The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2015, **116**, 1–12. With 3 figures.

Diversification of North American natricine snakes

JOHN D. MCVAY[1]*, OSCAR FLORES-VILLELA[2] and BRYAN CARSTENS[3]

[1]*Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA*
[2]*Facultad de Ciencias, Museo de Zoologia, Universidad Nacional Autonoma de Mexico, Circuito Extrior de CU, Mexico, DF, 04510, Mexico*
[3]*Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 318 W. 12th Avenue, Columbus, OH, 43210-1293, USA*

Received 14 November 2014; revised 31 March 2015; accepted for publication 31 March 2015

The phylogeny of thamnophiine snakes has not been fully resolved, largely because previous phylogenetic estimates have been based on incomplete taxon sampling or relied solely on mitochondrial sequence data. To address this deficiency, we collected data from multiple autosomal loci collected from 50 taxa before estimating the most robust phylogeny of Thamnophiini to date. Our findings clarify the relationships of taxa not previously included in molecular analyses and also lend evidence to previously recommended taxonomic revisions. Differences in topological estimates between competing models of evolution were minimal and not strongly supported; however, a multispecies coalescent model of evolution was highly favoured over a concatenated model based on marginal likelihood estimates. Additionally, we estimated the timing of divergence among the three major lineages to have occurred during the Miocene period (approximately 11–14 Mya), followed by a decline in speciation rates in all major lineages. © 2015 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2015, **116**, 1–12.

ADDITIONAL KEYWORDS: Bayes factors – multilocus nuclear – path sampling – Thamnophiini.

INTRODUCTION

Although molecular phylogenetic estimates provide vital data pertaining to the relationships among organisms, the utility of phylogenetic work is enhanced when the phylogenies are incorporated into downstream analyses. For example, comparative methods (i.e. the optimization of organismal features on the phylogenetic estimate) can provide insights regarding phenotypic evolution, particularly when phylogenetic independent contrasts (Felsenstein, 1985) are utilized. Understanding the timing of evolution by tracking rates of cladogensis can improve our understanding of species diversification. In addition to providing a historical context for interpreting the evolution of organismal features, phylogenies aid the researcher in understanding branching patterns and identifying the factors that promoted diversification. When combined, analytical tools that track character state evolution and the diversification of lineages through time improve our comprehension of both the pattern and process of evolutionary diversification. In the present study, we apply these tools to the thamnophiine snakes, a group of vertebrates that have diversified into a variety of feeding niches, aiming to learn about the timing of the radiation and the evolution of feeding specialization in this group.

Within the macrostomatan snakes, Thamnophiini (58 currently recognized species) represents the natricine subfamily of colubrid snakes in the Western hemisphere. This large radiation is traditionally classified (based largely on morphology) into nine genera that span from Canada to Costa Rica and occupy a variety of montane to estuarine habitats. Many thamnophiine snakes are diet specialists, including those whose prey choice is restricted to soft prey, such as earthworms and slugs, and those that prefer hard prey, such as crayfish. Most species are closely associated with water, either as their primary habitat or as a source of prey for both aquatic and terrestrial foragers (Gibbons & Dorcas, 2004; Rossman, Ford & Seigel, 1996). Molecular phylogenetic data not only support many previously hypothesized clades, but also suggest that several of the clades inferred from morphological data are paraphyletic. Notable examples of paraphyly include the inclusion

*Corresponding author. E-mail: johndmcvay@gmail.com

of *Thamnophis validus* in the genus *Nerodia* (de Queiroz & Lawson, 1994) and the non-monophyly of the genus *Regina* (Price, 1983; Rossman, 1985).

Previous genetic investigations focusing on the relationships among the thamnophiines include the allozyme studies by Lawson (1985) and de Queiroz & Lawson (1994), as well as the mitochondrial sequence-based works of Alfaro & Arnold (2001) and de Queiroz, Lawson & Lemos-Espinal (2002). Important generic-level taxonomic discoveries were made from each of these studies. For example, phylogenies presented by both Lawson (1985) and Alfaro & Arnold (2001) are inconsistent with the hypothesis that both *Regina* and *Virginia* are monophyletic, and evidence from de Queiroz *et al.* (2002) suggests that the distinctive mountain meadow snakes (genus *Adelophis*) are nested within the garter snakes (*Thamnophis*). Alfaro & Arnold (2001) designated three major lineages in the thamnophiine snakes: the garter snakes (*Thamnophis*), the water snakes (*Nerodia*, *Regina grahamii* and *Regina septemvittata*, and *Tropidoclonion*), and the semifossorial snakes (*Clonophis*, *Liodytes alleni* and *Liodytes rigida*, *Seminatrix*, *Storeria*, and *Virginia*). Despite these findings, key taxonomic hypotheses have not been confirmed using a multilocus phylogenetic estimation of the North American natricine snakes. We seek to generate such an estimate and use it to evaluate the key sources of conflict between morphological data and previous molecular work. Finally, we examine the pattern and timing of diversification across the tribe.

MATERIAL AND METHODS

Data collection

Tissues from 52 specimens representing 51 species (50 ingroup + *Natrix natrix*, a European natricine utilized as an outgroup) were obtained primarily from museum collections (see Appendix, Table A1). DNA was extracted from liver or muscle via a modified salt-saturation protocol (Aljanabi & Martinez, 1997), in which tissue was lysed using 300 μL of PureGene Cell Lysis solution (catalogue number 158906; Qiagen) followed by overnight incubation with proteinase K (P8102S; New England Biolabs).

We used the polymerase chain reaction (PCR) to amplify five (one mitochondrial and four nuclear) gene-coding loci for each individual (Table 1). Additionally, anonymous nuclear markers were developed for the present study by screening a fragment library previously prepared for microsatellite discovery, *sensu* Glenn & Schable (2005). Initial fragments were selected that were determined not to contain variable number tandem repeat regions, as detected by the *abblast* function of repeatmasker (Smit, Hubley & Green, 2013–2015), We then developed primers using PRIMER3 (Rozen & Skaletsky, 2000) and tested these primers in a four taxon test set (*Natrix*, *Nerodia*, *Storeria*, and *Thamnophis*). After testing for amplification across all taxa, five fragments were ultimately selected for sequencing.

PCR amplification of fragments was performed with reagent proportions: 0.4–1 ng μL^{-1} tDNA, 0.4 μM each primer, 0.2 μM dNTPs, 1 × Standard Taq reaction buffer (New England Biolabs) and 0.5 units of Taq DNA polymerase (M0267; New England Biolabs) per 25 μL final volume. Thermocycling conditions were optimized for primer melting temperature and target fragment length (Table 1). Bidirectional sequencing for both coding and anonymous fragments was performed using Big Dye, version 3.1 (Applied Biosystems) in accordance with the manufacturer's instructions. Sequencing was performed on an ABI 3130 genetic analyzer (Applied Biosystems) at the LSU Genomics Core Facility (Baton Rouge, LA, USA) and Beckman Coulter Genomics (Danvers, MA, USA), Chromatograms were examined by eye and edited using SEQUENCHER, version 4.8 (Gene Codes). Alignment of loci was conducted using MUSCLE (Edgar, 2004a, b), under the default settings.

Phylogeny estimation and divergence dating

A Bayesian estimate of phylogeny was generated using BEAST, version 1.8 (Drummond *et al.*, 2012). Prior to analysis, we selected site models for each locus using DT-MODSEL (Minin *et al.*, 2003) and PAUP* (Swofford, 2003). Optimal models for each locus were defined in BEAST, with the exception of models that contained both a gamma-distributed rates and invariant sites (Γ + I). Such models were simplified to use only the gamma distribution to describe rate variation as a result of the potential interference between variables (Yang, 2006). Each gene was allowed to evolve under an independent, uncorrelated relaxed lognormal clock, with each sample mean drawn from a uniform prior with range 0–100 to allow for differences in substitution rate among genes. Two identical Markov chain Monte Carlo (MCMC) runs of 2×10^8 steps (sampling every 2×10^4) were performed and posterior distributions of parameters were compared for convergence using TRACER, version 1.5 (Rambaut & Drummond, 2009). We recognize the importance of testing differing models of evolution when conducting phylogenetic studies; in particular, a coalescent-based species tree approach may be equally appropriate for these data (Edwards, 2009; Edwards, Liu & Pearl, 2007). To explore whether the model of evolution would significantly affect the

Table 1. Primer and thermocycler information for each locus

Gene	Oligo (5′- to 3′)	T_A	E	Reference
BDNF	F GACCATCCTTTTCCTKACTATGGTTATTTCATACTT R CTATCTTCCCCTTTTAATGGTCAGTGTACAAAC	50	:30	Leache & McGuire (2006)
FSHR	F CCDGATGCCTTCAACCCVTGTGA R CCRAAYTTRCTYAGYARRATGA	50	:30	Wiens *et al.* (2008)
ND4	F TGACTACCAAAAGCTCATGTAGAAGC R TTTTACTTGGATTTGCACCA	55	:30	Forstner, Davis, Arevalo (1995) Skinner *et al.* (2005)
NT3	F ATGTCCATCTTGTTTTATGTGATATTT R ACRAGTTTRTTGTTYTCTGAAGTC	50	:30	Wiens *et al.* (2008)
R35	F TCTAAGTGTGGATGATYTGAT R CATCATTGGRAGCCAAAGAA	50	:30	Fry *et al.* (2006)
'E'	F CTGGATCCATAGCTCCTGGT R ATTTTCAACCCAGCTTTTGG	52	:20	Present study
'I'	F GGGAAAAAGAGGGAAATTGG R GTGAAGGGTTTTGGGTGTTG	52	:20	Present study
'K'	F GCCACCCTGACACTAAAAACA R TTCCTGGAAGATGGTTTTGC	52	:20	Present study
'M'	F TGAATGAGGCTGCGAGATTA R AGGGGAGCCAGGTGTAACTT	52	:20	Present study

The locus, sequence of the primer (5′- to 3′), annealing temperature (T_A; °C), time of elongation (E), and source of the primer are shown.

topological outcome, we conducted two independent runs of coalescent-based species tree estimation using *BEAST; we then compared topologies from the estimates under both concatenation and multispecies coalescence to assess the degree of congruence and discordance supported by high posterior probabilities. To further assess to which model the data was a better fit, we employed the path sampling (Xie *et al.*, 2011) based marginal likelihood estimator, which allows for a direct comparison of competing models via Bayes factors, as implemented in BEAST (Baele *et al.*, 2012, 2013).

Fossil age estimates can serve as calibration points for estimating divergence times across phylogenies (Heath, Huelsenbeck & Stadler, 2014), which in turn can be incorporated into chronograms that can be used to track diversification through time. The fossil record of Thamnophiini is rich (Holman, 2000), with fossils known from each of the three major lineages as well as extinct lineages (e.g. *Neonatrix*). To estimate divergence times across lineages, we performed two additional MCMC runs in BEAST, allowing rates for all loci to be drawn from a uncorrelated relaxed lognormal clock, and incorporating two fossils representing the oldest known occurrences of *Nerodia* and *Thamnophis*, both from the Medial Barstovian (13–14.5 Mya) (Voorhies, 1990) fossil age in the Miocene (Holman, 2000) (see Appendix, Table A2). We enforced monophyly on each of the above nodes, and allowed ages to be drawn from a normal distribution (mean = 14 Mya, SD = 1). All other parameters were set identically to the initial, nonfossil-constrained analysis. Two identical chains were allowed to run until effective sample sizes were above 200, and we compared trace files of both runs to ensure that runs had converged.

Estimating rates of diversification

Modelling rates of speciation and extinction across chronograms allows for inference of shifts in rates of diversification and phenotypic evolution. We used BAMM (Rabosky *et al.*, 2013) and the R packages BAMMtools (Rabosky *et al.*, 2014) and APE (Paradis, Claude & Strimmer, 2004) to assess patterns of diversification rates among the ingroup. Specifically, we used the Bayesian software BAMM to test for evidence of shifts in the rate of speciation or extinction within and among clades. Utilizing the BEAST-generated chronogram, we ran the 'speciationextinction' MCMC for four chains of 1×10^6 generations, sampling from a poisson rate prior of 1.0 and a global sampling fraction at 0.9 to account for extant diversity missing from the analysis; all other prior and operators were set to default settings. Sufficient convergence of chains was assessed by estimated sample size, viewed in TRACER. The output from BAMM was then input into BAMMtools to compare competing rate-shift scenarios via Bayes factors, as well as to compare rates among clades.

RESULTS

Sequence analysis

The proportion of variable sites differed across loci (Table 2), with a mean length of 418 base pairs and a mean of 72 variable sites per locus. The overall quality of the gene tree estimates (as determined by the nodal support values) varied among gene trees (not shown) in a manner correlated with the amount of sequence variation in each locus. Of note, inspection of the DNA sequence data and gene tree estimate of the anonymous locus *K* suggested that multiple loci were being amplified because particular insertion/deletions were shared among polyphyletic assemblages spanning major lineages; therefore, this locus was not included in the concatenated analysis. A fifth anonymous fragment was sequenced but was excluded from further analysis because it contained a single heterozygous across all ingroup individuals, implying that paralogous loci were coamplified. Missing data were filled in where possible from Genbank data (see Appendix, Tables A1, A2).

Phylogeny, divergence, and diversification

We generated two maximum clade credibility trees from posterior topologies from the concatenated and multispecies coalescent BEAST analyses (Fig. 1). Topologies of these two estimates were largely concordant; topological conflict was not mutually well-supported. The concatenation-based estimate showed high support across most nodes in the tree; *Nerodia* was highly supported as monophyletic, with Bayesian posterior probabilities (BPP) = 1.0. *Thamnophis* was estimated as paraphyletic with *Adelophis* (BPP = 1.0); a third clade, including the genera *Clonophis*, *Storeria*, *Virginia*, *Haldea*, and *Liodytes*, was highly supported (BPP = 1.0). The remaining ingroup taxa, *R. grahamii*, *R. septemvittata* and *Tropidoclonion lineatum*, fell outside the three aforementioned lineages and were not highly supported in their placement. The phylogeny estimated using a multispecies coalescent model was similar with lower overall support across the tree, with the placement of the three taxa above rearranged with respect to the concatenated estimate. To determine which phylogenetic estimation method was most appropriate, we calculated the marginal likelihoods estimated via path sampling for both the concatenated estimate and the multispecies coalescent model estimate. The marginalized likelihood favoured the species tree over concatenation (mean concatenation m*ln*L = −14 327; mean species tree m*ln*L = −13 903; $K = 424$; for natural log Bayes factors, a value above $2 \times lnK > 10$ is considered as very strong support – Kass & Raftery, 1995).

The estimated timing of diversification among the major lineages indicates that the ancestors of extant thamnophiine snakes diversified quickly during the Miocene. The estimated divergence times among major lineages are broadly overlapping (Fig. 2), with most of the diversification estimated to have occurred between during the Miocene (approximately 11–14 Mya). The ancestral node of Thamnophiini is estimated to have occurred 15 ± 1 Mya, consistent with divergence events following the appearance of the oldest known North American natricine, *Neonatrix elongata* (Holman, 1973). Speciation rate analysis in BAMM detected a decrease in speciation rate subsequent to the initial divergence of the major lineages in Thamnophiini (Fig. 3A). Speciation (lambda) and extinction (mu) rates were similar across major lineages (Fig. 3B, C), with no detectable shifts (no scenarios with greater than zero shifts had appreciable Bayes factors support over a zero shift scenario) in diversification rate.

Table 2. Summary of the DNA fragments analyzed

Gene	bp	*s*	Model
BDNF	557	31	K80 + G
FSHR	511	39	HKY + I
ND4	614	270	K80 + G
NT3	561	85	K80 + G
R35	645	81	HKY + G
E	196	44	HKY
I	209	34	K80
K	243	37	HKY + I
M	228	31	HKY + I
Total	4199	708	N/A

bp, final length of edited fragment used in the present study; *s*, total number variable sites.

DISCUSSION

We estimated the most species- and data-rich molecular phylogeny of Thamnophiini to date. Our results largely agree with previously estimated molecular phylogenies based solely on mitochondrial data and many of the relationships within and among genera are well-supported by our analyses, with the exceptions discussed below. Our data represent the first nuclear sequence data published for most species represented in our dataset, some of which were previously without any molecular data.

Divergence and diversification

Despite the relatively young age and amount of diversity present in this clade, the results of the

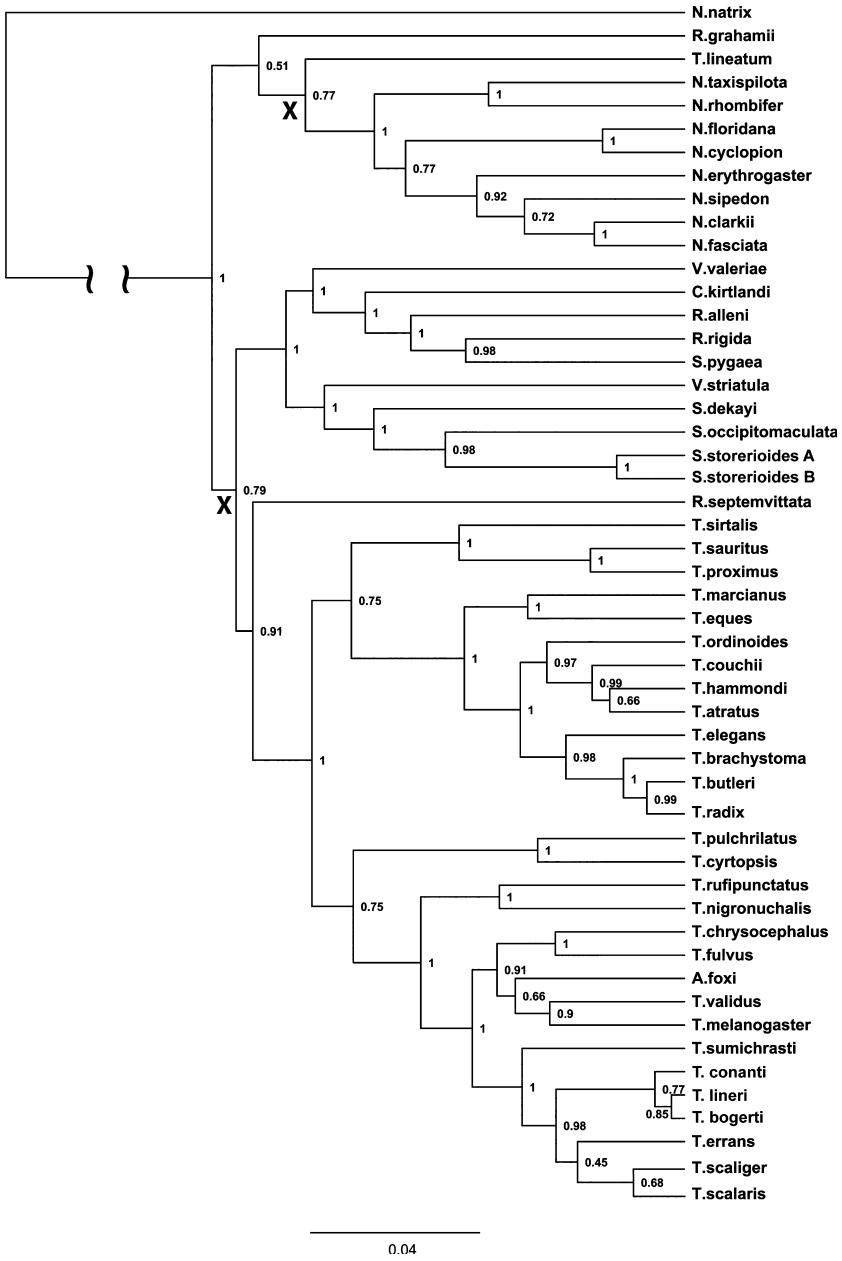

Figure 1. Multilocus Bayesian estimate of phylogeny (maximum clade credibility tree) of Thamnophiini. Branch lengths based on ND4 rate estimated by BEAST. An 'X' indicates conflicting topology with multispecies coalescent estimate.

divergence dating and rate analyses suggest that the rates of diversification have been declining subsequent to an initial increase. Burbrink & Pyron (2010) and Burbrink *et al.* (2012) estimated a decrease in diversification rates during the Pliocene for several squamate groups (including Thamnophinii; results presented here are consistent); all existing data are inconsistent with the Pleistocene speciation phenomenon seen in other North American taxa (Bermingham *et al.*, 1992; Knowles, 2000; Levsen, Tiffin & Olson, 2012). Burbrink *et al.* (2012) found this rate scheme in Thamnophiini and other squamate groups to be consistent with a deterministic diversification model (Nee, Mooers & Harvey, 1992) in which the rate of speciation slows as niches are filled across space. This pattern of rate decline appears to be repeated across each major lineage within Thamnophiini. However, the appearance of a decline may belie cryptic diversity (Ruane *et al.*, 2014) undiscoverable within the scope of the present

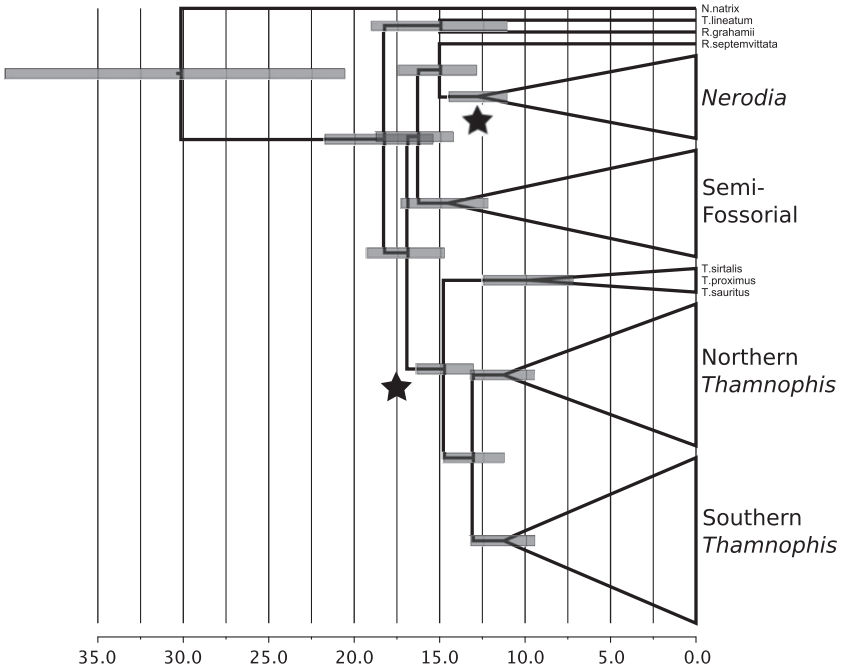

Figure 2. Chronogram of Thamnophiini, based on fossil calibrations of *Nerodia* and *Thamnophis* (indicated by stars). Error bars denote node age 95% posterior distribution. The estimated age of each node is given by the scale on the *x*-axis, in units of millions of years.

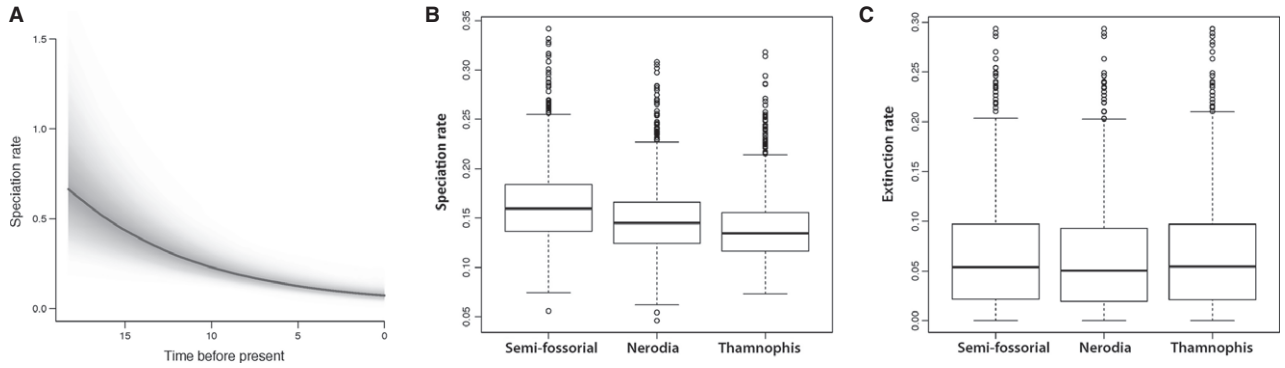

Figure 3. Diversification rate analyses visualized using BAMMtools. A, rate of diversification across Thamnophiini across time. Box plots representing posterior distributions of (B) speciation rate (lambda) and (C) extinction rate (μ) across each major clade.

study; such diversity may yet be revealed with more robust sampling within each taxon.

Phylogenetic estimation

A primary goal of the present study was to develop the most comprehensive phylogenetic estimate of the North American natricine snakes. We met this goal by including novel genetic data from eight independently evolving genetic loci.

The continual addition of taxa and characters to estimates of phylogeny serves to improve our model of the relationships of these organisms and, more broadly, our understanding of the nature of cladogenesis. Of equal importance is that a more robustly-estimated phylogeny often contains less uncertainty and can bolster the statistical confidence of studies (e.g. diversification and ancestral state estimation) that incorporate these estimates. In our case, the inclusion of multiple nuclear loci allows us to gather information from regions of the genome that are evolving at different rates, mitigating the potential effects of substitution saturation and increasing the potential for presence of parsimony informative sites across the depth of the tree. To assess the fit of two competing models of evolution,

we compared marginal likelihood estimate for each using Bayes factors. The coalescent model was highly favoured as a better fit to the data; however, the coalescent estimate was not well-resolved at multiple nodes across the tree. This uncertainty is likely a result of the low amount of variability in some of the nuclear markers sampled. For those nodes at which there was conflict with the concatenated estimate, posterior support was low in the coalescent estimate, suggesting limited information rather than effects of incomplete lineage sorting is a more likely cause of discordance. These results make for a difficult decision: whether to accept the poorly supported estimates of the better-fit model or the more highly supported estimate where the model is potentially mis-specified. We ultimately incorporated a concatenated model for divergence time estimation. The results of divergence estimation under a coalescent model (not shown) yielded what we considered to be poor estimates of deeper nodes in the phylogeny, such that the common ancestor of *Natrix* and Thamnophiini was estimated to have lived several million years (> 35 Mya) before the origin of the natricines (Rage, 1988; Guo *et al.*, 2012).

Despite inconsistencies, both estimates largely agree with the findings of the previously published molecular-based studies, with the exceptions noted below. Note also that we have partitioned our discussion of the taxonomic implications of the present study in accordance with the designations of Alfaro & Arnold (2001).

Garter snakes

Three well-supported clades were recovered within *Thamnophis*: two broadly-distributed clades, and one composed of species found mostly in México, Guatemala, and Honduras. Our samples also include three of the most recent additions to *Thamnophis*: *Thamnophis lineri*, *Thamnophis bogerti*, and *Thamnophis conanti*; formerly populations of *Thamnophis godmani* that were elevated to species status based on allopatry and morphological evidence (Rossman & Burbrink, 2005). Genetic divergence was evident but qualitatively lower than expected for distinct species (< 1% uncorrected pairwise sequence divergence for each comparison). However, these species are each represented by a single individual in the present study; more thorough genetic sampling would be needed to appropriately characterize the status of these lineages. Relationships estimated in the present study differed somewhat from those estimated by de Queiroz *et al.* (2002). Particularly, the pair of *Thamnophis cyrtopsis* and *Thamnophis pulchrilatus* are recovered as sister to the southern clade in the present study Fig. 1), whereas this pair is recovered as sister to the more northern clade in the previous study. In addition to the incorporation of different loci in each study, these topological inconsistencies may also be a result of differences in taxon sampling. Fox's mountain meadow snake (*Adelophis foxi*; Rossman & Blaney, 1968) was strongly supported (BPP > 0.95 at four internodes) as nested within *Thamnophis*. Despite these results, we are hesitant to make the suggestion that the genus *Adelophis* be synonymized with *Thamnophis* because this taxon is represented in our study by the specimen (LSUMZ 40848) used in de Queiroz *et al.* (2002) and thus subject to the same caveats discussed in their study. However, our DNA was processed from a different aliquot of tissue, dispelling the possibility of bias as a result of PCR contamination, as discussed by de Queiroz *et al.* (2002). The addition of sequence data from another individual of this species, as well as data from its congener, *Adelophis copei*, would serve to clarify the placement of this genus; however, these species are rarely encountered in the wild, and tissue for DNA extraction is unavailable in collections.

Water snakes

Nerodia was estimated as monophyletic with strong support. Our findings disagree with the relationships among species within *Nerodia* with those estimated by Alfaro & Arnold (2001); specifically, supported nodes from our estimate are in conflict with the hypothesis that *R. grahamii*, *R. septemvittata*, and *Tropidoclonion* are nested within *Nerodia*. This finding is not particularly unexpected given the lack of strong support for this particular grouping in the former study. However, we cannot reject the hypothesis that *Nerodia* and these three taxa form a monophyletic assemblage, although McVay & Carstens (2013) did reject a sister relationship of *R. grahamii* and *R. septemvittata* based on gene-by-gene tests of monophyly. This suggests that there are four independent origins of crayfish predation: *Liodytes*, *R. grahamii*, *R. septemvittata*, and a population of *Thamnophis melanogaster* in México (Manjarrez, 2005; Manjarrez, García & Drummond, 2013). Missing from the present study is *Nerodia harteri*, recovered as sister to *Nerodia sipedon* by Alfaro and Arnold. Qualitatively, both diet and habitat preference appear to be labile within this group. These two traits are undoubtedly linked, and body size, which is correlated with prey type in snakes (Pyron & Burbrink, 2009; Rodríguez-Robles, Bell & Greene, 1999), also appears across the major lineages of snakes to be qualitatively linked with habitat, particularly with degree of aquatic habitat use (the more aquatic habitat use, the larger the snakes; Burbrink *et al.*, 2012; Gibbons & Dorcas, 2004; Rossman *et al.*,

1996). Interestingly, qualitative ecological space appears to somewhat non-overlapping among the major lineages, suggesting a shift during initial divergence; however, there is a proclivity for *Thamnophis* to converge with the more aquatic *Nerodia* phenotype (e.g. *Thamnophis rufipunctatus*, *T. validus*). Unlike other major macrostomatan snake radiations in North America, there is little known specialization in this group to an ophiophagic (but see Manjarrez, Venegas-Barrera & GarcÍa-Guadarrama, 2007) or arboreal ecology, suggesting that, although species in this group are variable in niche space, they may be limited by the overall bauplan of the clade; this is apparently true of other radiations of Natricinae (African, European, Asian), although the diversity and biology of these clades is not understood to the extent of that of their North American counterpart.

Semi-fossorial clade

Although including more species than Alfaro and Arnold, we estimated a phylogeny consistent with their findings, including the paraphyletic nature of *Liodytes*. Our findings are also consistent with those of McVay & Carstens (2013), who rejected the monophyly of *Virginia*, based on multiple gene tree-based tests of monophyly. We recovered with high support that *Virginia valeriae* has a sister relationship to *Clonophis*, *Liodytes*, and *Seminatrix*, and *Haldea striatula* as sister to *Storeria*. Absent from our sampling is *Storeria hidalgoensis* (Taylor, 1942); however, the validity of this species is questionable because it is considered to be a synonym of *Storeria occipitomaculata* (Trapido, 1944).

Robustness of phylogeny estimation is dependent on the quantity and quality of the data employed and, although we have collected the largest dataset to date in Thamnophinii, we anticipate that the data available for phylogeny reconstruction in this group will increase dramatically as studies incorporate high-throughput sequencing (McCormack *et al.* 2013). Although genome-scale sequencing will likely improve our understanding of the broader relationships, it will also contribute to the need for finer scale genetic exploration both within and among species. To date, phylogeographical and/or population genetic results have been published for only a handful of the currently recognized species in this group, including *Thamnophis nigronuchalis* and *T. rufipunctatus* (Wood *et al.*, 2011), *Thamnophis sirtalis*, (Janzen *et al.*, 2002), *Thamnophis elegans* (Manier & Arnold, 2005), *Thamnophis proximus* (Allen, 2005), *T. validus* (de Queiroz & Lawson, 2008), *Nerodia clarkii* (Jansen, Mushinsky & Karl, 2008) *Nerodia erythrogaster* (Makowsky *et al.*, 2010), and *Nerodia rhombifer* (Brandley *et al.*, 2010). Almost nothing is known about the phylogeography of any members of Thamnophiini outside of *Nerodia* and *Thamnophis*. Of equal importance is the need for continued research into the ecology, morphology, and behaviour of this group. These data will be critical in developing a complete understanding of the Thamnophiini because they will (1) help to validate species boundaries and (2) hybrid zones, at the same time as leading to an increased understanding of how habitat and climatic change have influenced the evolution of this group.

ACKNOWLEDGEMENTS

We thank Christopher Blair, Jeff Boundy, Eric Green, Duncan Hauser, Paul Manos, Jose Meireles, and Alex Pyron for helpful comments and discussion. In addition, we thank Christopher Austin, James Maley, and Donna Dittmann for facilitating access to sample tissues; Carl Franklin and Jonathan Campbell for help with specimen identification; and Tara Pelletier and Jordan Satler for facilitating data collection. Finally, we thank two anonymous reviewers whose comments improved the manuscript.

REFERENCES

Alfaro ME, Arnold SJ. 2001. Molecular systematics and evolution of Regina and the Thamnophiine snakes. *Molecular Phylogenetics and Evolution* **21:** 408–423.

Aljanabi SM, Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25:** 4692–4693.

Allen L. 2005. Phylogeography of five subspecies of western ribbon snakes (*Thamnophis proximus*) in the U.S. and middle America. Unpublished MS, The University of Texas at Arlington.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29:** 2157–2167.

Baele G, Li WL, Drummond AJ, Suchard MA, Lemey P. 2013. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Molecular Biology and Evolution* **30:** 239–243.

Bermingham E, Rohwer S, Freeman S, Wood C. 1992. Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: a test of Mengel's model. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 6624–6628.

Brandley MC, Guiher TJ, Pyron RA, Winne CT, Burbrink FT. 2010. Does dispersal across an aquatic geographic barrier obscure phylogeographic structure in the diamond-backed watersnake (*Nerodia rhombifer*)? *Molecular Phylogenetics and Evolution* **57:** 552–560.

Burbrink FT, Pyron RA. 2010. How does ecological opportunity influence rates of speciation, extinction, and morphological diversification in New World ratsnakes (tribe Lampropeltini)? *Evolution* **64:** 934–943.

Burbrink FT, Chen X, Myers EA, Brandley MC, Pyron RA. 2012. Evidence for determinism in species diversification and contingency in phenotypic evolution during adaptive radiation. *Proceedings of the Royal Society of London Series B, Biological Sciences* **279:** 4817–4826.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Edgar RC. 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Edgar RC. 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32:** 1792–1797.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63:** 1–19.

Edwards SV, Liu L, Pearl DK. 2007. High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 5936–5941.

Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* **125:** 1–15.

Forstner MR, Davis SK, Arevalo E. 1995. Support for the hypothesis of anguimorph ancestry for the suborder Serpentes from phylogenetic analysis of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **4:** 93–102.

Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SFR, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E. 2006. Early evolution of the venom system in lizards and snakes. *Nature* **439:** 584–588.

Gibbons JW, Dorcas ME. 2004. *North American watersnakes: a natural history*. Norman, OK: University of Oklahoma.

Glenn TC, Schable NA. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* **395:** 202–222.

Guo P, Liu Q, Xu Y, Jiang K, Hou M, Ding L, Pyron RA, Burbrink FT. 2012. Out of Asia: natricine snakes support the Cenozoic Beringian dispersal hypothesis. *Molecular Phylogenetics and Evolution* **63:** 825–833.

Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences* **111:** E2957–E2966.

Holman JA. 1973. Reptiles of the Egelhoff local fauna (Upper Miocene) of Nebraska. *Contributions from the Museum of Paleontology* **24:** 125–134.

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

Jansen KP, Mushinsky HR, Karl SA. 2008. Population genetics of the mangrove salt marsh snake, *Nerodia clarkii* compressicauda, in a linear, fragmented habitat. *Conservation Genetics* **9:** 401–410.

Janzen FJ, Krenz JG, Haselkorn TS, Brodie ED. 2002. Molecular phylogeography of common garter snakes (*Thamnophis sirtalis*) in western North America: implications for regional historical forces. *Molecular Ecology* **11:** 1739–1751.

Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the American Statistical Association* **90:** 773–795.

Knowles LL. 2000. Tests of pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. *Evolution* **54:** 1337–1348.

Lawson R. 1985. Molecular studies of thamnophiine snakes. Unpublished PhD Dissertation, Louisiana State University.

Leache AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (Phrynosoma) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. *Molecular Phylogenetics and Evolution* **39:** 628–644.

Levsen ND, Tiffin P, Olson MS. 2012. Pleistocene speciation in the genus *Populus* (Salicaceae). *Systematic Biology* **61:** 401–412.

Makowsky R, Marshall JC Jr, McVay J, Chippindale PT, Rissler LJ. 2010. Phylogeographic analysis and environmental niche modeling of the plain-bellied watersnake (*Nerodia erythrogaster*) reveals low levels of genetic and ecological differentiation. *Molecular Phylogenetics and Evolution* **55:** 985–995.

Manier MK, Arnold SJ. 2005. Population genetic analysis identifies source-sink dynamics for two sympatric garter snake species (*Thamnophis elegans* and *Thamnophis sirtalis*). *Molecular Ecology* **14:** 3965–3976.

Manjarrez J. 2005. Posible invasión de unnicho alimentario nuevo y microevolución en una especie mexicana de serpiente. *Ciencia Ergo Sum* **12:** 275–281.

Manjarrez J, Venegas-Barrera CS, GarcÍa-Guadarrama T. 2007. Ecology of the Mexican alpine blotched garter snake (*Thamnophis scalaris*). *Southwestern Naturalist* **52:** 258–262.

Manjarrez J, García CM, Drummond H. 2013. Variation in the diet of the Mexican black-bellied gartersnake *Thamnophis melanogaster*: importance of prey availability and snake body size. *Journal of Herpetology* **47:** 413–420.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* **66:** 526–538.

McVay JD, Carstens B. 2013. Testing monophyly without well-supported gene trees: evidence from multi-locus nuclear data conflicts with existing taxonomy in the snake tribe Thamnophiini. *Molecular Phylogenetics and Evolution* **68:** 425–431.

Minin V, Abdo Z, Joyce P, Sullivan J. 2003. Performance-based selection of likelihood models for phylogeny estimation. *Systematic Biology* **52:** 674–683.

Nee S, Mooers AO, Harvey PH. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 8322–8326.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Price RM. 1983. Microdermatoglyphics: the Liodytes-Regina problem. *Journal of Herpetology* **17:** 292–294.

Pyron RA, Burbrink FT. 2009. Body size as a primary determinant of ecomorphological diversification and the evolution of mimicry in the lampropeltinine snakes (Serpentes: Colubridae). *Journal of Evolutionary Biology* **22:** 2057–2067.

de Queiroz A, Lawson R. 1994. Phylogenetic relationships of the garter snakes based on DNA sequence and allozyme variation. *Biological Journal of the Linnean Society* **53:** 209–229.

de Queiroz A, Lawson R. 2008. A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake *Thamnophis validus*. *Biological Journal of the Linnean Society* **95:** 409–424.

de Queiroz A, Lawson R, Lemos-Espinal JA. 2002. Phylogenetic relationships of North American garter snakes (Thamnophis) based on four mitochondrial genes: how much DNA sequence is enough? *Molecular Phylogenetics and Evolution* **22:** 315–329.

Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, Alfaro ME. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* **4:** 1958.

Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, Huang H, Larson JG. 2014. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* **5:** 701–707.

Rage JC. 1988. The oldest known colubrid snakes. The state of the art. *Acta Zoologica Cracoviensia* **31:** 457–474.

Rambaut A, Drummond AJ. 2009. *Tracer, Version 1.5*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Rodríguez-Robles JA, Bell CJ, Greene HW. 1999. Gape size and evolution of diet in snakes: feeding ecology of erycine boas. *Journal of Zoology* **248:** 49–58.

Rossman DA. 1985. Liodytes resurrected, reexamined, and reinterred. *Journal of Herpetology* **19:** 169–171.

Rossman DA, Blaney RM. 1968. A new natricine snake of the genus *Adelophis* from western Mexico. *Occasional Papers of the Museum of Zoology Louisiana State University* **35:** 1–12.

Rossman DA, Burbrink FT. 2005. Species limits withing the Mexican garter snakes of the *Thamnophis godmani* complex. *Occasional Papers of the Museum of Natural Science Louisiana State University* **79:** 1–44.

Rossman DA, Ford NB, Seigel RA. 1996. *The garter snakes: evolution and ecology*. Norman, OK: University of Oklahoma Press.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. *Bioinformatics methods and protocols: methods in molecular biology*. Totowa, NJ: Humana Press, 365–386.

Ruane S, Bryson RW, Pyron RA, Burbrink FT. 2014. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. *Systematic Biology* **63:** 231–250.

Skinner A, Donnellan SC, Hutchinson MN, Hutchinson RG. 2005. A phylogenetic analysis of Pseudonaja (Hydrophiinae, Elapidae, Serpentes) based on mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **37:** 558–571.

Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0. Available at: http://www.repeatmasker.org

Swofford DL. 2003. *PAUP*. Phylogenetic analysis using parsimony (*and other methods), 4th edn*. Sunderland, MA: Sinauer Associates.

Taylor EH. 1942. Mexican snakes of the genera Adelophis and Storeria. *Herpetologica* **2:** 75–79.

Trapido H. 1944. The snakes of the genus *Storeria*. *American Midland Naturalist* **31:** 1–84.

Voorhies M. 1990. Vertebrate biostratigraphy of the Ogallala Group in Nebraska. In: Gustavson TC, ed. *Geologic Framework and Regional Hydrology*. Austin, Texas: University of Texas Bureau of Economic Geology, 115–151.

Wiens JJ, Kuczynski CA, Smith SA, Mulcahy DG, Sites JW Jr, Townsend TM, Reeder TW. 2008. Branch lengths, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. *Systematic Biology* **57:** 420–431.

Wood DA, Vandergast AG, Lemos Espinal JA, Fisher RN, Holycross AT. 2011. Refugial isolation and divergence in the Narrowheaded Gartersnake species complex (*Thamnophis rufipunctatus*) as revealed by multilocus DNA sequence data. *Molecular Ecology* **20:** 3856–3878.

Xie W, Lewis PO, Fan Y, Kuo L, Chen MH. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* **60:** 150–160.

Yang Z. 2006. *Computational molecular evolution*. Oxford: Oxford University Press.

APPENDIX

TABLE A1. MATERIAL EXAMINED

Scientific name	Specimen number	Country	State
Adelophis foxi	LSUMZ 40846	México	Durango
Clonophis kirtlandii	LSUMZ 39566	USA	Illinois
Natrix natrix	H05128	na	na
Nerodia clarkii	LSUMZ 43426	USA	Alabama
Nerodia cyclopion	JDM 1034	USA	Louisiana
Nerodia erythrogaster	JDM 1004	USA	Texas
Nerodia fasciata	LSUMZ 40040	USA	Alabama
Nerodia floridana	LSUMZ 40090	USA	Florida
Nerodia rhombifer	H21296	USA	Louisiana
Nerodia sipedon	LSUMZ 40906	USA	Georgia
Nerodia taxispilota	LSUMZ 40308	USA	Florida
Regina alleni	LSUMZ 40570	USA	Florida
Regina grahamii	LSUMZ 40330	USA	Louisiana
Regina rigida	LSUMZ 40503	USA	Louisiana
Regina septemvittata	LSUMZ 40101	USA	na
Seminatrix pygaea	LSUMZ 42686	USA	Georgia
Storeria dekayi	LSUMZ 39878	USA	Pennsylvania
Storeria storerioides A	JAC 23435	México	Jalisco
Storeria occipitomaculata	LSUMZ 80971	USA	Louisiana
Storeria storerioides B	LSUMZ 40790	México	México
Thamnophis atratus	LSUMZ 44386	USA	California
Thamnophis brachystoma	LSUMZ 58447	USA	Pennsylvania
Thamnophis butleri	LSUMZ 39656	Canada	Ontario
Thamnophis chrysocephalus[1]	HCD7310	México	na
Thamnophis couchii	H08146	na	na
Thamnophis cyrtopsis	LSUMZ 40426	USA	New Mexico
Thamnophis elegans	LSUMZ 39641	USA	New Mexico
Thamnophis eques	LSUMZ 40752	México	Durango
Thamnophis errans[2]	LSUMZ 16999	México	Durango
Thamnophis fulvus	LSUMZ 57127	Guatemala	Jalapa
Thamnophis lineri	JAC 21406	México	Oaxaca
Thamnophis bogerti	JAC 21416	México	Oaxaca
Thamnophis conanti	JAC 22810	México	Puebla
Thamnophis couchii	LSUMZ 37179	USA	California
Thamnophis marcianus	LSUMZ 48745	USA	Texas
Thamnophis melanogaster	LSUMZ 37429	México	Michoacán
Thamnophis nigronuchalis	LSUMZ 40849	México	Durango
Thamnophis ordinoides	LSUMZ 40130	Canada	British Columbia
Thamnophis proximus	LSUMZ 87348	USA	Louisiana
Thamnophis pulchrilatus	LSUMZ 35379	México	Durango
Thamnophis radix	H02935	USA	Wisconsin
Thamnophis rufipunctatus	LSUMZ 40853	México	Chihuahua
Thamnophis sauritus	LSUMZ 41508	USA	Florida
Thamnophis scalaris	LSUMZ 42639	México	México
Thamnophis scaliger	LSUMZ 42640	México	México
Thamnophis sirtalis	LSUMZ 41181	USA	Maine
Thamnophis sumichrasti[3]	LSUMZ 11114	México	Hidalgo
Thamnophis validus	JRM 4541	México	Sinaloa
Tropidoclonion lineatum	H13044	USA	na
Haldea striatula	LSUMZ 83481	USA	Louisiana
Virginia valeriae	LSUMZ 81173	USA	Louisiana

LSUMZ, Louisiana State University Museum of Zoology; H, LSUMZ tissue catalogue; JAC, Jonathan A. Campbell field catalogue; JDM, John D. McVay field catalogue; JRM, Joseph R. Mendelson III field catalogue. Where noted, the ND4 sequence data were taken from Genbank: [1]AF420098; [2]EF417363; [3]AF420200.

TABLE A2. OLDEST DISCOVERED FOSSILS OF NATRICINE SNAKES IN NORTH AMERICA

Genus	Species	Oldest period	Age range (Mya)
Neonatrix	*elongata*	Hemingfordian	20.6–16.3
Neonatrix	*intera*	Early Barstovian	16.3–13.6
Neonatrix	*magna*	Medial Barstovian	16.3–13.6
Nerodia		Medial Barstovian	16.3–13.6
Nerodia	*erythrogaster*	Irvingtonian I	1.9–0.9
Nerodia	*fasciata*	Irvingtonian I	1.9–0.9
Nerodia	*rhombifer*	Blancan V	
Nerodia	*floridana*	Irvingtonian I	1.9–0.9
Nerodia	*sipedon*	Blancan V	
Nerodia	*taxispilota*	Rancholabrean II	0.15–0.01
Nerodia	*hibbardi*	Blancan III	
Nerodia	*hillmani*	Clarendonian I	
Regina	sp.	Irvingtonian I	1.9–0.9
Regina	*alleni*	Irvingtonian I	1.9–0.9
Regina	*grahamii*	Blancan V	
Regina	*intermedia*	Irvingtonian I	1.9–0.9
Regina	*septemvittata*	Rancholabrean II	0.15–0.01
Storeria	sp.	Irvingtonian II	0.9–0.4
Storeria	cf. *dekayi*	Rancholabrean I	0.4–0.15
Storeria	*dekayi*	Rancholabrean II	0.15–0.01
Storeria	*occipitomaculata*	Rancholabrean II	0.15–0.01
Thamnophis		Medial Barstovian	16.3–13.6
Thamnophis	*brachystoma*	Rancholabrean II	0.15–0.01
Thamnophis	*couchii*	Rancholabrean II	0.15–0.01
Thamnophis	cf. *cyrtopsis*	Rancholabrean II	0.15–0.01
Thamnophis	*elegans*	Irvingtonian II	0.9–0.4
Thamnophis	*marcianus*	Blancan III	
Thamnophis	*proximus*	Blancan V	
Thamnophis	cf. *sirtalis*	Early Hemphillian	
Thamnophis	cf. *sauritus*	Blancan IV	
Thamnophis	*radix*	Blancan II	
Thamnophis	*sirtalis*	Blancan II	
Tropidoclonion	*lineatum*	Irvingtonian I	1.9–0.9
Virginia		Irvingtonian I	1.9–0.9
Haldea	*striatula*	Rancholabrean II	0.15–0.01
Virginia	*valeriae*	Rancholabrean II	0.15–0.01