae | yes (86%) | — | — | yes (78%) | — | yes |
| atractaspidids nested in Colubridae | yes (86%) | — | — | yes (78%) | — | yes |

[a] The Pseudoxyrhophiinae as defined here includes *Dromicodryas* (previously Boodontinae) and *Ditypophis*.
[b] *Ditypophis* clustered within Malagasy pseudoxyrhophiines.
[c] The colubrine clade as defined here includes *Macroprotodon* (previously Boodontinae) but does not include *Prosymna* (in Colubrinae according to Zaher 2000).

## (b) *Alternative trees*

The alternative hypotheses of monophyletic origins of either the Malagasy or the Socotran taxa could be significantly excluded both by parametric bootstrapping and by SH tests in a reduced set of taxa. In the reduced set of taxa, the tree topology corresponded to the overall ML tree (figure 1; likelihood score $-\ln L = -8000.69$), with three main clades that were all supported by bootstrap values of at least 98%: *Ditypophis* being the sister group of the Malagasy *Pseudoxyrhopus*, *Mimophis* being sister to *Psammophis* and *Hemerophis* being sister to *Platyceps*. The SH test excluded most of the 953 alternative topologies (with likelihoods of $-\ln L > -8046$) at a significance level of $p < 0.05$. The 35 trees that were not significantly worse had $-\ln L$ values lower than $-8045$. They were characterized by paraphyletic arrangements of the taxa of one or two of these clades but never contained either the two Malagasy or the two Socotran taxa in an exclusive monophyletic group.

The topologies with the highest likelihoods under the constraint of monophyly of the Malagasy or Socotran clades, or both, required 43, 62 and 85 additional steps compared with the unconstrained ML topology. The parametric bootstrapping procedures under the null hypotheses returned maximum values of 20, 21 and 38 (figure 2). The null hypothesis that the observed differences may be reached or exceeded by chance alone could therefore be rejected at $p < 0.01$. Therefore, the results argue strongly against monophyly of *Mimophis* and the other Malagasy genera and against monophyly of the Socotran genera.

## (c) *Evolutionary rate and age estimates*

A likelihood ratio test indicated that the log-likelihood of the more complex (non-clock-like) tree was significantly greater than that of the clock-like tree ($p < 0.05$). The application of a rate-smoothing procedure was therefore necessary. In the following we give the estimates of divergence times (Myr ago) after non-parametric rate smoothing. The values refer to the mean of the combined analysis using non-saturated partitions of all gene fragments (figure 3). In parentheses we furthermore give the global minimum and maximum values from the estimates based on the combined dataset, of each of the separate datasets, and (if applicable) on the 95% confidence intervals of all combined and separate analyses (see electronic Appendix E for the detailed estimates): divergence between *Ditypophis* and Malagasy Pseudoxyrhophiinae, 30.8 (21.5–75.9); initial divergence among Malagasy pseudoxyrhophiines, 25.6 (17.5–62.5); divergence between *Mimophis* and *Psammophis*, 12.9 (8.2–29.5); divergence between *Hemerophis* and its sister clade, 21.8 (16.2–30.9). The three separate analyses based on conservative regions of the 16S rRNA gene and on second positions of cytochrome *b* and *c-mos* provided similar age estimates, which were all younger than the plate-tectonic separation of Africa and Madagascar (see electronic Appendix E).

## 4. DISCUSSION

Our results indicate monophyly of a clade containing atractaspidids, colubrids and elapids and excluding viperids, in agreement with the results of previous studies (Rage 1987; Heise *et al.* 1995; Dowling *et al.* 1996; Kraus & Brown 1998; Slowinski & Lawson 2002). Although not significantly supported by our data, the paraphyly of colubrids (with atractaspidids and elapids nested within this family) has also been observed previously (Cadle 1994; Dowling *et al.* 1996; Kraus & Brown 1998; Kelly *et al.* 2003) and seems to warrant further familial partitioning of the Colubridae.

The double origins of both the Socotran and the Malagasy colubrids were well supported in our analysis. Alternatives implying single origins of either Socotran or Malagasy taxa were significantly excluded by SH tests and
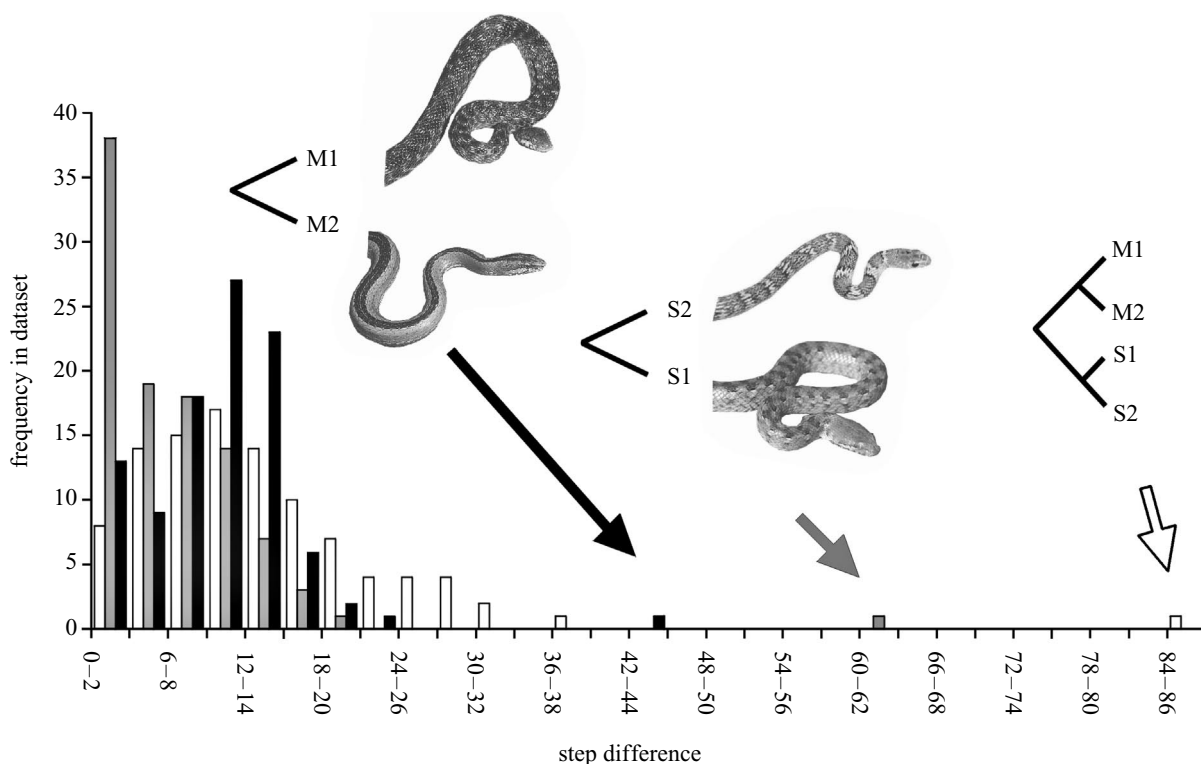
Figure 2. Results of the parametric bootstrap test of non-monophyly of Socotran and Malagasy colubrids. The bars show the distribution of step differences between constrained and unconstrained analyses of 100 simulated datasets, under the constraints of monophyly of the two Malagasy clades (black bars; M1 and M2, symbolized by pictures of *Madagascarophis* and *Mimophis*), monophyly of the two Socotran clades (grey bars; S1 and S2, *Ditypophis* and *Hemerophis*) and both (white bars). The arrows point to the step differences from the real datasets; because these are far beyond the range observed in the simulated data, the null hypothesis (monophyly) can be rejected at $p < 0.01$.
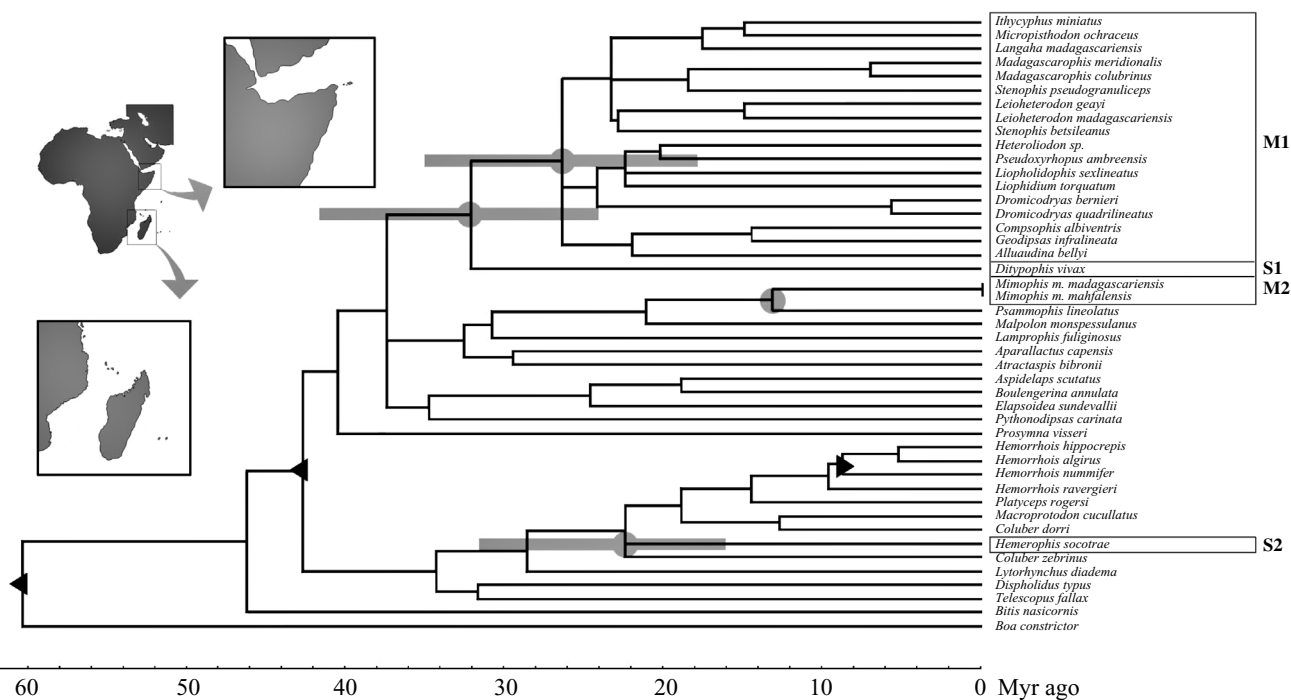


Figure 3. Chronogram of the evolution of the colubroid taxa studied. Non-parametric rate-smoothing was performed with the r8s program (Sanderson 2002) on the ML topology, with branch lengths calculated on the basis of non-saturated data partitions only. The tree was rooted with a mammal (*Sus*) and the root calibrated at 310 Myr ago. Further calibrations were entered as minimum and maximum constraints and are symbolized by black triangles (see § 2). Grey bars indicate 95% confidence intervals calculated by cross-validation (not possible for the *Mimophis–Psammophis* split). The inset picture shows the location of Madagascar and Socotra.

parametric bootstrapping. The placement of *Mimophis* among the psammophiines agrees with recent classification based on hemipenial morphology (Zaher 2000), morphology (de Haan 2003) and molecular data (Vidal & Hedges 2002). In addition, the relationships of *Hemerophis* were reliably resolved: *H. socotrae* had previously been placed in the paraphyletic genus *Coluber* (s.l.), and our data confirm affinities to the Old World representatives of this group (see also Schätti & Utiger 2001). The monophyly of the Malagasy pseudoxyrhophiines and the placement of the Socotran *Ditypophis* as their sister group within Pseudoxyrhophiinae received relevant bootstrap values in the *c-mos* and the combined trees, and we consider these affinities as probable (see also Zaher 2000).

Biogeographically, our data provide strong support for the suggestion that two independent events of transoceanic dispersal were involved in the colonization of Madagascar by colubrids. The young age of these clades (older 95% confidence interval less than 76 Myr ago in all estimates) provides convincing evidence against ancient origins by vicariance. The accepted time estimate for the Madagascar–Africa separation (165–121 Myr ago; Rabinowitz *et al.* 1983) predates our estimate by more than 40 Myr. The preferred phylogenetic topology (figure 1) suggests that in both cases the direction of the dispersal was towards Madagascar, but alternative hypotheses with one dispersal to Madagascar and a second dispersal out of Madagascar could not be rejected by means of the SH tests.

For Socotran colubrids, the situation is more ambiguous. Their Cenozoic age is also supported by the molecular-clock analyses. However, because the time of the last connection of Socotra with the mainland is disputed, it is not possible to decide whether *Hemerophis* and *Ditypophis* originated by vicariance or by transoceanic dispersal.

The placement of *Ditypophis* in the Pseudoxyrhophiinae points to the possibility of biogeographical connections between Madagascar and Socotra. The Cenozoic age of the taxa involved would imply direct transoceanic dispersal between these land masses, which seems less likely because of the large distance between them. If our tree is correct, a hypothetical colonization of Socotra from Madagascar should have occurred prior to any intra-Malagasy radiation and within a relatively narrow time window. The reverse scenario (colonization of Madagascar from Socotra) seems unlikely because the prevailing ocean currents would favour dispersal from Madagascar northwards (Arnold 2000).

Considering the general patterns of reptile dispersal in the Indian Ocean region (e.g. Caccone *et al.* 1999; Mausfeld *et al.* 2000; Raxworthy *et al.* 2002; Palkovacs *et al.* 2002), we favour another—admittedly speculative—scenario in which both Madagascar and Socotra were colonized independently from the African plate. According to our phylogenies, pseudoxyrhophiines diverged early in the initial colubrid radiation. The preferred molecular age estimates indicate that their radiation took place 9–18 Myr earlier than the divergences of *Mimophis* and *Hemerophis* from their sister groups. Representatives of the pseudoxyrhophiine clade might have been more widespread in the dawn of colubrid evolution. Socotra was reached by the ancestor of *Ditypophis*, while the specimens dispersing overseas to Madagascar gave rise to a remarkable endemic radiation. On the African and Asian mainland, these snakes might have become extinct when more successful colubrid radiations appeared. Madagascar and Socotra acted as biogeographical refuges for these ancient lineages, and were subsequently also reached by more modern colubrids: *Mimophis* and *Hemerophis*, respectively.

A similar hypothesis may also apply to Malagasy cichlid fishes (Vences *et al.* 2001), and possibly to gekkonid lizards (U. Joger, personal observation). The Seychellean colubrids may also consist of an older (*Lycognathophis*) and a more recent colonist (*Lamprophis geometricus*), which would constitute a remarkable parallel to the Socotran and Malagasy serpent fauna. Islands and archipelagos of the Indian Ocean may therefore parallel ancient lakes (e.g. Nishida 1991; Sherbakov *et al.* 1998) in being evolutionary reservoirs for old Cenozoic radiations that disappeared elsewhere.

## REFERENCES

Arnold, E. N. 2000 Using fossils and phylogenies to understand evolution of reptile communities on islands. In *Isolated vertebrate communities in the tropics* (ed. G. Rheinwald), pp. 309–324. *Bonner Zool. Monogr.* **46**.

Burbrink, F. T., Lawson, R. & Slowinski, J. B. 2000 Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. *Evolution* **54**, 2107–2118.

Caccone, A., Amato, G., Gratry, O. C., Behler, J. & Powell, J. R. 1999 A molecular phylogeny of four endangered Madagascar tortoises based on mtDNA sequences. *Mol. Phylogenet. Evol.* **12**, 1–9.

Cadle, J. E. 1994 The colubrid radiation in Africa (Serpentes: Colubridae): phylogenetic relationships and evolutionary patterns based on immunological data. *Zool. J. Linn. Soc.* **110**, 103–140.

Carroll, R. L. 1988 *Vertebrate paleontology and evolution*. New York: Freeman.

Crother, B. I. & Guyer, C. 1996 Caribbean historical biogeography: was the dispersal–vicariance debate eliminated by an extraterrestrial bolide? *Herpetologica* **52**, 440–465.

de Haan, C. C. 2003 Extrabuccal infralabial secretion outlets in *Dromophis*, *Mimophis* and *Psammophis* species (Serpentes, Colubridae, Psammophiini). A probable substitute for 'self-rubbing' and cloacal scent gland functions, and a cue for a taxonomic account. *C. R. Biologies* **326**, 275–286.

de Queiroz, A., Lawson, R. & Lemos-Espinal, J. A. 2002 Phylogenetic relationships of North American garter snakes: how much DNA is enough? *Mol. Phylogenet. Evol.* **22**, 315–329.

Dowling, H. G., Highton, R., Maha, G. C. & Maxson, L. R. 1983 Biochemical evaluation of colubrid snake phylogeny. *J. Zool.* **201**, 309–329.

Dowling, H. G., Hass, C. A., Hedges, S. B. & Highton, R. 1996 Snake relationships revealed by slow-evolving proteins: a preliminary survey. *J. Zool.* **240**, 1–28.

Farias, I. P., Orti, G., Sampaio, I., Schneider, H. & Meyer, A. 2001 The cytochrome *b* gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. *J. Mol. Evol.* **53**, 89–103.

Glaw, F. & Vences, M. 1994 *A fieldguide to the amphibians and reptiles of Madagascar*, 2nd edn. Cologne, Germany: Vences & Glaw Verlag.

Harris, D. J. 2003 Codon bias variation in C-*mos* between squamate families might distort phylogenetic inference. *Mol. Phylogenet. Evol.* **27**, 540–544.

Hedges, S. B. 1996 Vicariance and dispersal in Caribbean biogeography. *Herpetologica* **52**, 466–473.

Heise, P. J., Maxson, L. R., Dowling, H. G. & Hedges, S. B. 1995 Higher-level snake phylogeny inferred from mitochondrial DNA sequences of 12S rRNA and 16S rRNA genes. *Mol. Biol. Evol.* **12**, 259–265.

Huelsenbeck, J. P. & Ronquist, F. 2001 MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Huelsenbeck, J. P., Hillis, D. M. & Jones, R. 1996 Parametric bootstrapping in molecular phylogenetics: applications and performance. In *Molecular zoology: advances, strategies and protocols* (ed. J. D. Ferraris & S. R. Palumbi), pp. 19–45. New York: John Wiley & Sons.

Ivanov, M., Rage, J.-C., Szyndlar, Z. & Venczel, M. 2000 Histoire et origine géographique des faunes de serpents en Europe. *Bull. Soc. Herpétol. Fr.* **96**, 15–24.

Jansa, S. A., Goodman, S. M. & Tucker, P. K. 1999 Molecular phylogeny and biogeography of the native rodents of Madagascar (Muridae: Nesomyinae): a test of the single-origin hypothesis. *Cladistics* **15**, 253–270.

Joger, U. 1987 An interpretation of reptile zoogeography in Arabia, with special reference to Arabian herpetofaunal relations with Africa. In *Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East* (ed. R. Kinzelbach, F. Krupp & W. Schneider), pp. 257–271. Tübinger Atlas des Vorderen Orients. Wiesbaden: Reichert Verlag.

Joger, U. 2000 The reptile fauna of the Socotra Archipelago. In *Isolated vertebrate communities in the tropics* (ed. G. Rheinwald), pp. 337–356. *Bonner Zool. Monogr.* **46**.

Kelly, C. M. R., Barker, N. P. & Villet, M. H. 2003 Phylogenetics of advanced snakes (Caenophidia) based on four mitochondrial genes. *Syst. Biol.* **52**, 439–459.

Kopp, H. 1999 Abiotische Geofaktoren. In *Sokotra—Mensch und Natur* (ed. W. Wranik), pp. 3–22. Jemen-Studien monograph series, vol. 14. Wiesbaden.

Kraus, F. & Brown, W. M. 1998 Phylogenetic relationships of colubroid snakes based on mitochondrial DNA sequences. *Zool. J. Linn. Soc.* **122**, 455–487.

Krause, D. W., Hartman, J. H. & Wells, N. A. 1997 Late Cretaceous vertebrates from Madagascar. Implications for biotic changes in deep time. In *Natural change and human impact in Madagascar* (ed. S. M. Goodman & B. D. Patterson), pp. 3–43. Washington and London: Smithsonian Institution Press.

Krupp, F., Apel, M. & Zajonz, U. 2002 Meeresbiologische Untersuchungen im Sokotra-Archipel. *Natur Mus.* **132**, 273–294.

Kumar, S. & Hedges, S. B. 1998 A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920.

Laughton, A. S., Whitmarsh, A. B. & Jones, M. T. 1970 The evolution of the Gulf of Aden. *Phil. Trans. R. Soc. Lond.* A **267**, 227–266.

Mausfeld, P., Vences, M., Schmitz, A. & Veith, M. 2000 First data on the molecular phylogeography of scincid lizards of the genus *Mabuya*. *Mol. Phylogenet. Evol.* **17**, 11–14.

Murphy, W. J. & Collier, G. E. 1997 A molecular phylogeny for aplocheiloid fishes (Atherinomorpha, Cyprinodontiformes): the role of vicariance and the origins of annualism. *Mol. Biol. Evol.* **14**, 790–799.

Nishida, M. 1991 Lake Tanganyika as an evolutionary reservoir of old lineages of East African cichlid fishes: inferences from allozyme data. *Experentia* **47**, 974–979.

Nussbaum, R. A. 1984 Snakes of the Seychelles. In *Biogeography and ecology of the Seychelles Islands* (ed. D. R. Stoddart), pp. 361–377. The Hague: W. Junk Publisher.

Palkovacs, E. P., Gerlach, J. & Caccone, A. 2002 The evolutionary origin of Indian Ocean tortoises (*Dipsochelys*). *Mol. Phylogenet. Evol.* **24**, 216–227.

Palumbi, S. R., Martin, A., Romano, S., McMillan, W. O., Stice, L. & Grabowski, G. 1991 *The simple fool's guide to PCR, Version 2.0.* University of Hawaii.

Posada, D. & Crandall, K. A. 1998 Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Rabinowitz, P. D., Coffin, M. F. & Falvey, D. 1983 The separation of Madagascar and Africa. *Science* **220**, 67–69.

Rage, J. C. 1987 Fossil history. In *Snakes—ecology and evolutionary biology* (ed. R. A. Seigel, J. T. Collins & S. S. Novak), pp. 51–76. New York: Macmillan.

Rage, J. C. 1988 The oldest known colubrid snakes. *Acta Zool. Cracov.* **31**, 457–474.

Rage, J. C. 2001 Fossil snakes from the Paleocene of São José de Itaboraí, Brazil. II. Boidae. *Palaeovertebrata* **30**, 111–150.

Rambaut, A. & Grassly, N. C. 1997 Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* **13**, 235–238.

Raxworthy, C. J., Forstner, M. R. J. & Nussbaum, R. A. 2002 Chameleon radiation by oceanic dispersal. *Nature* **415**, 784–786.

Richards, C. M. & Moore, W. S. 1996 A phylogeny for the African treefrog family Hyperoliidae based on mitochondrial rDNA. *Mol. Phylogenet. Evol.* **5**, 522–532.

Ruedi, M., Auberson, M. & Savolainen, V. 1998 Biogeography of Sulawesian shrews: testing for their origin with a parametric bootstrap on molecular data. *Mol. Phylogenet. Evol.* **9**, 567–571.

Sanderson, M. J. 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1231.

Sanderson, M. J. 2002 r8s, version 1.05 (beta). University of California, Davis.

Schätti, B. & Utiger, U. 2001 *Hemerophis*, a new genus for *Zamenis socotrae* Günther, and a contribution to the phylogeny of Old World racers, whip snakes, and related genera (Reptilia: Squamata: Colubrinae). *Revue suisse Zool.* **108**, 919–948.

Schmidt, H. A., Strimmer, K., Vingron, M. & von Haeseler, A. 2002 Tree-Puzzle: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502–504.

Sherbakov, D. Y., Kamaltynov, R. M., Ogarkov, O. B. & Verheyen, E. 1998 Patterns of evolutionary change in Baikalian gammarids inferred from DNA sequences (Crustacea, Amphipoda). *Mol. Phylogenet. Evol.* **10**, 160–167.

Shimodaira, H. & Hasegawa, M. 1999 Multiple comparison of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116.

Slowinski, J. B. & Lawson, R. 2002 Snake phylogeny: evidence from nuclear and mitochondrial genes. *Mol. Phylogenet. Evol.* **24**, 194–202.

Storey, M., Mahoney, J. J., Saunders, A. D., Duncan, R. A., Kelley, S. P. & Coffin, M. F. 1995 Timing of hot spot-related volcanism and the breakup of Madagascar and India. *Science* **167**, 852–855.

Swofford, D. L. 2002 *PAUP*. Phylogenetic analysis using parsimony (*and other methods*), v. 4b10. Sunderland, MA: Sinauer.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997 The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882.

Vences, M., Freyhof, J., Sonnenberg, R., Kosuch, J. & Veith, M. 2001 Reconciling fossils and molecules: cenozoic divergence of cichlid fishes and the biogeography of Madagascar. *J. Biogeogr.* **28**, 1091–1099.

Vidal, N. & Hedges, S. B. 2002 Higher-level relationships of caenophidian snakes inferred from four nuclear and mitochondrial genes. *C. R. Biologies* **325**, 987–995.

Vidal, N., Kindl, S. G., Wong, A. & Hedges, S. B. 2000 Phylogenetic relationships of xenodontine snakes inferred from 12S and 16S ribosomal RNA sequences. *Mol. Phylogenet. Evol.* **14**, 389–402.

Whelan, S., Liò, P. & Goldman, N. 2001 Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends Genet.* **5**, 262–272.

Zaher, H. 2000 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.

Visit http://www.pubs.royalsoc.ac.uk to see electronic appendices to this paper.