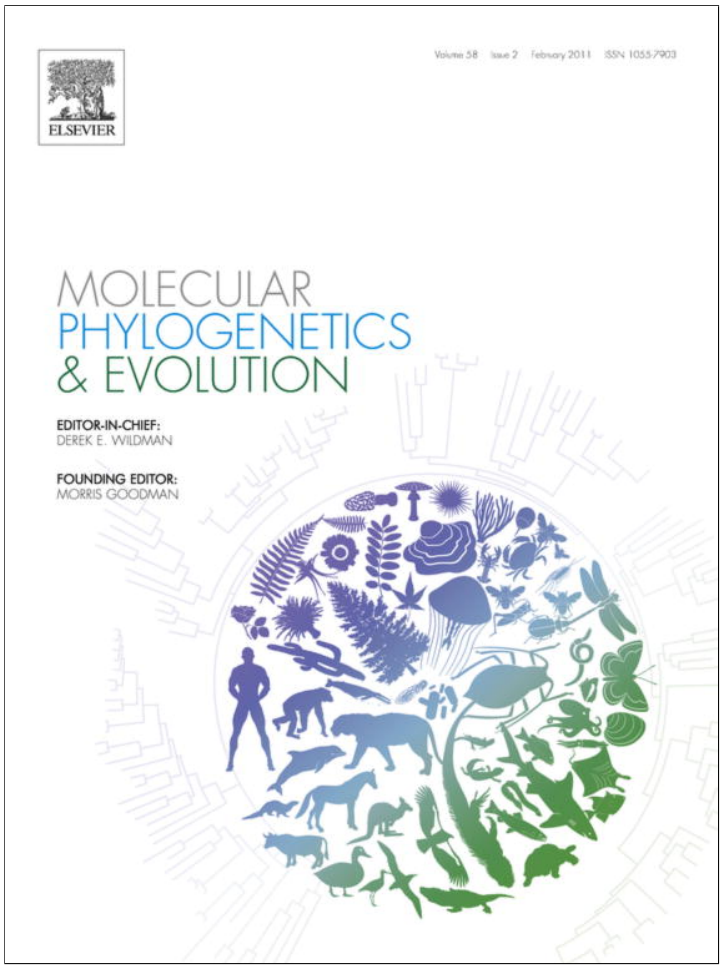

Molecular Phylogenetics and Evolution 58 (2011) 329–342

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

The phylogeny of advanced snakes (Colubroidea), with discovery of a new subfamily and comparison of support methods for likelihood trees

R. Alexander Pyron [a,b,*], Frank T. Burbrink [b,f], Guarino R. Colli [c], Adrian Nieto Montes de Oca [d], Laurie J. Vitt [e], Caitlin A. Kuczynski [a], John J. Wiens [a]

[a] Dept. of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794-5245, United States
[b] Dept. of Biology, The Graduate School and University Center, The City University of New York, 365 Fifth Ave., NY, NY 10016, United States
[c] Departamento de Zoologia, Universidade de Brasília, Brasília, DF Brasil 70910-900, Brazil
[d] Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México, México, Distrito Federal 04510, Mexico
[e] Sam Noble Oklahoma Museum of Natural History and Zoology Dept., 2401 Chautauqua Avenue, University of Oklahoma, Norman, OK 73072, United States
[f] Dept. of Biology, The College of Staten Island, The City University of New York, 2800 Victory Blvd., Staten Island, NY 10314, USA

ARTICLE INFO

Article history:
Received 14 August 2010
Revised 28 October 2010
Accepted 7 November 2010
Available online 11 November 2010

Keywords:
Bootstrapping
Colubroidea
Combined analysis
Phylogenetic methods
Snakes
Supermatrix

ABSTRACT

The superfamily Colubroidea (>2500 species) includes the majority of snake species and is one of the most conspicuous and well-known radiations of terrestrial vertebrates. However, many aspects of the phylogeny of the group remain contentious, and dozens of genera have yet to be included in molecular phylogenetic analyses. We present a new, large-scale, likelihood-based phylogeny for the colubroids, including 761 species sampled for up to five genes: cytochrome *b* (93% of 761 species sampled), ND4 (69%), ND2 (28%), c-mos (54%), and RAG-1 (13%), totaling up to 5814 bp per species. We also compare likelihood bootstrapping and a recently proposed ultra-fast measure of branch support (Shimodaira-Hasegawa-like [SHL] approximate likelihood ratio), and find that the SHL test shows strong support for several clades that were weakly-supported by bootstrapping in this or previous analyses (e.g., Dipsadinae, Lamprophiidae). We find that SHL values are positively related to branch lengths, but show stronger support for shorter branches than bootstrapping. Despite extensive missing data for many taxa (mean = 67% per species), neither bootstrap nor SHL support values for terminal species are related to their incompleteness, and that most highly incomplete taxa are placed in the expected families from previous taxonomy, typically with very strong support. The phylogeny indicates that the Neotropical colubrine genus *Scaphiodontophis* represents an unexpectedly ancient lineage within Colubridae. We present a revised higher-level classification of Colubroidea, which includes a new subfamily for *Scaphiodontophis* (Scaphiodontophiinae). Our study provides the most comprehensive phylogeny of Colubroidea to date, and suggests that SHL values may provide a useful complement to bootstrapping for estimating support on likelihood-based trees.

1. Introduction

Dense taxon sampling is extremely important for phylogenetic and evolutionary studies. For example, extensive taxon sampling may greatly increase phylogenetic accuracy under some conditions (e.g., Rannala et al., 1998; Zwickl and Hillis, 2002) and allow more accurate estimates of diversification rates (e.g., Heath et al., 2008; Cusimano and Renner, 2010). However, major challenges to inferring large-scale phylogenies remain. One is the expense and time necessary to obtain tissue samples and comparable character sampling for hundreds of species and many genes. Another is the difficulty of estimating trees and support values using sophisticated model-based methods (e.g., maximum likelihood) on large-scale data matrices in a reasonable amount of time.

Some recent advances have made inferring large-scale phylogenies more tractable for many groups. One is the finding that large matrices with extensive missing data can yield well-supported trees that are largely congruent with traditional taxonomy (e.g., Driskell et al., 2004; Philippe et al., 2004; Wiens et al., 2005; Thomson and Shaffer, 2010). The supermatrix approach to phylogenetics involves gathering all or most available data and analyzing it simultaneously (de Queiroz and Gatesy, 2007). This easily permits new sequence data to be combined with existing information from databases such as GenBank (e.g., Sanderson et al., 2003) to yield densely sampled supermatrices. Second, recent computational innovations have greatly facilitated estimating large-scale,

* Corresponding author. Present address: Department of Biological Sciences, The George Washington University, 2023 G St. NW, Washington, DC 20052, United States.

E-mail addresses: rpyron@colubroid.org (R.A. Pyron), frank.burbrink@csi.cuny.edu (F.T. Burbrink), grcolli@unb.br (G.R. Colli), anietomontesdeoca@me.com (A.N.M. de Oca), vitt@ou.edu (L.J. Vitt), kate@life.bio.sunysb.edu (C.A. Kuczynski), wiensj@life.bio.sunysb.edu (J.J. Wiens).

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.11.006

likelihood-based phylogenies (e.g., Stamatakis et al., 2008; Guindon et al., 2010; Price et al., 2010). Third, fast new methods for assessing clade confidence for likelihood trees have been developed that provide an alternative to traditional, time-intensive methods such as non-parametric bootstrapping (e.g., Anisimova and Gascuel, 2006; Guindon et al., 2010).

For many species-rich groups, supermatrix strategies (e.g., Wiens et al., 2005; de Queiroz and Gatesy, 2007; Thomson and Shaffer, 2010) offer a useful approach for inferring large-scale phylogenies when different amounts of character data are available across taxa. Other potential methods for large-scale phylogenetic inference include supertree construction (e.g., Bininda-Emonds, 2004) and mega-phylogeny approaches (e.g., Smith et al., 2009). The supertree method involves grafting trees inferred from different datasets into a single phylogeny, but it suffers from the need for *a priori* assumptions about which species belong to which taxa, limiting the potential for new discoveries about the phylogeny. The mega-phylogeny approach is similar to the supermatrix strategy, but uses an automated pipeline to identify gene regions and homologous sequence clusters of interest. We prefer the supermatrix strategy, as it most directly incorporates the largest amount of sequence data into the phylogenetic analysis, without assuming placement of species within groups *a priori*.

Here, we produce a new, large-scale phylogeny for the superfamily Colubroidea, the advanced snakes (*sensu* Lawson et al., 2005). Colubroids are among the most diverse groups of extant terrestrial vertebrates (>2500 species; Lawson et al., 2005) despite their relatively recent origin in the Cenozoic (Burbrink and Pyron, 2008; Vidal et al., 2009). They occur on every continent except Antarctica (Vitt and Caldwell, 2009) and include many common and familiar groups (e.g., racers, garter, rat, king, and milk snakes), and all known dangerously venomous snake species, such as elapids (cobras, sea snakes, and mambas) and viperids (e.g., rattlesnakes, adders, and vipers). These venomous colubroids are responsible for ~20,000–94,000 human fatalities every year (Kasturiratne et al., 2008). Given their diversity and broad distribution, colubroids have been the focus of many phylogeny-based studies in historical biogeography (e.g., Keogh, 1998; Pinou et al., 2004; Alfaro et al., 2008; Pyron and Burbrink, 2009a; Daza et al., 2010) and evolutionary biology (e.g., Fry and Wüster, 2004; Lynch, 2009; Pyron and Burbrink, 2009b,c; Burbrink and Pyron, 2010). However, despite the great biological and medical significance of this group, no study has offered a comprehensive assessment of the higher-level phylogeny of Colubroidea. For example, none has included representatives of all currently recognized subfamilies in a single analysis.

Several recent authors have addressed relationships within Colubroidea using DNA sequence data (e.g., Lawson et al., 2005; Burbrink and Pyron, 2008; Wiens et al., 2008; Kelly et al., 2009; Vidal et al., 2009; Zaher et al., 2009), typically sampling either many genes for relatively few taxa (e.g., seven genes for 24 species in Vidal et al., 2007; 20 genes for 29 species in Wiens et al., 2008) or many taxa for fewer genes (e.g., three genes for 131 species in Zaher et al., 2009). Major changes to colubroid taxonomy have been proposed based on these studies (e.g., Lawson et al., 2005; Burbrink et al., 2007; Vidal et al., 2007; Zaher et al., 2009). Yet, relatively few species and genera were included in these phylogenies, leaving the classification of many genera in question. These gaps in taxon sampling may hide radical differences between traditional taxonomies and molecular phylogenies. For example, the genus *Oxyrhabdium* was traditionally thought to belong to Xenodermatidae (Vitt and Caldwell, 2009), but molecular phylogenetic analyses showed it to be nested within Lamprophiidae (see Lawson et al., 2005; Kelly et al., 2009; Zaher et al., 2009).

Many of these recent molecular phylogenies agree regarding some relationships, such as monophyly of Homalopsidae and Viperidae. However, substantive disagreements remain regarding many parts of the phylogeny. One is the monophyly of the predominantly African assemblage "Lamprophiidae" (Kelly et al., 2009). Some studies have supported the monophyly of this group (Vidal et al., 2007; Burbrink and Pyron, 2008; Wiens et al., 2008; Zaher et al., 2009), whereas others have found it to be paraphyletic with respect to Elapidae (Kelly et al., 2009). Another is the placement of the family Homalopsidae as the sister taxon either to Elapidae + Lamprophiidae (Burbrink and Pyron, 2008), or to Elapidae + Lamprophiidae + Colubridae (Lawson et al., 2005; Vidal et al., 2007; Wiens et al., 2008). Yet another is the placement of the colubrid subfamily Dipsadinae as the sister taxon either to Natricinae (Vidal et al., 2008; Kelly et al., 2009; Zaher et al., 2009), Colubrinae (Vidal et al., 2007), or Colubrinae + Natricinae (Lawson et al., 2005; Wiens et al., 2008), with other colubrid subfamilies, such as Calamariinae and Pseudoxenodontinae (if sampled), often found interdigitated among these clades (Lawson et al., 2005; Zaher et al., 2009). These issues are important for numerous reasons, including understanding the relationships among medically significant taxa, and the interpretation of historical biogeographic scenarios. All of these questions are best addressed through a large-scale phylogenetic analysis of Colubroidea, using as many taxa as possible to resolve relationships within the group.

Here, we address colubroid relationships and classification using a supermatrix approach that combines data for two nuclear genes, three mitochondrial genes, and 761 colubroid species in 299 genera, totaling 70% of the 426 known genera and 29% of the 2654 identified species (The Reptile Database: Uetz, 2009; http://www.reptile-database.org/). In previous studies, dozens of researchers have generated sequences for hundreds of colubroid species, with five genes being commonly used (mitochondrial cytochrome *b*, ND2, and ND4; and nuclear c-mos and RAG-1). We also present new sequence data for 41 additional species (38 genera) from the two most species-rich colubroid subfamilies (Dipsadinae and Colubrinae), most of which have never been included in a molecular phylogenetic analysis. We combine these new sequences with existing data from previous studies to produce the largest analysis of Colubroidea to date, containing nearly six times as many species as any previous estimate, and including all known families and subfamilies in the same analysis for the first time.

We also compare two methods for estimating clade support for large-scale, likelihood-based phylogenies. Specifically, we compare support values from traditional, non-parametric bootstrapping (BS hereafter; Felsenstein, 1985, 2004) and the non-parametric Shimodaira-Hasegawa-like approximation of the likelihood-ratio test statistic (SHL hereafter; Anisimova and Gascuel, 2006; Guindon et al., 2010). Using SHL support values may be desirable, especially for large trees, as calculating them can be several orders of magnitude faster than assessing traditional BS support (Anisimova and Gascuel, 2006) and both measures seem to give similar values (Guindon et al., 2010). Although the thorough study by Guindon et al. (2010) addressed many aspects of the relative performance of these methods using empirical and simulated datasets, some key questions remain. Here, we assess the relationship between these values and branch lengths and missing data (in terminal taxa), questions that were not addressed in previous studies. We hypothesize that SHL values will be positively related to branch lengths (as shown for likelihood BS values; Wiens et al., 2008), but will show higher values than BS values on shorter branches (suggested but not explicitly tested by Guindon et al., 2010). We also hypothesize that there will be little or no relationship between SHL support values for clades and the proportion of missing data in the terminal taxa in those clades, nor for likelihood BS values (as shown for parsimony BS values and Bayesian posterior probabilities; Wiens et al., 2005). A key assumption of the supermatrix approach is that extensive missing data in terminal taxa need not prevent them from being placed in the tree with strong support.

2. Materials and methods

2.1. Baseline taxonomy

The initial taxonomy used in this paper follows Lawson et al. (2005). However, we differ from that paper in following the traditional, restricted usage of Elapidae (i.e., the clade of dangerously venomous taxa, including subfamilies Elapinae, Hydrophiinae, and Laticaudinae), and in referring to the putative sister group of Elapidae as Lamprophiidae (following Vidal et al., 2007). The lamprophiids occur primarily in Africa, Madagascar, and the Middle East, and include the subfamilies Aparallactinae, Atractaspidinae, Lamprophiinae, Prosymninae, Psammophiinae, Pseudaspidinae, and Pseudoxyrhophiinae (following Vidal et al., 2008; Kelly et al., 2009). We also follow Vidal et al. (2007) and recognize the subfamily Grayiinae (family Colubridae), considered part of Colubrinae by Lawson et al. (2005). Finally, we follow Zaher et al. (2009) in using the name Dipsadinae (Bonaparte, 1840) to refer to Xenodontinae (Bonaparte, 1845) of Lawson et al. (2005) and previous authors, based on priority.

2.2. Sequence acquisition

We obtained sequences for 761 colubroid species (299 genera) and six outgroup species from the nuclear protein-coding genes oocyte maturation factor Mos (c-mos; 572 base pairs [bp]; 414 taxa) and recombination-activating gene 1 (RAG-1; up to 2604 bp; 106 taxa), and the mitochondrial protein-coding genes cytochrome *b* (cyt-b; 1000 bp; 716 taxa), NADH subunit 4 (ND4; 684 bp; 530 taxa), and NADH subunit 2 (ND2; 954 bp; 218 taxa). Sequences for 720 of the colubroid species were obtained from GenBank. We searched GenBank regularly up to September, 2009, using subfamilies as the search terms, and supplementing these searches with literature surveys to ensure that all available sequences were included from existing datasets. Outgroup taxa were from the families Acrochordidae, Aniliidae, Boidae, Cylindrophiidae, Tropidophiidae, and Uropeltidae and were selected to represent the other major groups of alethinophidian snakes (e.g., Wiens et al., 2008). Within these families, species were selected for which sequences from all five genes were available.

We sequenced c-mos and cyt-b for 41 additional species (38 genera), using standard methods of DNA extraction, amplification, and sequencing (see Lawson et al., 2005). These species were selected because most of the genera that they represent were not present in GenBank when we initiated our study, at least for the standard protein-coding loci analyzed here. Amplification and sequencing used the following primers: for cyt-b, H14910 and THRSN2 (Burbrink et al., 2000); for c-mos, G77 and G78 (Lawson et al., 2005). Voucher and GenBank accession numbers are given in Appendix A (supplementary material).

Sequences were aligned using the MUSCLE algorithm (Edgar, 2004) with the default parameters in the program Geneious v5.0.3 (Genematters Corp.). Alignments were checked by eye, and were unambiguous for all taxa, as all sequences were protein coding and maintained an open reading frame. Indels consisted of a single 3 bp indel occurring in c-mos in approximately half of the taxa. The final matrix comprises 767 terminal taxa and 5814 bp. We excluded taxa represented only by c-mos because this gene fragment shows limited sequence variation, and including these taxa led to poorly resolved relationships in preliminary analyses. The alignment and trees are available at TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S10982).

Not all taxa had sequences for all five genes. On average, each species had 1935 bp present (range 274–5794 bp) and 67% missing data (range 0.003–95%). Much of the missing data are associated with the lack of RAG-1 (~60% of the characters) in most taxa. Although missing data may be a cause for concern (Lemmon et al., 2009), simulations and empirical studies suggest that even highly incomplete taxa (e.g., 90–95% missing data) can be accurately placed in parsimony, likelihood, and Bayesian analyses, if the overall number of informative characters in the analysis is large (e.g., Wiens, 2003; Driskell et al., 2004; Philippe et al., 2004; Wiens et al., 2005; Wiens and Moen, 2008). Indeed, our results here (see below) show that even the most highly incomplete taxa are consistently placed in the expected higher-level clades (e.g., incomplete taxa traditionally classified as viperids are nested inside Viperidae), typically with very strong support.

2.3. Phylogenetic analyses

We performed phylogenetic inference using ML and assessed support using three separate strategies. First, we performed ML tree inference and non-parametric bootstrapping using the program RAxMLv7.0.4 (Stamatakis, 2006) with the five-gene concatenated matrix. We used the GTRGAMMA model for all genes and partitions because GTR is the only substitution model implemented in RAxML. Previous phylogenetic analyses of snakes suggest that GTR + Γ + I is the best-fitting model for these genes, and that these genes should be partitioned by codon positions (e.g., Lawson et al., 2005; Bryson et al., 2007; Wiens et al., 2008). Further, the GTRGAMMA model in RAxML is recommended over the GTR + Γ + I because the 25 rate categories account for potentially invariant sites (Stamatakis, 2006). We used the rapid-bootstrapping algorithm (1000 non-parametric bootstrap replicates) with the thorough ML search option (200 independent searches, starting from every fifth bootstrap replicate). We performed two additional ML tree inferences without bootstrapping to ensure that the analysis was not stuck on local optima, based on qualitative similarity in topology and likelihood values. Given that BS values generally appear to be biased but conservative (Felsenstein, 2004), we considered clades with values of 70% or greater to be well-supported (see Taylor and Piel, 2004).

Second, to provide an alternative to bootstrapping, we assessed branch support using the SHL test (Anisimova and Gascuel, 2006; Guindon et al., 2010), implemented in RAxMLv7.2.6. For a given branch, this approach compares the estimated ML branch to the next two most likely nearest-neighbor interchange (NNI) rearrangements of that branch. Support is measured as $1 - P$, where P is equal to the probability of the null hypothesis (i.e., that the reconstructed branch is not significantly more likely than alternative rearrangements); thus, SHL values have a relatively straightforward statistical interpretation (Guindon et al., 2010). The support values provided by BS and the SHL method have been shown to be highly congruent for both simulated and empirical data when strong "phylogenetic signal" is present in a dataset, where phylogenetic signal is measured based on the number of sites with non-missing data and the internal branch lengths (Guindon et al., 2010). Selection thresholds of 0.8–0.9 have been shown by simulation to yield sufficient power ($1 - \beta$ [type II error rate] = ~0.85) for both simulated and empirical data (Guindon et al., 2010); therefore, we use 0.85 as a cutoff for "strong" support. We estimated SHL support values for 200 independent replicates in RAxMLv7.2.6, using the fast ML search algorithm implemented using the '-f E' option. We used RAxMLv7.0.4 for the rapid-bootstrapping ML analysis, as it is the most recent, stable release version that is implemented in parallel for supercomputing clusters (Stamatakis et al., 2008), while the SHL algorithm is available only in alpha versions of RAxMLv7.2.x.

Third, we compared SHLvalues implemented in RAxMLv7.2.6 to those in both PHYML v3.0 (Guindon et al., 2010) and FastTree v2.0 (Price et al., 2010) to ensure that SHL values were not an artifact of

the implementation of this approach in a particular package. Results were quantitatively similar (in terms of significance) for all analyses described below, and we prefer RAxML for estimating SHL values because the other two packages do not support model partitioning for multi-gene alignments, so only RAxML results are reported. Correlation with the RAxML SHL values was high for both FastTreev2.0 ($r_s = 0.81$, $P < 0.001$) and PHYMLv3.0 ($r_s = 0.83$, $P < 0.001$). Note that this and all other correlations involving BS and SHL support values were assessed using Spearman's rank correlation in R 2.11.0 (R Core Development Team, 2010).

The analyses estimating SHL values required generating a new ML phylogeny, which was mostly congruent with the tree estimated using the rapid-bootstrapping, intensive ML search algorithm. We plotted the SHL values for congruent nodes on the tree from the more intensive ML BS search. We assessed concordance between the BS and SHL support values by testing the correlation between them across congruent nodes (using Spearman's rank correlation), which would be positive if they are highly concordant. We determined which measure was larger (on average) by calculating the mean of the two sets of values. Finally, we calculated the SHL/BS ratio, equal to the number of branches with both strong SHL and BS support divided by the number of branches strongly supported by either method (Guindon et al., 2010). The SHL/BS ratio is equal to one when both methods strongly support all the same branches throughout the tree, and 0 when they strongly support completely different sets of branches. However, this may be misleading if congruence is low between the two trees (i.e., if only two branches are well-supported and congruent, then the ratio is high but uninformative).

We place high confidence in clades that are strongly supported by both high BS and SHL values. Clades with strong support from one measure and weak support from the other are more difficult to interpret. In general, we expect that shorter branches will have lower support from both measures, due to the decreased probability that informative mutations will accumulate along those branches and the increased frequency of conflicts between gene trees on shorter branches (see Wiens et al., 2008). This is likely to be particularly true for BS, which relies on the frequency of informative sites and congruence among characters (Guindon et al., 2010), but may be less important for SHL values, which are based on the likelihoods of alternative topologies. We tested for a relationship between the branch lengths of the primary ML tree and the BS and SHL values for branches congruent between both searches.

In addition, Guindon et al. (2010) suggested that SHL values are likely to be more accurate than BS values for shorter branches, simply because a short branch may be correct but supported by relatively few sites (which may yield a high SHL value but low BS). They suggested that the opposite may also occur, such that SHL values are correctly low, but BS proportions erroneously give strong support, possibly due to the "star-tree paradox" (the tendency for some branches to be strongly resolved even in the absence of signal, or other sources of hidden determinism in the algorithm; Guindon et al., 2010). Finally, it is also possible that SHL values will be inappropriately high in some cases due to the limited consideration of alternate topologies, though this may be less problematic if the data contain clear signal and a good approximation of the ML topology has been found (Guindon et al., 2010). Thus, we tested for a relationship between the difference (SHL-BS) between the support values for each congruent branch, and the length of that branch. We expect a negative relationship if SHL > BS for shorter branches.

We also tested for a relationship between support for the placement of terminal taxa (using both BS and SHL), and the proportion of sequence data present in the matrix for those species. A previous analysis suggested that parsimony BS proportions and Bayesian posterior probabilities for the placement of terminal taxa are uncorrelated with the proportion of missing data in these taxa, and that support values are instead related to the data that are present in these taxa, not the data that are absent (Wiens et al., 2005). However, it is not clear if this is also true for SHL values and likelihood BS values.

To address this topic, we first modified the protocol of Wiens et al. (2005) for testing the relationship between support for the placement of terminal taxa and their completeness. For sister species, we used the support for that pair and the lowest completeness of the two, rather than using both and duplicating the support value as in Wiens et al. (2005). When a single species was the sister taxon to a clade containing more than one species, we used the completeness of that taxon and the product of the support values for (a) the node subtending that taxon and its sister clade and (b) the node subtending the sister clade (rather than using the mean of these two values, as used by Wiens et al. (2005)).

This modification reduces some of the pseudoreplication inherent in the protocol of Wiens et al. (2005), but may suffer from considering only the most highly incomplete taxa (e.g., a pair of terminal taxa in which both have 95% missing data receives the same score as a pair in which one has 95% missing data and the other is 100% complete) and still leaves some pseudoreplication in the use of support values in non-sister species. We also ran these analyses using the original protocol of Wiens et al. (2005). A significant result using either would indicate that missing data negatively affect branch support. However, a non-significant result does not mean that all incomplete taxa are placed correctly, nor does a significant result mean that incomplete taxa are misplaced.

Like all supermatrix studies to date, ours is based on a concatenated (combined) analysis of different genes. Recent studies have suggested that explicit species-tree analyses may provide more accurate phylogenetic inferences in cases where there is strong discordance between gene trees and the underlying species tree (e.g., Edwards et al., 2007; Degnan and Rosenberg, 2009). However, these methods currently are not practical for an analysis of 761 taxa, with data from only three independent loci.

3. Results

A summary of the ML tree based on rapid-bootstrapping analysis from RAxMLv7.0.4 ($-\ln L = -370658.05$) is shown in Fig. 1, and the full tree is shown in Fig. 2. Between the rapid-BS ML (RAxMLv7.0.4) and NNI-optimized (RAxMLv7.2.6) searches, 89% (679 of 766) branches were congruent. We show the BS and SHL values for nodes at the subfamily level and above (Fig. 1). For clarity of presentation, we show only BS proportions for all nodes (Fig. 2), but the NNI-optimized topology with SHL support values is provided as Supplementary material (online Appendix B). The ML analysis yields strong support (BS and SHL) for Xenodermatidae as the sister group to all other colubroids and Pareatidae as the sister group to all Colubroidea excluding xenodermatids (Figs. 1 and 2). Viperids are strongly placed as the sister group to other colubroids exclusive of xenodermatids and pareatids. Within Viperidae, Viperinae is strongly supported as sister group to Azemiopinae and Crotalinae. In contrast to other recent studies (e.g., Vidal et al., 2007; Wiens et al., 2008), our results suggest a sister-group relationship between Homalopsidae and Elapidae + Lamprophiidae (Figs. 1 and 2B; Burbrink and Pyron, 2008). This is the only major disagreement found between this analysis and the phylogenomic results of Wiens et al. (2008), and is not strongly supported in our study.

We find strong support for Elapidae + Lamprophiidae, but poor BS support for the monophyly of Lamprophiidae. However, SHL values are high for the monophyly of Lamprophiidae (98%)

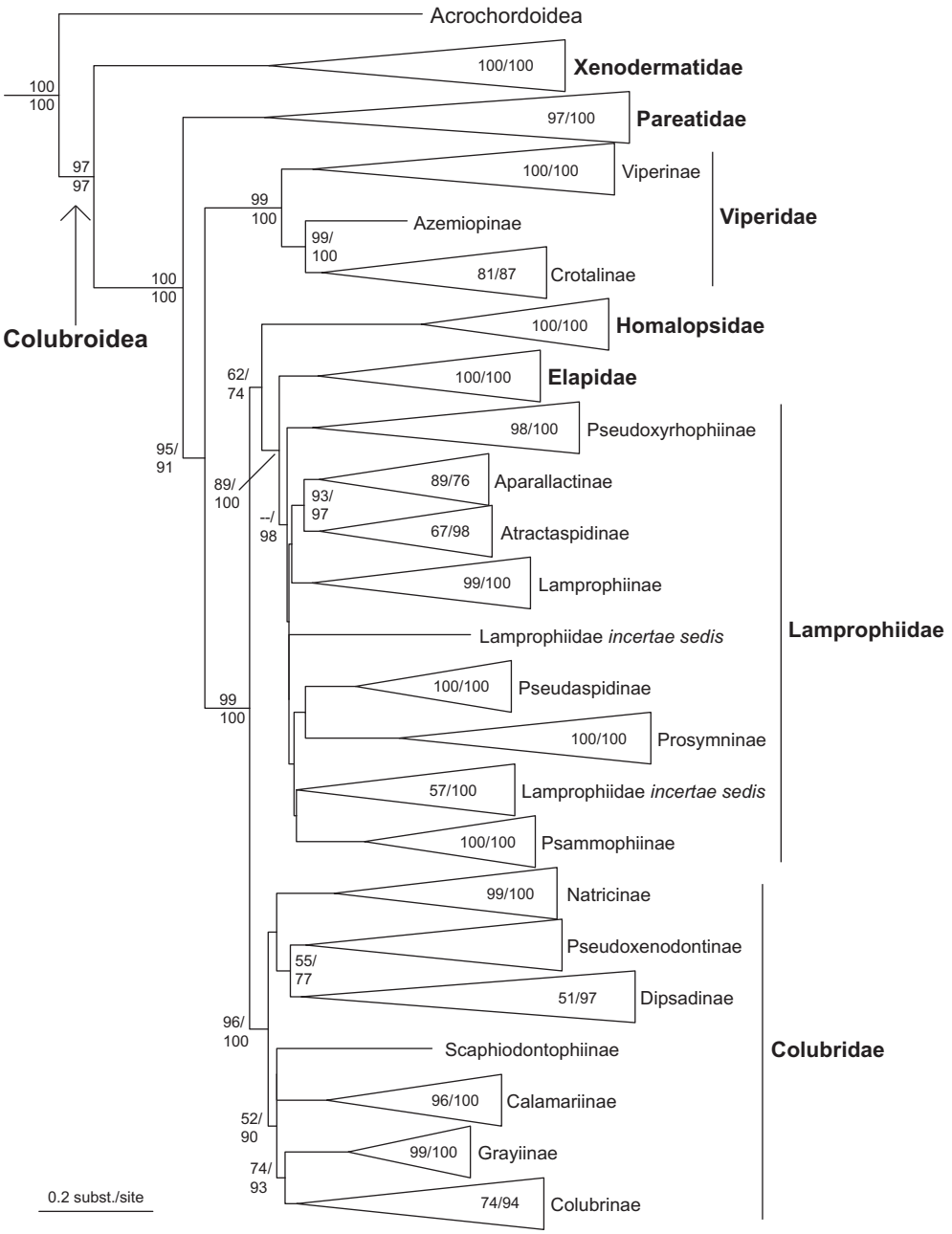

Fig. 1. Summary phylogeny of 761 colubroid snake species based on a concatenated maximum likelihood analysis of five genes (5814 bp). Only families and subfamilies are shown (full topology in Fig. 2). Triangles indicate clades with >1 species sampled, but the size of triangles is arbitrary and constant between clades. Numbers next to branches or in terminal triangles indicate BS/SHL support values. Only values >50% are shown. Five additional outgroup species are not shown.

compared to the next two most likely arrangements of that node. We find strong support for the monophyly of five of the seven subfamilies of Lamprophiidae (Lamprophiinae, Psammophiinae, Pseudaspidinae, Pseudoxyrhophiinae, and Prosymninae), whereas Aparallactinae and Atractaspidinae are each only supported by BS or SHL values, respectively. Relationships among these subfamilies are only weakly supported by both methods. Further, the placement of several taxa of uncertain subfamilial assignment within Lamprophiidae (i.e., *Buhoma*, *Oxyrhabdium*, and *Psammodynastes*) remains poorly resolved (Figs. 1, 2B), as in previous studies (e.g., Lawson et al., 2005; Kelly et al., 2009).

Monophyly of Colubridae is strongly supported (BS = 96%, SHL = 100%), but relationships among the five subfamilies are not. Monophyly is well-supported for each of the subfamilies Calamariinae, Colubrinae, Grayiinae, Natricinae, and Pseudoxenodontinae (excepting the genera *Scaphiodontophis* and *Thermophis*, see below). We infer a weakly-supported sister-group relationship between Natricinae and Dipsadinae + Pseudoxenodontinae. Monophyly of Dipsadinae is weakly-supported (51%) by BS proportion, but strongly supported (97%) by SHL values (Fig. 1). The enigmatic genus *Thermophis* from Tibet is weakly placed as the sister taxon to Pseudoxenodontinae, contrary to other analyses placing it within Dipsadinae (e.g., Huang et al., 2009), and we tentatively place it in Pseudoxenodontinae. This clade (Pseudoxenodontinae) is weakly-supported as the sister taxon to Dipsadinae (e.g., Lawson et al., 2005; Vidal et al., 2007).

Numerous colubrine and dipsadine genera (e.g., *Adelphicos*, *Chrysopelea*, *Conopsis*, *Dendrophidion*, *Drymobius*, *Drymoluber*,

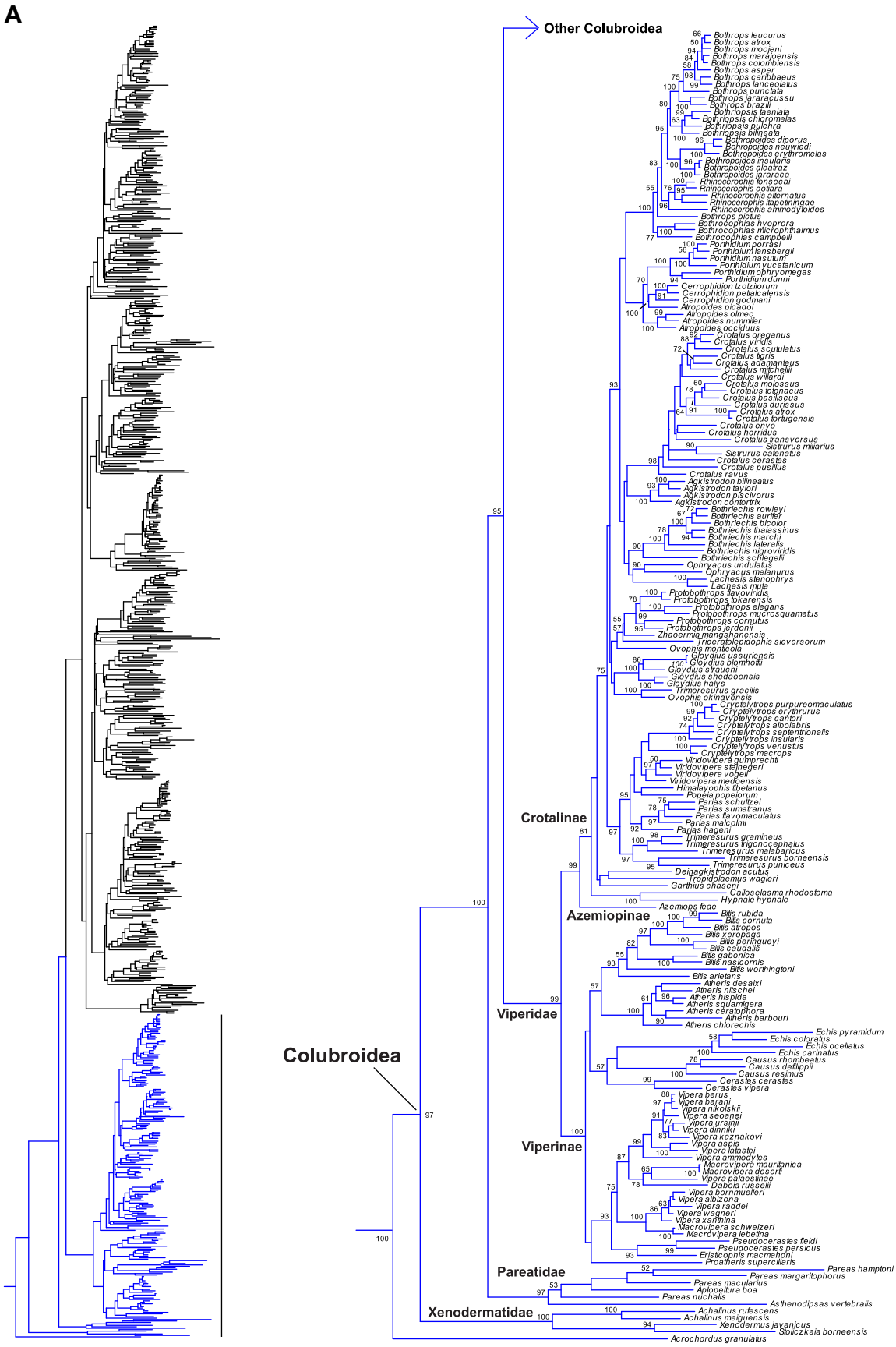

Fig. 2. Phylogeny of 761 colubroid snake species (see summary in Fig. 1). The tree was estimated based on a likelihood analysis of five concatenated genes. Five outgroup species are not shown. BS values greater than 50% are shown. Colors of clades indicate their position in the overall tree, shown at left.

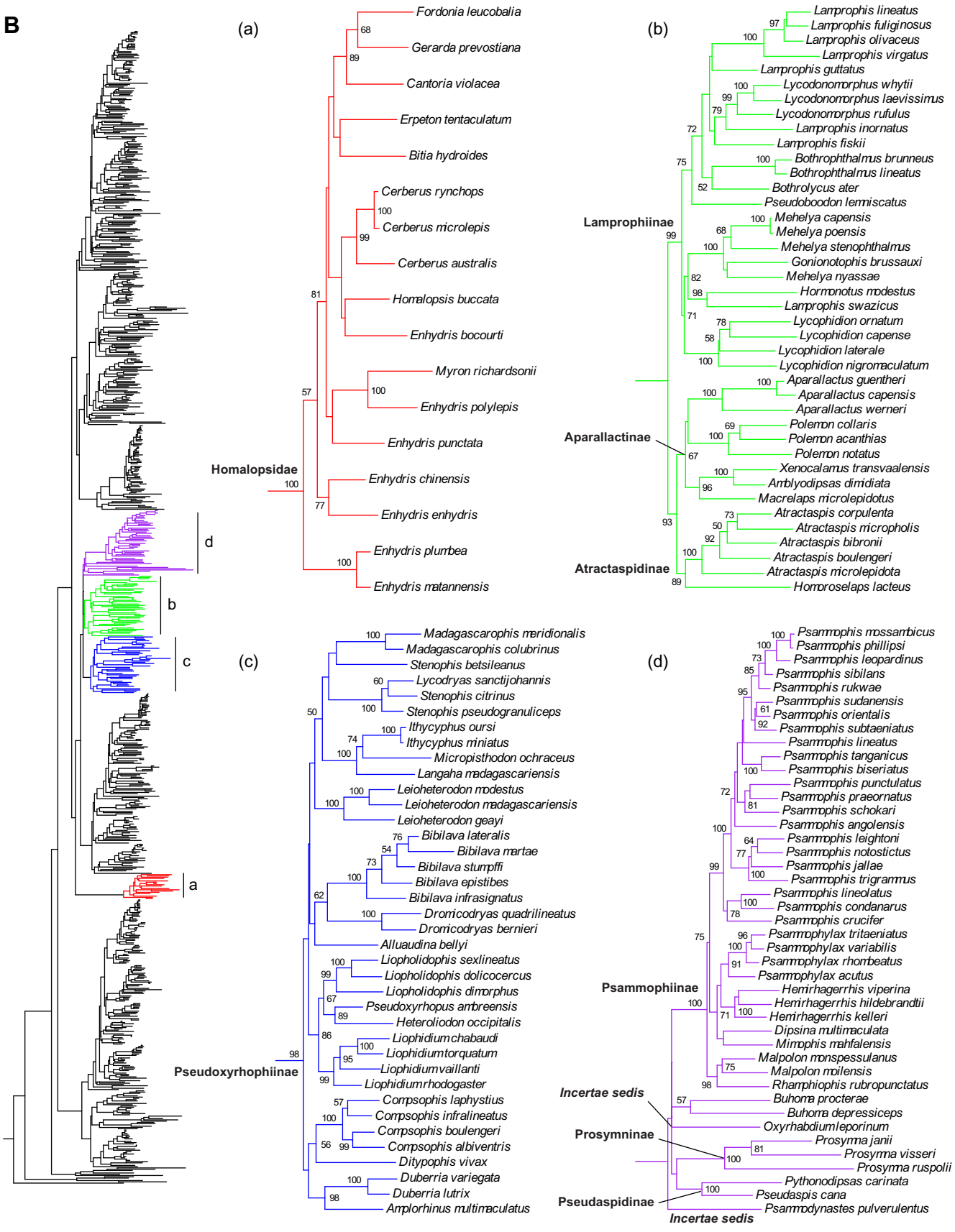

Fig. 2 (*continued*)

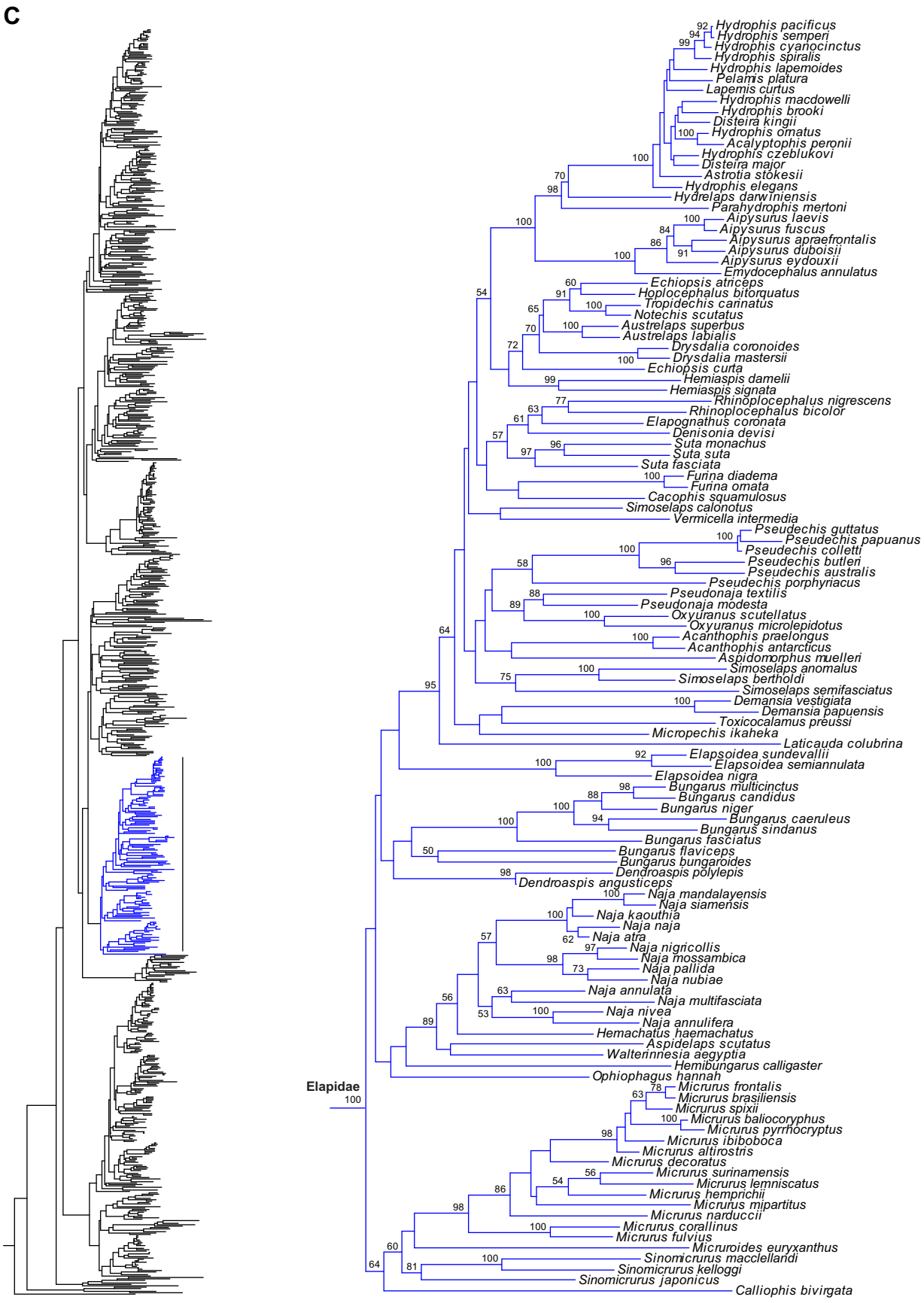

Fig. 2 (*continued*)

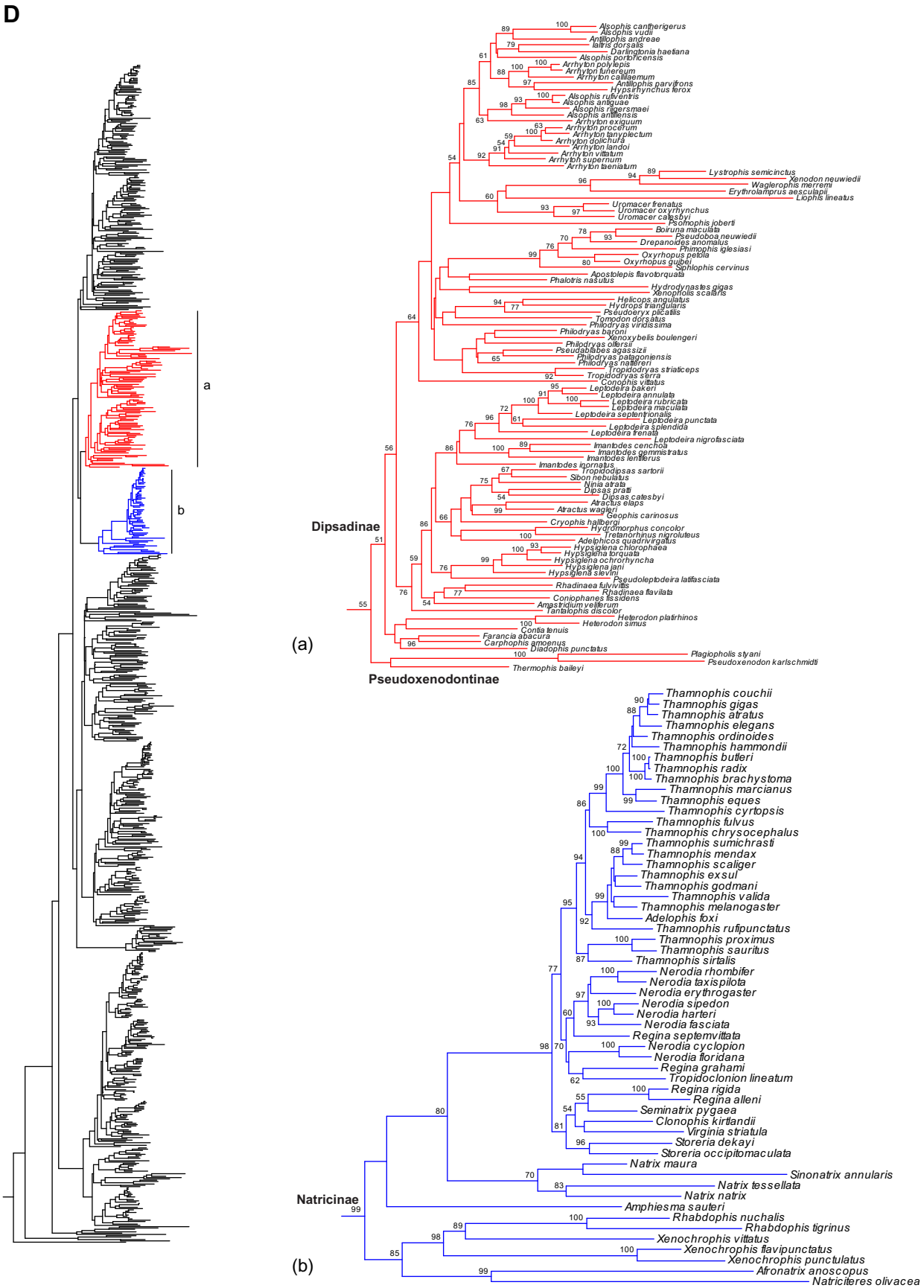

Fig. 2 (*continued*)

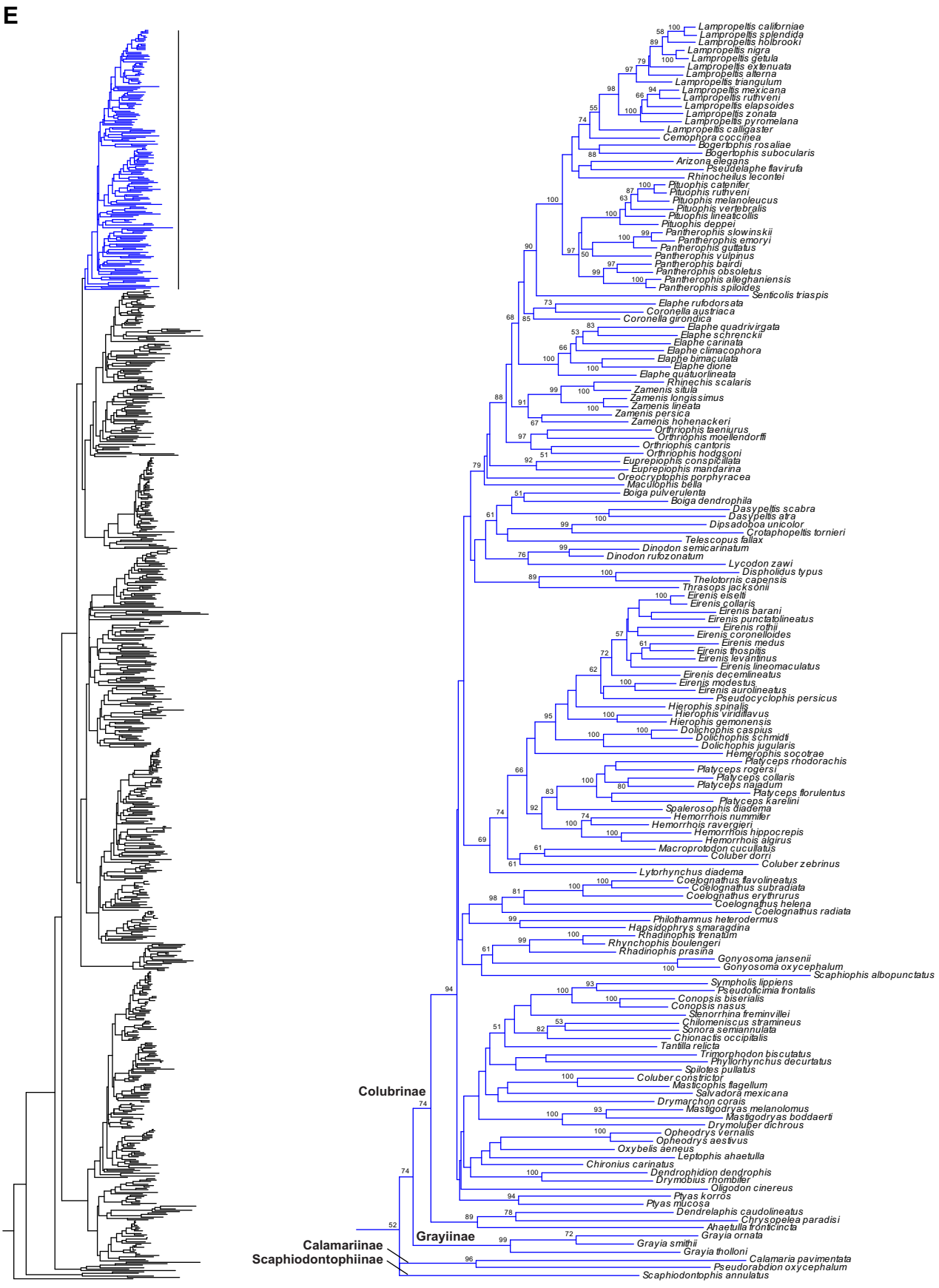

Fig. 2 (*continued*)

Geophis, *Hydromorphus*, *Leptophis*, *Pseudoficimia*, *Scaphiodontophis*, *Stenorrhina*, and *Sympholis*; Appendix A) are included in a molecular analysis for the first time here. Almost all of these genera are placed with strong support in the subfamilies expected based on previous taxonomy (Fig. 2D–E). However, we find that the Neotropical genus *Scaphiodontophis* is not placed inside of Colubrinae, where it has traditionally been classified (e.g., Lawson et al., 2005). This taxon has not been included in any previous molecular phylogenies, and it apparently represents an unexpectedly ancient lineage within Colubridae (Fig. 2E), which we recognize as a distinct subfamily (see below). This species is included based on data from cyt-b, c-mos, and RAG-1, and is considerably more complete than average (55% vs. 33% mean across all species), strongly suggesting that the surprising placement of this taxon is not an artifact of missing data. Remarkably, the Neotropical genus *Scaphiodontophis* is weakly placed as the sister taxon to the Asian subfamily Calamariinae, rather than the other NW colubrid groups.

The SHL and BS support values exhibit a strong positive correlation ($r_s = 0.82$, $P < 0.001$) across the nodes of the phylogeny, indicating generally high concordance (Fig. 3A). The relationship is triangular (constraining); branches well-supported by BS proportion are typically well-supported by SHL, though branches with lower BS support may have strong SHL support. The mean SHL values are higher on average (86) than the BS proportions (70). The SH-like/BS ratio is 0.37, indicating that 37% of branches receiving strong support from either method receive strong support from both. As noted above, most of the higher-level nodes are well-supported using both measures (Fig. 1).

For both BS and SHL values, we observe a significant positive relationship between branch support and branch length (SHL in Fig. 3B, BS not shown; $r_s = 0.72$ for both, $P < 0.001$), with both measures of support dropping off sharply for branches with less than 0.05 substitutions per site. On average, we find that SHL values are higher than BS proportions for these shorter branches, though they become increasingly congruent as branch lengths increase ($r_s = -0.44$, $P < 0.001$; Fig. 3C). However, there is a small subset of these very short branches (<0.05 subst./site) for which BS > SHL (43 of 472, 9%).

We do not observe a significant relationship between the support for the placement of terminal taxa and their completeness (SHL in Fig. 3D; BS not shown; $r_s = 0.058$ and -0.008, $P = 0.19$ and 0.85, respectively). Results are similar using the unmodified protocol of Wiens et al. (2005) for both SHL ($r_s = 0.035$, $P = 0.33$) and BS ($r_s = 0.037$, $P = 0.31$) values. These results suggest that highly incomplete taxa (mean = 67% missing data across all taxa) can be placed in the phylogeny with strong BS and SHL support. Furthermore, we find that all 761 species (regardless of completeness) are placed in the family-level clades predicted by previous, primarily morphological, taxonomies, suggesting that their overall phylogenetic placement is accurate despite extensive missing data in most species. In addition, all families but Lamprophiidae have very strong BS and SHL support for their monophyly.

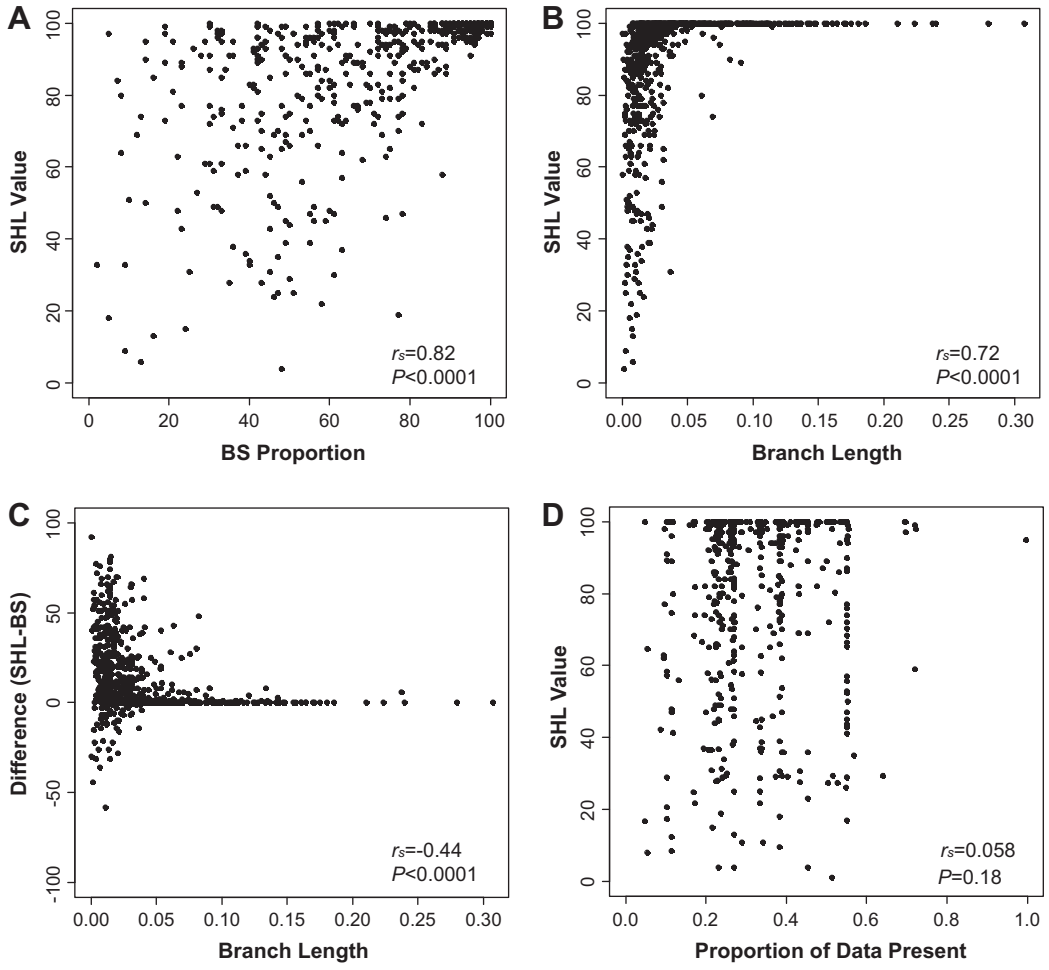

Fig. 3. Behavior of SHL values with regard to a number of variables: (A) relationship between BS proportions and SHL support values for congruent nodes, (B) relationship of SHL values to branch lengths, (C) relationship of difference between support values (SHL-BS) to branch lengths, and (D) Spearman's correlation between the completeness of a terminal taxon (1 – proportion missing data) and SHL support for its placement.

4. Discussion

4.1. Large-scale phylogenetics and branch support

In this paper, we reconstruct a large-scale phylogeny of colubroid snakes utilizing a supermatrix approach to taxon and character sampling and new, fast methods for estimating likelihood trees and evaluating their support. We also compare two methods for assessing support on large-scale likelihood trees (BS and SHL). Admittedly, our empirical analyses cannot address the crucial relationship between support values and accuracy, where accuracy is the probability of clades being correctly reconstructed (see Felsenstein, 2004). Nevertheless, we address two critical questions about the performance of these methods that have not been addressed before. Previous studies, based on simulated and empirical datasets (Anisimova and Gascuel, 2006; Guindon et al., 2010), have found that BS and SHL values are correlated, suggesting that SHL values are often likely to be accurate under the same set of conditions as BS proportions. We also find such a correlation between SHL and BS values here. However, previous studies have not addressed the relationship between SHL values and branch lengths, nor how SHL and likelihood BS values may be related to missing data.

We found that both BS and SHL values are positively correlated with branch lengths, as might be expected given that longer branches should have more changes and greater congruence among genes. Thus, branch lengths seem to explain much of the variation in support values across nodes, and similar responses to branch length variation may explain much of the congruence between BS and SHL values. We also found that SHL values tend to be higher on shorter branches, which may explain many of the observed cases where SHL values differed from BS support on a given branch. However, this trend for higher SHL on short branches was not universal, and these results alone do not necessarily mean that short branches with high SHL and low BS values are likely to be accurately reconstructed. We also found a small set of short branches that receive high BS and low SHL support (see also Guindon et al., 2010). Whether this is due to recognized issues such as the star-tree paradox (e.g., Susko, 2008), or hidden algorithmic flaws is unclear, and the support for these branches should perhaps be considered weak (Guindon et al., 2010). Exploration of the relationship between accuracy and SHL and BS support values under these short-branch conditions should be a priority for future simulation studies.

We also found that both SHL and likelihood BS values are unrelated to the amount of missing data in terminal taxa. This result confirms previous findings for parsimony BS and Bayesian analysis (e.g., Wiens et al., 2005), and suggests that SHL values are not unusually sensitive to missing data. In addition, we found that even the most highly incomplete taxa (>90% missing data) were placed in the higher taxa (family or subfamily) expected based on previous taxonomy, and that most of these traditionally recognized groups had very strong support for their monophyly, based on both BS and SHL values. Overall, these results support previous simulation and empirical results suggesting that highly incomplete taxa can be accurately placed in supermatrix analyses, provided that the overall number of characters in the matrix is high (e.g., Wiens, 2003; Driskell et al., 2004; Philippe et al., 2004; Wiens et al., 2005; Wiens and Moen, 2008).

While further simulation studies are necessary to evaluate the performance and utility of the SHL method, our results suggest that it may provide a useful complement to BS for assessing branch support, particularly on very large trees. Importantly, it is much faster to compute than BS proportions or Bayesian posterior probabilities (Anisimova and Gascuel, 2006; Guindon et al., 2010; Price et al., 2010). Indeed, the rapid-BS analysis (1000 replicates, 200 ML searches) took 2.6 days using 16 cores of a 240-core Dell PowerEdge supercomputing system at the High Performance Computing Center at the City University of New York. In comparison, the NNI-optimized SHL analysis (200 searches) took 3 days on a single 3.4Ghz desktop machine belonging to the senior author, and would presumably take only a small fraction of that time on a supercomputer.

4.2. Systematics of Colubroidea

The higher-level classification of Colubroidea has been in flux as new molecular results contradict traditional taxonomy, and new phylogenies and taxonomies contradict each other (e.g., Lawson et al., 2005; Vidal et al., 2007, 2010; Burbrink et al., 2007; Wiens et al., 2008; Zaher et al., 2009). By using the most comprehensive colubroid phylogeny so far and a conservative approach to taxonomic changes, we here attempt to stabilize higher-level colubroid classification, focusing on the subfamily level and above. Our new classification follows from our phylogenetic results, a few simple guidelines, and the taxonomic recommendations of Lawson et al. (2005), Vidal et al. (2007), Wiens et al. (2008), and Zaher et al. (2009). Most importantly, we strive to retain traditional taxonomy as much as possible, while still maintaining monophyletic higher taxa. Our phylogeny also suggests paraphyly of many genera (e.g., *Crotalus*, *Enhydris*, *Nerodia*, *Rhadinophis*, *Stenophis*, *Thamnophis*, *Vipera*, *Zamenis*, etc.), though we refrain from addressing generic-level taxonomy, pending more complete sampling.

First, we define crown-group Colubroidea to consist of the extant families Colubridae, Elapidae, Homalopsidae, Lamprophiidae, Pareatidae, Viperidae, and Xenodermatidae, comprising the most recent common ancestor of *Coluber constrictor* and *Xenodermus javanicus*, and all descendants of that ancestor. Colubroidea is the sister taxon to Acrochordoidea. The colubroid stem-group thus includes all species more closely related to *C. constrictor* than to *Acrochordus javanicus*. This usage of Colubroidea dates back to Romer (1956), and has been widely used by herpetologists (e.g., Dowling and Duellman, 1978; Greene, 1997; Zaher, 1999; Lawson et al., 2005; Wiens et al., 2008; Vitt and Caldwell, 2009). This long-standing definition differs from recent proposals by Vidal et al. (2007) and Zaher et al. (2009), who make Colubroidea equivalent to Colubridae (sensu Lawson et al., 2005). These authors elevated the subfamilies of Colubridae to the family level (i.e., Calamariidae, Colubridae, Dipsadidae [or Xenodontidae], Natricidae, and Pseudoxenodontidae), requiring that the traditional Colubridae be ranked as a superfamily (Colubroidea). However, there is no need for recognizing these subfamilies as families, or for changing the long-standing definition of Colubroidea, as these are merely changes in the rank of monophyletic groups. Thus, we retain the traditional meaning of Colubroidea, and recognize these clades as subfamilies in the family Colubridae.

Traditionally, the clades Homalopsidae, Pareatidae, and Xenodermatidae were also considered subfamilies within Colubridae (e.g., Greene, 1997; Vitt and Caldwell, 2009). However, recent phylogenetic analyses and classifications all concur that they are only distantly related to other Colubridae, and thus must be recognized as separate families (e.g., Lawson et al., 2005; Vidal et al., 2007; Wiens et al., 2008; Zaher et al., 2009; this study). Viperidae retains its traditional designation, consisting of subfamilies Viperinae, Crotalinae, and Azemiopinae. We follow the traditional usage of Elapidae, including only the medically significant venomous snakes variously placed in the subfamilies Elapinae, Hydrophiinae, and Laticaudinae (e.g., Greene, 1997; Vitt and Caldwell, 2009). However, these subfamilies do not form well-supported monophyletic groups in our phylogeny (Fig. 2C), and we do not recognize any subfamilies in Elapidae at present. This choice contrasts with Lawson et al. (2005), who favored an expanded Elapidae, but

agrees with Vidal et al. (2007), Wiens et al. (2008), and Zaher et al. (2009), among others.

We consider the most difficult aspect of higher-level colubroid taxonomy to be Lamprophiidae, the assemblage of mostly African snakes related to Elapidae (Vidal et al., 2008; Kelly et al., 2009). The monophyly of Elapidae + Lamprophiidae is strongly supported in all analyses, but the monophyly of Lamprophiidae is weakly-supported by BS values, though strongly supported by SHL values (Figs. 1 and 2). We follow Vidal et al. (2007) in tentatively recognizing Lamprophiidae as a single family, including Aparallactinae, Atractaspidinae, Lamprophiinae, Psammophiinae, and Pseudoxyrophiinae. Our results agree with Kelly et al. (2009) in recognizing the subfamily Pseudaspidinae for the genera *Pseudaspis* and *Pythonodipsas*, recognizing Prosymninae for the genus *Prosymna*, and in including these two subfamilies in Lamprophiidae. The genera *Buhoma*, *Oxyrhabdium*, and *Psammodynastes* cannot be placed confidently within the existing subfamilies of Lamprophiidae. Additionally, the genus *Montaspis* was not included in our study, but was also classified as Lamprophiidae *incertae sedis* by Kelly et al. (2009). Relationships among these groups remain poorly supported, and should be a priority for future studies.

Our results support monophyly of Colubridae, containing the traditionally recognized subfamilies Calamariinae, Colubrinae, Natricinae, Pseudoxenodontinae, and Dipsadinae (Figs. 1 and 2D–E). The placement of the Old World subfamily Pseudoxenodontinae remains poorly supported, though a sister-group relationship with Dipsadinae has been suggested by other authors (e.g., Vidal et al., 2007). The enigmatic Tibetan taxon *Thermophis* (putatively a dipsadine; Huang et al., 2009) is placed as the sister taxon to the pseudoxenodontine genera *Plagiopholis* and *Pseudoxenodon*, and we here consider *Thermophis* to be part of Pseudoxenodontinae. We also find support for recognizing the African subfamily Grayiinae (see Vidal et al., 2007), which is placed as the sister taxon to Colubrinae and strongly supported by BS and SHL support values (Fig. 2E). Finally, we find that the New World colubrine genus *Scaphiodontophis* represents an ancient lineage with a phylogenetic placement that renders Colubrinae paraphyletic (Figs. 1 and 2E). This genus has not been included in previous molecular phylogenies, but it has traditionally been placed in Colubrinae (e.g., Lawson et al., 2005; Zaher et al., 2009). To avoid paraphyly of Colubrinae, we recognize a new subfamily for *Scaphiodontophis*.

Scaphiodontophiinae *subfam. nov.*

Type: Genus *Scaphiodontophis* (Taylor and Smith, 1943), species *S. annulatus* (Duméril et al., 1854).
Etymology: No etymology is given by Taylor and Smith (1943). Apparently from the Greek for "spade-toothed snake;" from "skaphe (σκφη)," meaning "spade- " or "boat-like," "odonto- (οδοντω)," meaning "teeth," and "ophis (όφις)," for snake; in reference to the spatulate (i.e., spade- or boat-shaped) maxillary teeth found in all species, some of which are also hinged.
Content: The two species of *Scaphiodontophis* (Taylor and Smith, 1943), *S. annulatus* (Duméril et al., 1854) and *S. venustissimus* (Günther, 1885). Numerous other species have been recognized historically (Taylor and Smith, 1943).
Diagnosis and Definition: Monophyly of scaphiodontophiines is supported by a very unusual synapomorphy: in both species, at least some maxillary teeth are hinged (Taylor and Smith, 1943; Savage and Slowinski, 1996). Another synapomorphy is the presence of fracture planes between the caudal vertebrae, rendering the snakes capable of non-regenerative inter-vertebral pseudoautotomy (i.e., their tails are modified such that they break off easily, but between the vertebrae, rather than within them as in lizards; Savage and Slowinski, 1996). Both traits are very rare in snakes, and are almost certainly derived within Colubroidea (Savitzky, 1981; Savage and Slowinski, 1996).
Other morphological traits, not necessarily derived, include (all data from Taylor and Smith, 1943): small body size (<900 mm from snout to tail tip); dorsal scales smooth, lacking apical pits, in 17 rows; subcaudals single; anal plate divided; one loreal, one preocular, usually two postoculars; one or two anterior temporals; tail long, approaching the body length in some specimens; hemipenis short (10 or 11 subcaudals in length), sulcus single. The color pattern in life is highly variable, typically consisting of uniform (black or brown) dorsal color with faint dark stripes, interspersed with or dominated by tricolored bands (monads or dyads) of black, red, and white or yellow, especially anteriorly, potentially mimicking venomous coral snakes (*Micrurus*).
Distribution: Scaphiodontophiines occur from southern Mexico to northern South America (The Reptile Database: Uetz, 2009; http://www.reptile-database.org).

Acknowledgments

For use of tissues in their care, we thank R.M. Brown and W.E. Duellman (University of Kansas), B.D. Hollingsworth (San Diego Natural History Museum), T.W. Reeder (San Diego State University), J. Vindum and R. Lawson (California Academy of Sciences), T.J. LaDuc (University of Texas, Austin), and C. Austin, J. Boundy, and D. Dittman (Louisiana State University Museum of Natural Sciences). We thank F.M. Fontanella (Brigham Young University) for generating sequences of several taxa. We thank all of the researchers whose time and effort went in to producing the GenBank data used in this project. We are grateful to A. Larson, T. Guiher, D. Frost, and two anonymous reviewers for their comments on the manuscript. Analyses were facilitated by a grant of computer time from the City University of New York High Performance Computing Center, which is supported by US National Science Foundation grants CNS-0855217 and CNS-0958379. Financial support for lab work was provided by US National Science Foundation grant EF 0334923 to J.J.W. R.A.P. was supported during preparation of this manuscript by US National Science Foundation grant DBI-0905765. Additional support for R.A.P. and F.T.B. was provided by the Graduate School and University Center and the College of Staten Island, both of the City University of New York. Financial support for field studies to collect many of the specimens and tissues used in this study was provided by US National Science Foundation grants DEB-9200779, DEB-9505518, and DEB-0415430 to L.J.V. and J.P. Caldwell, and DEB-0613802 to J.A. Campbell and O. Flores-Villela.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.11.006.

References

Alfaro, M.E., Karns, D.R., Voris, H.K., Brock, C.D., Stuart, B.L., 2008. Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. Mol. Phylogenet. Evol. 46, 576–593.
Anisimova, M., Gascuel, O., 2006. Approximate likelihood ratio test for branches: a fast, accurate and powerful alternative. Syst. Biol. 55, 539–552.
Bininda-Emonds, O., 2004. The evolution of supertrees. Trends Ecol. Evol. 19, 315–322.
Bonaparte, C.L., 1840. Amphibia europaea ad systema nostrum vertebratorum ordinata. Memorie della Reale Accademia delle Scienze di Torino 2, 385–456.
Bonaparte, C.L., 1845. Specchio general dei Sistemi Erpetologico ed Amfibiologico. Atti Congr. Sci. Ital. 6, 376–378.

Bryson, R.W., Pastorini, J., Burbrink, F.T., Forstner, M.R.J., 2007. A phylogeny of the *Lampropeltis mexicana* complex (Serpentes: Colubridae) based on mitochondrial DNA sequences suggests evidence for species-level polyphyly within *Lampropeltis*. Mol. Phylogenet. Evol. 43, 674–684.

Burbrink, F.T., Lawson, R., Slowinski, J.B., 2000. Molecular phylogeography of the North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. Evolution 54, 2107–2114.

Burbrink, F.T., Crother, B.I., Lawson, R., 2007. The destabilization of North American snake taxonomy. Herp. Rev. 38, 273–278.

Burbrink, F.T., Pyron, R.A., 2008. The taming of the skew: estimating proper confidence intervals for divergence dates. Syst. Biol. 57, 317–328.

Burbrink, F.T., Pyron, R.A., 2010. How does ecological opportunity influence rates of speciation, extinction, and morphological diversification in New World ratsnakes (tribe Lampropeltini)? Evolution 64, 934–943.

Cusimano, N., Renner, S.S., 2010. Slowdowns in diversification rates from real phylogenies may not be real. Syst. Biol. 59, 458–464.

Daza, J.M., Castoe, T.A., Parkinson, C.L., 2010. Using regional-scale comparative phylogeographic data to infer historical processes in Middle America. Ecography 33, 343–354.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340.

de Queiroz, A., Gatesy, J., 2007. The supermatrix approach to systematics. Trends Ecol. Evol. 22, 34–41.

Dowling, H.G., Duellman, W.E., 1978. Systematic Herpetology: A Synopsis of Families and Higher Categories. HISS Publications, New York.

Driskell, A.C., Ané, C., Burleigh, J.G., McMahon, M.M., O'Meara, B.C., Sanderson, M.J., 2004. Prospects for building the Tree of Life from large sequence databases. Science 306, 1172–1174.

Duméril, A.M.C., Bibron, G., Duméril, A.H.A., 1854. Erpétologie générale ou Histoire Naturelle complète des Reptiles, vol. 7 (partie 1), Paris.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA 104, 5936–5941.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Sunderland, MA.

Fry, B.G., Wüster, W., 2004. Assembling an arsenal: origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. Mol. Biol. Evol. 21, 870–883.

Greene, H.W., 1997. Snakes, The Evolution of Mystery in Nature. University of California Press, Berkeley.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Günther, A.C.L.G., 1885. Reptilia and Batrachia, Biologia Centrali-Américana. Taylor and Francis, London.

Heath, T.A., Zwickl, D.J., Kim, J., Hillis, D.M., 2008. Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. Syst. Biol. 57, 160–166.

Huang, S., Liu, S.Y., Guo, P., Zhang, Y.P., Zhao, E.M., 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.

Kasturiratne, A., Wickremasinghe, A.R., de Silva, N., Gunawardena, N.K., Pathmeswaran, A., et al., 2008. Estimation of the global burden of snakebite. PloS Med. 5, e218.

Kelly, C.M.R., Barker, N.P., Villet, M.H., Broadley, D.G., 2009. Phylogeny, biogeography and classification of the snake superfamily Elapoidea: a rapid radiation in the late Eocene. Cladistics 25, 38–63.

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

Lawson, R., Slowinski, J.B., Crother, B.I., Burbrink, F.T., 2005. Phylogeny of the Colubroidea (Serpentes): new evidence from mitochondrial and nuclear genes. Mol. Phylogenet. Evol. 37, 581–601.

Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of missing data on phylogenetic estimates obtained by maximum-likelihood and Bayesian inference. Syst. Biol. 58, 130–145.

Lynch, V.J., 2009. Live-birth in vipers (Viperidae) is a key innovation and adaptation to global cooling during the Cenozoic. Evolution 63, 2457–2465.

Philippe, H., Snell, E.A., Bapteste, E., Lopez, P., Holland, P.W.H., Casane, D., 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol. Biol. Evol. 21, 1740–1752.

Pinou, T., Vicario, S., Marscher, M., Caccone, A., 2004. Relict snakes of North America and their relationships to the Caenophidia using Bayesian methods on mitochondrial sequences. Mol. Phylogenet. Evol. 32, 563–574.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE 5, e9490.

Pyron, R.A., Burbrink, F.T., 2009a. Can the Tropical Conservatism Hypothesis explain temperate species richness patterns? An inverse latitudinal biodiversity gradient in the New World snake tribe Lampropeltini. Glob. Ecol. Biogeogr. 18, 406–415.

Pyron, R.A., Burbrink, F.T., 2009b. Body size as a primary determinant of ecomorphological diversification and the evolution of mimicry in the lampropeltinine snakes. J. Evol. Biol. 22, 2057–2067.

Pyron, R.A., Burbrink, F.T., 2009c. Neogene diversification and taxonomic stability in the snake tribe Lampropeltini (Serpentes: Colubridae). Mol. Phylogenet. Evol. 52, 524–529.

Rannala, B., Huelsenbeck, J.P., Yang, Z., Nielsen, R., 1998. Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47, 702–710.

Romer, A.S., 1956. Osteology of the Reptiles. University of Chicago Press, Chicago.

Sanderson, M.J., Driskell, A.C., Ree, R.H., Eulenstein, O., Langley, S., 2003. Obtaining maximally concatenated phylogenetic data sets from large sequence databases. Mol. Biol. Evol. 20, 1036–1042.

Savage, J.M., Slowinski, J.B., 1996. Evolution of coloration, urotomy and coral snake mimicry in the snake genus *Scaphiodontophis* (Serpentes: Colubridae). Biol. J. Linn. Soc. 57, 129–194.

Savitzky, A.H., 1981. Hinged teeth in snakes: an adaptation for swallowing hard-bodied prey. Science 212, 346–349.

Smith, S.A., Beaulieu, J.M., Donoghue, M.J., 2009. Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. BMC Evol. Biol. 9, 37.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A fast bootstrapping algorithm for the RAxML web-servers. Syst. Biol. 57, 758–771.

Susko, E., 2008. On the distributions of bootstrap support and posterior distributions for a star tree. Syst. Biol. 57, 602–612.

Taylor, E.H., Smith, H.M., 1943. A review of American sibynophine snakes, with a proposal of a new genus. Univ. Kansas Sci. Bull. 29, 301–337.

Taylor, D.J., Piel, W.H., 2004. An assessment of accuracy, error, and conflict with support values from genome-scale phylogenetic data. Mol. Biol. Evol. 21, 1534–1537.

Thomson, R.C., Shaffer, H.B., 2010. Sparse supermatrices for phylogenetic inference: taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. Syst. Biol. 59, 42–58.

Uetz, P., 2009. The Reptile Database. <http://www.reptile-database.org/>.

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

Vidal, N., Branch, W.R., Pauwels, O.S.G., Hedges, S.B., Broadley, D.G., Wink, M., Cruaud, C., Joger, U., Nagy, Z.T., 2008. Dissecting the major African snake radiation: a molecular phylogeny of the Lamprophiidae Fitzinger (Serpentes, Caenophidia). Zootaxa 1945, 51–66.

Vidal, N., Rage, J.-C., Couloux, A., Hedges, S.B., 2009. Snakes (Serpentes). In: Hedges, S.B., Kumar, S. (Eds.), The Timetree of Life. Oxford University Press, New York, pp. 390–397.

Vidal, N., Dewynter, M., Gower, D.J., 2010. Dissecting the major American snake radiation: a molecular phylogeny of the Dipsadidae Bonaparte (Serpentes, Caenophidia). C.R. Biologies 333, 48–55.

Vitt, L.J., Caldwell, J.P., 2009. Herpetology: An Introductory Biology of Amphibians and Reptiles, third ed. Academic Press, San Diego.

Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.

Wiens, J.J., Moen, D.S., 2008. Missing data and the accuracy of Bayesian phylogenetics. J. Syst. Evol. 46, 307–314.

Wiens, J.J., Fetzner, J.W., Parkinson, C.L., Reeder, T.W., 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Syst. Biol. 54, 719–748.

Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D.G., Sites Jr., J.W., Townsend, T.M., Reeder, T.W., 2008. Branch lengths, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. Syst. Biol. 57, 420–431.

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.

Zaher, H., Grazziotin, F.G., Cadle, J.E., Murphy, R.W., Moura-Leite, J.C., Bonatto, S.L., 2009. Molecular phylogeny of advanced snakes (Serpentes, Caenophidia) with an emphasis on South America xenodontines: a revised classification and descriptions of new taxa. Pap. Av. Zool. 49, 115–153.

Zwickl, D.J., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598.