Forked Tongues Revisited: Molecular Apomorphies Support Morphological Hypotheses of Squamate Evolution

Caleb D. McMahan[1,2], Layla R. Freeborn[3], Ward C. Wheeler[4], and Brian I. Crother[5]

Incongruence between morphological and molecular-based phylogenetic hypotheses has been reported across a wide range of taxa. Specifically, morphological and molecular hypotheses of squamate phylogeny have been consistently incongruent and have been notoriously difficult to reconcile. With the ever-rising popular view of the superiority of molecular data over morphological data, studies have been biased toward seeking explanations for homoplasy in the morphological data. We propose considering approaches that do not make *a priori* assumptions about the superiority of one type of data over another. We present two case studies using the proto-oncogene *c-mos*, as well as a large multilocus dataset, to examine apomorphy distributions across hypotheses. This approach reveals novel insights into data incongruence currently plaguing squamate phylogeny, with direct implications for other scenarios of phylogenetic incongruence.

INCONGRUENCE between molecular and morphological-based phylogenetic hypotheses have been reported across a wide range of taxa, including woodcreepers (Irestedt et al., 2004), cetaceans and artiodactylans (O'Leary et al., 2003; O'Leary and Gatesy, 2008), placental mammals (Lee and Camens, 2009), haplosporidians (Burreson and Reece, 2006), pickerelweed (Graham et al., 1998), hexapods (Bitsch et al., 2004), fishes (Betancur-R., 2009), and squamates (Vidal and Hedges, 2005; Losos et al., 2012). Numerous aspects of analyses (e.g., taxon sampling, outgroup selection, number and types of characters/genes included, chosen method of analysis, species vs. gene trees) may help explain the discordance between phylogenies derived from these different data types; however, the issue still receives considerable attention in the literature, and deservedly so, as it is largely unresolved (Hermsen and Hendricks, 2008; Lee and Camens, 2009; Losos et al., 2012). Nevertheless, many evolutionary biologists are quick to prefer molecular hypotheses to morphological, which begs the conclusion that the molecular data and inferred phylogeny are somehow superior to the morphological (e.g., Crother and Presch, 1992; Hedges and Sibley, 1994; Hedges and Maxson, 1996; Scotland et al., 2003; Wiens, 2004). For example, it is common practice in the study of trait evolution to map characters over molecular trees (termed the molecular scaffold approach; Hermsen and Hendricks, 2008). However, this method makes the implicit and unwarranted assumption that the molecular hypothesis is the best-inferred hypothesis (Mooi and Gill, 2010).

Squamata (lizards, snakes, amphisbaenians) comprises a group of approximately 8,000 species with a diversity of ecological, behavioral, and physiological adaptations. Since Camp's initial, pre-cladistic work (Camp, 1923), numerous morphology-based analyses, often including numerous fossil taxa, have obtained results strikingly similar to the landmark analysis (Wu et al., 1996; Evans and Barbadillo, 1998; Lee, 1998; Caldwell, 1999; Lee and Caldwell, 2000;

Evans et al., 2005; Conrad, 2008). However, the results of molecular studies (Townsend et al., 2004; Vidal and Hedges, 2005; Albert et al., 2009; Eckstut et al., 2009; Wiens et al., 2010, 2012; Pyron et al., 2013) sharply contradict morphology-based hypotheses. Losos et al. (2012) published a thought-provoking overview of the stark incongruence between morphological and molecular phylogenetic hypotheses of squamates and brought to light the need for investigation into the factors leading to this incongruence. In short, all morphological studies of squamate phylogeny (regardless of taxon sampling and characters) have recovered a sister relationship between the Iguania and Scleroglossa; however, no molecular study has ever recovered this relationship. Furthermore, there is little to no consensus across molecular studies as to the earliest divergence within squamates. By way of assessing the incongruence, Conrad (2008) fit his morphological data onto two molecular phylogenies (Towsend et al., 2004; Vidal and Hedges, 2005) of squamates and found hypotheses less parsimonious by 175 and 171 steps, respectively. Gauthier et al. (2012) were also unsuccessful in finding morphological characters from their dataset to support molecular hypotheses of squamate evolution, noting a necessary 51–77 reversals ("cases in which iguanians must re-evolve the unmodified ancestral lepidosaurian conditions . . . ") for molecular hypotheses of the position of Iguania to be accurate (Gauthier et al., 2012). Additonally, Gauthier et al. (2012) found not a single synapomorphy (ambiguous nor unambiguous) that supported gene tree hypotheses for a nested Iguania within either Autarchoglossa or Scleroglossa, much less any part of Anguimorpha.

Here, we aim to address the incongruence between molecular and morphology-based hypotheses of squamate interrelationships. We investigate character support for the two major clades (i.e., earliest divergence within Squamata) by employing the opposite of the molecular scaffold approach (thus mapping molecular data onto

[1] The Field Museum of Natural History, 1400 S. Lake Shore Drive, Chicago, Illinois 60605; E-mail: cmcmahan@fieldmuseum.org. Send reprint requests to this address.

[2] LSU Museum of Natural Science, Department of Biological Sciences, Louisiana State University, 119 Foster Hall, Baton Rouge, Louisiana 70803.

[3] Department of Ecology and Evolutionary Biology, Tulane University, Boggs 400, New Orleans, Louisiana 70118; E-mail: lfreebor@tulane.edu.

[4] Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th St., New York, New York 10024; E-mail: wheeler@amnh.org.

[5] Department of Biological Sciences, Southeastern Louisiana University, SLU 10736, Hammond, Louisiana 70402; E-mail: bcrother@selu.edu.
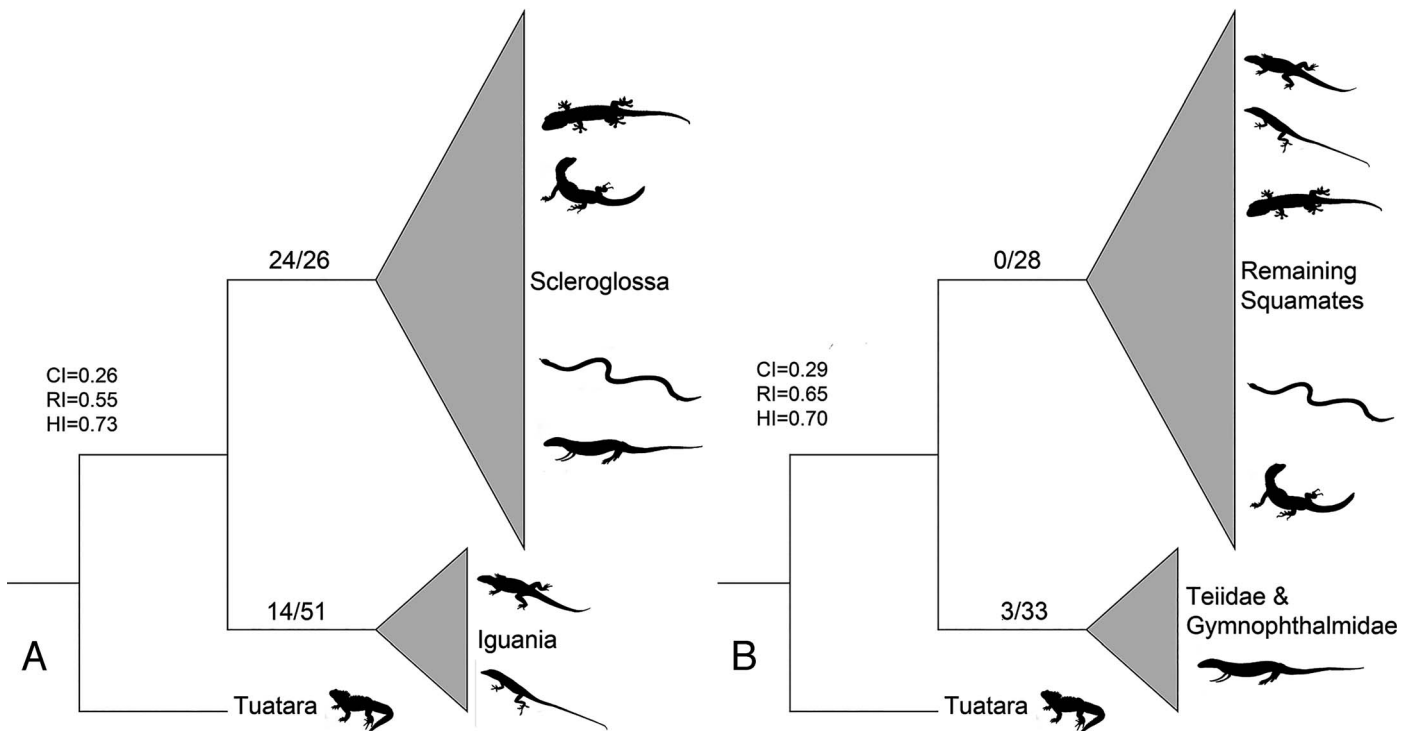
Fig. 1. Apomorphies from *c-mos* sequence data mapped onto (A) the morphological phylogeny of Conrad (2008) and (B) the molecular phylogeny of Eckstut et al. (2009). Numbers on branches indicate unambiguous/ambiguous character state changes.

a morphological topology). We first address whether there are molecular apomorphies for morphological hypotheses of squamate phylogeny, then compare results to support on molecular trees. By investigating morphological/molecular incongruence in squamate relationships, we present a method for examining similar scenarios of phylogenetic incongruence in other taxa.

MATERIALS AND METHODS

To employ our method, we chose two large molecular datasets. The molecular study of Eckstut et al. (2009) uses only the highly conserved proto-onco gene *c-mos*, but we chose to use the dataset and phylogeny from this study given that it is one of the most complete in terms of taxon sampling (>600 spp.) for a single gene. Second, we used the molecular study of Pyron et al. (2013), as this multi-locus dataset consists of 12 genes for 4161 species of squamates. The morphological phylogeny of Conrad (2008) was utilized because this study has the greatest depth of taxon sampling of both extant and extinct lineages. Given that the *c-mos* dataset is only one gene, taxa were compared between Conrad (2008) and Eckstut et al. (2009) to determine genera present in both analyses (*n* = 35). *C-mos* sequences for individuals (one per species for as many available species per genus) were taken from the original, aligned data file from Eckstut et al. (2009), and these sequences were compiled. Given the breadth of taxon and locus sampling, the entire Pyron et al. (2013) dataset was used.

Two tree topologies were constructed for each dataset using Mesquite (Maddison and Maddison, 2011)—one with the morphological Iguania + Scleroglossa relationship, followed by: Teiidae/Gymnophthalmidae + remaining Squamata (following *c-mos* tree) and Dibamidae + remaining Squamata (following Pyron et al., [2013] tree). Topologies are shown in Figures 1 and 2. Using PAUP* version 4.0a14

(Swofford, 2002), each molecular dataset was mapped onto both the molecular and morphological topologies via the Describe Trees > List of apomorphies option (ACCTRAN optimization). The numbers of unambiguous and ambiguous character state changes were recorded for each clade, as well as standard descriptive statistics for each topology (Consistency Index [CI], Homoplasy Index [HI], Retention Index [RI]).

RESULTS

Descriptive statistics (CI, HI, and RI) are relatively similar for both datasets, regardless of topology (Figs. 1, 2). For the *c-mos* dataset for squamates, results show that a number of molecular apomorphies support an Iguania clade and sister Scleroglossa clade (Fig. 1A), the most internal sister relationship consistently recovered in all morphological squamate phylogenies to date. There are 14 unambiguous (thus unique and unreversed) and 51 ambiguous character state changes diagnosing the Iguania, and 24 unambiguous and 26 ambiguous changes diagnosing its sister Scleroglossa. In stark contrast, however, there is far less support for the Teiidae/Gymnophthalmidae and remaining squamate clades recovered in the *c-mos* phylogeny (Fig. 1B). There are only three unambiguous and 33 ambiguous character state changes supporting the Teiidae/Gymnophthalmidae clade, and zero unambiguous and 28 ambiguous changes supporting the sister clade comprising the remaining Squamata.

For the multi-locus dataset of Pyron et al. (2013), a similar pattern was observed as for analysis of the *c-mos* dataset. Sixteen unambiguous and 954 ambiguous character state changes diagnose the Iguania, and 19 unambiguous and 977 ambiguous character state changes diagnose a sister Scleroglossa (Fig. 2A). Once again, there is less support for the Dibamidae and sister clade consisting of remaining
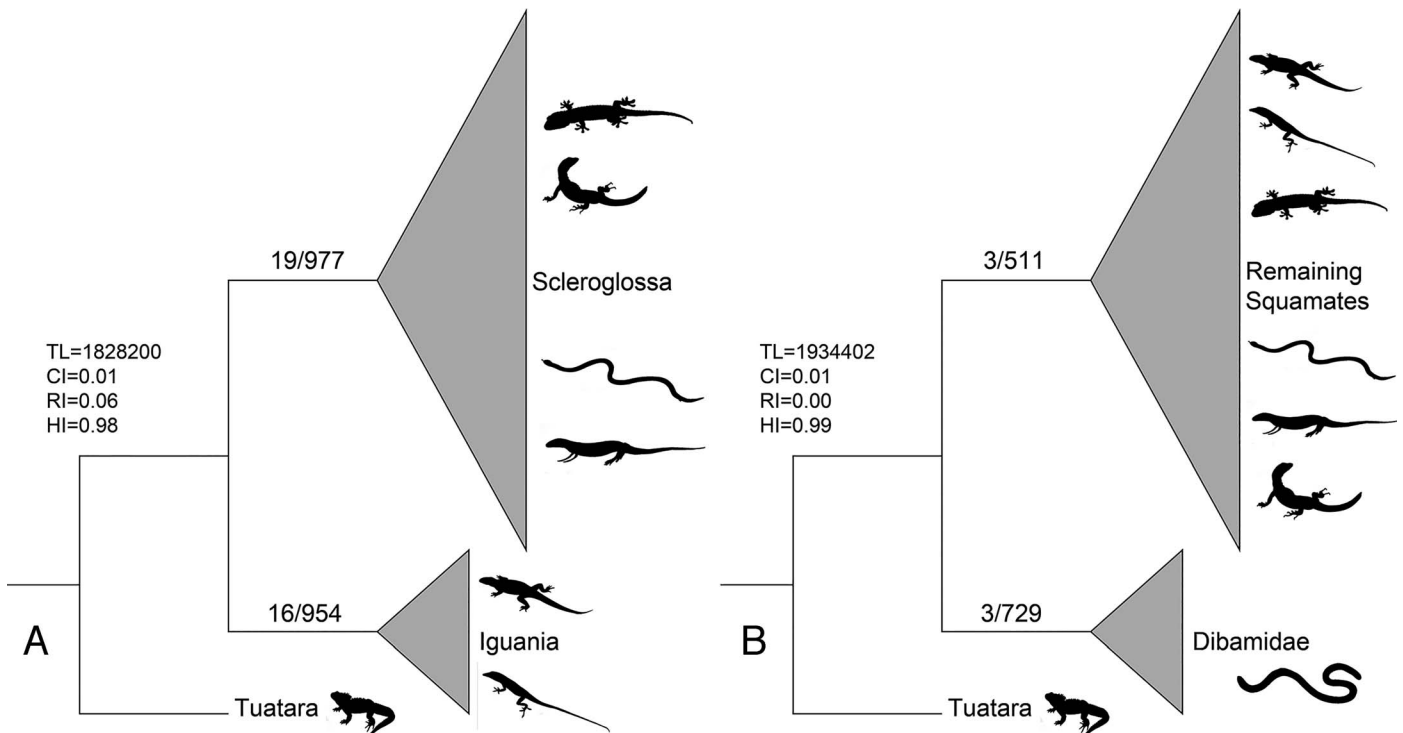
Fig. 2. Apomorphies from multi-locus dataset of Pyron et al. (2013) mapped onto (A) the morphological phylogeny of Conrad (2008) and (B) the molecular phylogeny of Pyron et al. (2013). Numbers on branches indicate unambiguous/ambiguous character state changes.

squamates, as recovered in Pyron et al. (2013). Only three unambiguous and 729 ambiguous character state changes support this position of Dibamidae, with three unambiguous and 511 ambiguous apomorphies supporting a sister clade consisting of other squamates (Fig. 2B).

DISCUSSION

Distant outgroups.—Our results show that numerous apomorphies from both the *c-mos* sequence dataset and multi-locus dataset unambiguously support a monophyletic lineage of Iguania, as well as its sister clade, the Scleroglossa. Despite different topologies inferred from phylogenetic analysis of the *c-mos* dataset and multi-locus dataset of Pyron et al. (2013), in both cases these molecular data seem to better fit the morphological topology (Conrad, 2008; and others), at least for the most internal relationships within squamates. We contend that molecular-based squamate phylogenies may be misled by the genetic divergence between the only reasonable outgroup (tuatara; Rhynchocephalia: *Sphenodon* spp.) and the ingroup. For the molecular datasets used in the present study, tuatara constitutes the longest branch of both respective trees. Furthermore, although Rhynchocephalia is still considered the closest extant sister lineage to Squamata within the Lepidosauromorpha, recent work has shown that the numerous differences between squamates and rhynchocephalians are no longer interpreted as plesiomorphic, but instead are considered derived (Conrad, 2008). The latter interpretation suggests a long evolutionary separation between rhynchocephalians and squamates, with the potential for several extinct lepidosauromorphan lineages; indeed, a number of fossil taxa exist that are phylogenetically intermediate between tuataras and squamates (e.g., Eolacertilia; however, detailed position of these lineages between tuataras and

squamates remains uncertain; Conrad, 2008). These critical taxa would be impossible to be accounted for in a molecular dataset and obviously contribute to the divergence between tuatara and squamates.

Wheeler (1990), regarding the inference of phylogeny with DNA sequence data, clearly showed that distant outgroups are essentially random data connections to the ingroup. The immediate consequence of such a connection would be a random placement of the root, and consequences of this are discussed at length in Hillis and Wiens (2000). Given that the distance from tuataras to Squamata is greater than the average distance across all Squamata, we can consider this outgroup distant and consider the root determined by it to be suspect. If *Sphenodon* is a functionally random outgroup, it could root the tree at the longest branch of the ingroup (Wheeler, 1990). To us, this is further convincing evidence that *Sphenodon* is a functionally random outgroup for these two molecular datasets and may explain the unusual relationships in this and other DNA-based trees. One important caveat to consider is that like most squamate phylogenetic studies, we have only a single outgroup taxon. In this case, the use of additional outgroup taxa (e.g., Wiens et al., 2012) only introduces even more distantly related lineages and exacerbates this issue.

Homoplasy.—A second possible explanation for the incongruence between molecular and morphological-based topologies concerns homoplasy. The claim has often been made that morphology is inferior to molecular data for phylogeny reconstruction because morphology is rife with homoplasy (Hedges and Sibley, 1994; Hedges and Maxson, 1996; Scotland et al., 2003; Wiens et al., 2010). However, concluding convergence of morphological traits based on the mapping of these traits onto a molecular tree (i.e., the molecular scaffold approach) only indicates convergence

based on that topology the data were fitted to—only discovering incongruence, which is not evidence of superiority of one dataset over another for phylogeny reconstruction. Based on our reverse scaffold approach, it appears there is so much homoplasy in the *c-mos* and multi-locus datasets that those characters easily distribute on the morphological tree as unambiguous and ambiguous synapomorphies for the earliest divergence within Squamata, with more characters than on the trees inferred from each dataset themselves. Additionally, numerous authors have argued that these patterns of homoplasy are what one would expect for misplaced taxa in a phylogeny (Gauthier et al., 1988, 2012; Donoghue et al., 1989).

Of course additional explanations for the incongruence (see Introduction above) are equally plausible and should also be investigated. Clearly the overwhelming evidence gathered from phylogenetic incongruence among hypotheses of squamate evolution, indicates that the amount of data (i.e., number of loci) is not the problem, nor is the taxon sampling (e.g., Gauthier et al., 2012; Pyron et al., 2013). Regardless of how these variables are adjusted, corroboration is never achieved between the two sources of data. Reeder et al. (2015) utilize a total evidence approach to study the phylogeny of Squamata and claimed to have solved the incongruence problem. However, our results show the problem is not resolved. It is interesting to us that they interpret the "unexpected placements for fossil taxa" as a better hypothesis than wondering if there may be a problem.

Most prior work has asked the question "why do morphological data not recover molecular relationships?" We ask the reverse "why do molecular data not recover the morphological relationships?" We argue that consideration of both these questions is an important and crucial step in attempting to understand incongruence between these two groups of hypotheses. The morphological scaffold approach identifies 50 *c-mos* characters and 996 characters across 12 loci supporting the Scleroglossa, thus excluding the Iguania from the Scleroglossa. While there are clearly synapomorphies supporting some of the internal molecular clades (thus similar descriptive statistics for both topologies within each dataset), the stark incongruence in the position of Iguania still begs explanation. We propose the hypothesis that the molecular data are incongruent with morphological hypotheses of squamate evolutionary relationships not because the molecular data are correct or that the morphology is rife with homoplasy, but because there may be an inherent analytical problem with the molecular data. What this problem is, we are uncertain, but we can point to the distant outgroup as a starting place.

These results should be of immediate interest to systematists, as well as others who utilize phylogenies, as we suspect that if this morphological scaffold method were employed on other scenarios of molecular-morphological phylogenetic incongruence, similar results may be obtained. Future work could be aimed at including vigorous statistical analyses to investigate this issue. However, the purpose of this study is to bring this issue more clearly to light and offer a starting hypothesis for discussion and future work. These results are not to say that molecular data lack merit or are uninformative, as numerous studies of phylogeny, phylogeography, and population differentiation have clearly illustrated; only, that there is without doubt evolutionary signal in analysis of morphological characters as clearly demonstrated here and in countless morphology-based phylogenetic analyses. Incongruence among phylogenetic hypotheses is not something that should be handled by simply dismissing the morphological ones, in preference for those based on molecules. We posit that this approach can be used to gain further insight into sources of discordance between morphological and molecular hypotheses of phylogeny, and allow for more accurate investigations into the evolutionary history of organisms.

LITERATURE CITED

Albert, E. M., D. San Mauro, M. García-París, L. Rüber, and R. Zardoya. 2009. Effect of taxon sampling on recovering the phylogeny of squamate reptiles based on complete mitochondrial genome and nuclear gene sequence data. Gene 15:12–21.

Betancur-R., R. 2009. Molecular phylogenetics and evolutionary history of ariid catfishes revisted: a comprehensive sampling. BMC Evolutionary Biology 9:175.

Bitsch, J., C. Bitsch, T. Bourgoin, and C. D'Haese. 2004. The phylogenetic position of early hexapod lineages: morphological data contradict molecular data. Systematic Entomology 29:433–440.

Burreson, E. M., and K. S. Reece. 2006. Spore ornamentation of *Haplosporidium nelsoni* and *Haplosporidium costale* (Haplosporidia), and incongruence of molecular phylogeny and spore ornamentation in the Haplosporidia. Journal of Parasitology 92:1295–1301.

Caldwell, M. W. 1999. Squamate phylogeny and the relationships of snakes and mosasauroids. Zoological Journal of the Linnean Society 125:115–147.

Camp, C. L. 1923. Classification of the lizards. Bulletin of the American Museum of Natural History 48:289–481.

Conrad, J. L. 2008. Phylogeny and systematics of Squamata (Reptilia) based on morphology. Bulletin of the American Museum of Natural History 310:1–182.

Crother, B. I., and W. F. Presch. 1992. The phylogeny of xantusiid lizards: the concern for analysis in the search for a best estimate of phylogeny. Molecular Phylogenetics and Evolution 1:289–294.

Donoghue, M. J., J. A. Doyle, J. Gauthier, A. G. Kluge, and T. Rowe. 1989. The importance of fossils in phylogeny reconstruction. Annual Review of Ecology, Evolution, and Systematics 20:431–460.

Eckstut, M. E., D. M. Sever, M. E. White, and B. I. Crother. 2009. Phylogenetic analysis of sperm storage in female squamates, p. 185–218. *In*: Animal Reproduction: New Research Developments. L. T. Dahnof (ed.). Nova Science Publishers, Inc., Happauge, New York.

Evans, S. E., and L. J. Barbadillo. 1998. An unusual lizard (Reptilia: Squamata) from the Early Cretaceous of Las Hoyas, Spain. Zoological Journal of the Linnean Society 124:235–265.

Evans, S. E., Y. Wang, and C. Li. 2005. The Early Cretaceous Chinese lizard, *Yabeinosaurus*: resolving an enigma. Journal of Systematic Paleontology 3:319–335.

Gauthier, J. A., M. Kearney, J. A. Maisano, O. Rieppel, and A. D. B. Behlke. 2012. Assembling the squamate tree of life: perspectives from the phenotype and the fossil record. Bulletin of the Peabody Museum of Natural History 53:3–308.

Gauthier, J., A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4:105–209.

Graham, S. W., J. R. Kohn, B. R. Morton, J. E. Eckenwalder, and S. C. H. Barrett. 1998. Phylogenetic congruence and disconrdance among one morphological and three molecular datasets from Pontederiaceae. Systematic Biology 47:545–567.

Hedges, S. B., and L. R. Maxson. 1996. Molecules and morphology in amniote phylogeny. Molecular Phylogenetics and Evolution 6:312–314.

Hedges, S. B., and C. G. Sibley. 1994. Molecules vs. morphology in avian evolution: The case of the ''pelecaniform'' birds. Proceedings of the National Academy of Sciences of the United States of America 91:9861–9865.

Hermsen, E. J., and J. R. Hendricks. 2008. W(h)ither fossils? Studying morphological character evolution in the age of molecular sequences. Annals of the Missouri Botanical Garden 95:72–100.

Hillis, D. M., and J. J. Wiens. 2000. Molecules versus morphology in systematics, p. 1–19. *In*: Phylogenetic Analysis of Morphological Data. J. Wiens (ed.). Smithsonian Institution Press, Washington, D.C.

Irestedt, M., M. Fjeldså, and G. G. P. Ericson. 2004. Phylogenetic relationships of woodcreepers (Aves: Dendrocolaptinae)—incongruence between molecular and morphological data. Journal of Avian Biology 35:280–288.

Lee, M. S. Y. 1998. Convergent evolution and character correlation in burrowing reptiles: towards a resolution of squamate relationships. Biological Journal of the Linnean Society 65:369–453.

Lee, M. S. Y., and M. J. Caldwell. 2000. *Adriosaurus* and the affinities of mosasaurs, dolichosaurs, and snakes. Organisms, Diversity, and Evolution 5:25–45.

Lee, M. S. Y., and A. B. Camens. 2009. Strong and surprising morphological support for the molecular evolutionary tree of mammals. Journal of Evolutionary Biology 22:2243–2257.

Losos, J. B., D. M. Hillis, and H. W. Greene. 2012. Who speaks with a forked tongue? Science 338:1428–1429.

Maddison, W. P., and D. R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. http://mesquiteproject.org

Mooi, R. D., and A. C. Gill. 2010. Phylogenies without synapomorphies—a crisis in fish systematics: time to show some character. Zootaxa 2450:26–40.

O'Leary, M. A., and A. Gatesy. 2008. Impact of increased character sampling on the phylogeny of Cetartiodactyla (Mammalia): combined analysis including fossils. Cladistics 24:397–442.

O'Leary, M. A., A. Gatesy, and M. J. Novacek. 2003. Are the fossil data really at odds with the molecular data? Morphological evidence for whale phylogeny (re)examined. Systematic Biology 52:853–864.

Pyron, R. A., F. T. Burbrink, and J. J. Wiens. 2013. A phylogeny and updated classification of Squamata, including 4161 species of lizards and snakes. BMC Evolutionary Biology 13:93.

Reeder, T. W, T. M. Townsend, D. G. Mulcahy, B. P. Noonan, P. L. Wood, Jr., J. W. Sites, Jr., and J. J. Wiens. 2015. Integrated analyses resolve conflicts over squamate reptile phylogeny and reveal unexpected placements for fossil taxa. PLOS ONE 10(3):e0118199. DOI: 10.1371/journal.pone.0118199

Scotland, R. W., R. G. Olmstead, and J. R. Bennett. 2003. Phylogeny reconstruction: the role of morphology. Systematic Biology 52:539–548.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0a14. Sinauer Associates, Sunderland, Massachusetts. http://paup.csit.fsu.edu

Townsend, T. M., A. Larson, E. Louis, and J. R. Macey. 2004. Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. Systematic Biology 53:735–757.

Vidal, N., and S. B. Hedges. 2005. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. Comptes Rendus Biologies 328:1000–1008.

Wheeler, W. C. 1990. Nucleic acid sequence phylogeny and random outgroups. Cladistics 6:363–367.

Wiens, J. J. 2004. The role of morphological data in phylogeny reconstruction. Systematic Biology 53:653–661.

Wiens, J. J., C. R. Hutter, D. G. Mulcahy, B. P. Noonan, T. M. Townsend, J. W. Sites, Jr., and T. W. Reeder. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. Biology Letters 8:1043–1046.

Wiens, J. J., C. A. Kuczynski, T. Townsend, T. W. Reeder, D. G. Mulcahy, and J. W. Sites, Jr. 2010. Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the placement of fossil taxa. Systematic Biology 59:674–688.

Wu, X.-C., D. B. Brinkman, and A. P. Russell. 1996. *Sineoamphisbaena hexatabularis*, an amphisbaenian (Diapsida: Squamata) from the Upper Cretaceous redbeds at Bayan Mandahu (Inner Mongolia, People's Republic of China), and comments on the phylogenetic relationships of the Amphisbaenia. Canadian Journal of Earth Science 33:541–577.