Hidden diversity within the depauperate genera of the snake tribe Lampropeltini (Serpentes, Colubridae)

Hollis A. Dahn\textsuperscript{a,c}, Jason L. Strickland\textsuperscript{a,d}, Alejandra Osorio\textsuperscript{a,e}, Timothy J. Colston\textsuperscript{b,f}, Christopher L. Parkinson\textsuperscript{a,g,*}

\textsuperscript{a} Department of Biology, University of Central Florida, 4110 Libra Drive, Orlando, FL 32816, USA
\textsuperscript{b} Department of Biological Sciences, The George Washington University, 2029 G St NW, Room 302, Washington DC 20052, USA
\textsuperscript{c} Department of Ecology and Evolutionary Biology, University of Toronto, 25 Wilcocks St., Toronto, ON M5S 3B2, Canada
\textsuperscript{d} Department of Biological Sciences, Clemson University, 190 Collings St., Clemson, SC 29634, USA
\textsuperscript{e} College of Medicine, University of Florida, 1600 SW Archer Rd, Gainesville, FL 32610, USA
\textsuperscript{f} Department of Biological Sciences, Florida State University, 319 Stadium Drive, Tallahassee, FL 32306, USA
\textsuperscript{g} Department of Biological Sciences and Department of Forestry and Environmental Conservation, Clemson University, 190 Collings St., Clemson, SC 29634, USA
\textsuperscript{*} Corresponding author at: Dept. of Biological Sciences & Dept. of Forestry and Environmental Conservation, Clemson University, 190 Collings St., 157B Life Sciences Facility, Clemson, SC 29634, USA.

E-mail addresses: h.dahn@mail.utoronto.ca (H.A. Dahn), jason.strickland@knights.ucf.edu (J.L. Strickland), aosorio5@knights.ucf.edu (A. Osorio), tim@maddreptiles.com (T.J. Colston), viper@clemson.edu (C.L. Parkinson).

\textsuperscript{1} Current address.

https://doi.org/10.1016/j.ympev.2018.08.018

Received 6 October 2017; Received in revised form 31 July 2018; Accepted 29 August 2018

Available online 03 September 2018

1055-7903/ © 2018 Elsevier Inc. All rights reserved.

\textbf{1. Introduction}

Having a thorough understanding of lineage diversity within a given study system is a crucial first step in understanding the evolution of biodiversity. Drivers of biodiversity patterns cannot be accurately understood when lineage diversity is unknown or under-sampled, and inadequate understanding of biodiversity can negatively affect a diverse range of scientific and conservation interests (Wilson, 1988; Blackmore, 1996; Lawton et al., 1998; Daly et al., 2001; Gotelli, 2004; Hedtke et al., 2006; Linkem et al., 2010). There are many potential causes for unsampled diversity in evolutionary studies. For example, sampling may be purely informed by current taxonomy with yet unassessed groups or cryptic species, the included individuals may be geographically clustered and not address biogeographic barriers, or adequate numbers or
Among the nine genera of Lampropeltini, six are taxonomically depauperate (Arizona, Bogertophis, Cemophora, Pseudelaphe, Rhinocheilus, and Senticolis) while three are comparably speciose (Lampropeltis, Pantherophis, and Pitophis). The depauperate genera each contain 1–2 species compared to the 7–24 observed in Lampropeltis, Pantherophis, and Pitophis. Sampling for previous phylogenetic studies investigating higher-order relationships of Lampropeltini have been taxonomically informed, and as such, include 1–2 samples for the six depauperate genera and a broader representation for the more speciose genera (e.g., Burbrink and Lawson, 2007; Pyron and Burbrink, 2009a; Chen et al., 2017). One potential reason for poor sampling of the depauperate genera both phylogenetically and phylogeographically is the large proportion of the distribution of these genera in Mexico and Central America, covering areas particularly challenging to sample (Hansen and Salmon, 2017). Existing studies of other taxa with sampling in the region have found high levels of unrecognized diversity due to the area’s complex biogeographic history (Devitt, 2006; Bryson et al., 2011).

Recent taxonomic revisions in this group have further increased the total number of species and have been largely clustered within Lampropeltis and Pantherophis. Since 2000, the number of species of Lampropeltis has been recommended for increase from 9 to 24 (Bryson et al., 2005; Pyron and Burbrink, 2009b; Myers et al., 2013; Ruane et al., 2014; McKelvy and Burbrink, 2017). The clade that is now Pantherophis was increased from 3 species to 9 (Burbrink, 2001; Burbrink, 2002; Crother et al., 2011). Pitophis was increased from 4 to 7 (Crother, 2000; Wallach et al., 2014). In contrast, the remaining genera, all with two or fewer described species, received fewer evaluations. Recommended increases were from one species to two in Arizona (Grissmer, 2002), Cemophora (Weinell and Austin, 2017), Rhinocheilus (Grissmer, 1999) and Pseudelaphe (Flores-Villela and Canseco-Márquez, 2004). As many analyses of this tribe inform their sampling through this taxonomy, it begs the question of whether or not this dichotomy of speciose and depauperate genera is truly representative of existing lineage diversity.

Despite the long history of investigations estimating the relationships between genera of this tribe (e.g.: Cope, 1895; Dunn, 1928; Underwood, 1967; Keogh, 1996; Utiger et al., 2002; Burbrink and Lawson, 2007; Pyron and Burbrink, 2009a; Pyron et al., 2013; Chen et al., 2017), the placement of the depauperate genera has a particular lack of consensus. For example, the genus Arizona in recent iterations has been sister to Pseudelaphe (Utiger et al., 2002), Bogertophis (Burbrink and Lawson, 2007), Rhinocheilus (Pyron and Burbrink, 2009a), and Pseudelaphe again (Pyron et al. 2013). The genus Bogertophis has been sister to Lampropeltis (Utiger et al., 2002), Arizona (Burbrink and Lawson, 2007), Pseudelaphe (Pyron and Burbrink, 2009a), and the Cemophora-Lampropeltis clade (Pyron et al. 2013). Most recently, a next-generation approach by Chen et al. (2017) yielded a well-resolved phylogeny of the tribe using hundreds of markers. The relationships they recovered again differ from the next most recent phylogeny by Pyron et al. (2013). The study by Chen et al. (2017) represents a sorely needed use of an expanded set of loci to investigate evolutionary relationships within this group. However, they too did not extend sampling of the depauperate genera beyond one individual per most of the six depauperate genera (two for Bogertophis). Because the understudied depauperate genera may contain unknown lineage diversity, it is unclear whether or not this sampling strategy hindered attempts to assess speciation rates and higher-order relationships in this tribe.

In order to investigate the extent of undocumented lineage diversity within the depauperate genera and its effect on phylogenetic analyses, we assembled what is thus far the most representative sampling of the six depauperate genera within Lampropeltini. We assess genetic diversity via genetic distance measurements in each genus of Lampropeltini to determine if genetic diversity reflects the lineage diversity described by current taxonomy with the assumption that large
amounts of genetic diversity are generally correlated with lineage diversity. Additionally, we use concatenated phylogenetic analyses to investigate the substructure of each genus as further indication of any potential undocumented lineage diversity they contain. We use both coalescent and concatenated phylogenetic analyses as well as fossil-calibrated node dating to determine how this additional sampling may affect phylogenetic conclusions compared to previous studies as well as to view how this diversity is distributed geographically and temporally. With these data, we were also able to test the monophyly of two putative species, *Arizona pacata* and *Pseudelaphe phaeasces*. Lastly, we assess whether or not an ancient rapid radiation may be contributing to the group's topological instability with divergence dating analyses.

2. Materials and methods

2.1. Taxon sampling and DNA extraction

We consider Lampropeltini to include 51 putative species of the genera *Arizona* (2 sp.), *Bogortophis* (2 sp.), *Cemophora* (2 sp.), *Lampropeltis* (24 sp.), *Pantherophis* (9 sp.), *Pituophis* (7 sp.), *Rhinoaechus* (2 sp.), *Senticolis* (1 sp.: Ruane et al., 2014; Wallach et al., 2014; Weinnell and Austin, 2017; Myers et al., 2013; McKelvy and Burbrink, 2017). Our sampling for this investigation was comprised of tissues representing 173 individuals of 20 (17 Lampropeltini) species. To supplement this, additional Lampropeltini and outgroup sequences were downloaded from GenBank for a total of 260 terminals representing 67 (32 Lampropeltini) species (Table S1).

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.08.018.

In order to investigate the diversity of lampropeltines beyond previous studies, we sought to increase geographic and taxonomic sampling within the less-studied clades of Lampropeltini: *Arizona* (49 individuals, 2 of 2 species and 8 of 8 subspecies), *Bogortophis* (22 individuals, 2 of 2 described species and 2 of 3 subspecies), *Rhinoaechus* (46 individuals, 1 of 2 species and 3 of 4 subspecies), *Cemophora* (30 individuals, 2 of 2 species and 3 of 3 subspecies), *Pseudelaphe* (7 individuals, 2 of 2 species and 2 of 4 subspecies), and *Senticolis* (15 individuals, 2 of 3 subspecies). In addition to these individuals, our phylogenetic analyses included a large number of outgroup colubrid taxa (35 species of 14 genera) from successive sister groups based on previous molecular analyses (Table S1, Luo et al., 2010; Pyron et al., 2013; Wilberg, 2015). This outgroup strategy was chosen due to the previously incongruent phylogenetic estimates of higher-order relationships within Lampropeltini and its allied clades (Utiger et al., 2002; Burbrink and Lawson, 2007; Pyron and Burbrink, 2009a; Pyron et al., 2013).

Tissue samples we collected were acquired through fieldwork in the region with greatest lampropeltini diversity and abundance, the southwestern United States and Mexico. Individuals were euthanized, preserved, vouchered, and deposited in museum collections when permitted; or released at site of capture after processing (Table S1). Tissue samples were collected via scale clipping, tail doc, blood draw, or liver dissection and stored in 95% ethanol. Low-quality samples including dried skin sheds and salvaged muscle from dead-on-road specimens were incorporated where corresponding tissue was not available. Supplemental tissue loans were acquired from 13 natural history collections and 4 private collections. From all tissue samples gathered, whole genomic DNA was extracted using the Serapure bead extraction protocol of Rohland and Reich (2012) modified by Faircloth (2014).

2.2. Amplification, sequencing, and alignment

We amplified 4 mitochondrial (mtDNA) and 3 nuclear (nDNA) loci via PCR following standard conditions (Table 1). All primer sequences were obtained from previously published research (Table 1). Amplicons were purified using 0.05 μL Exonuclease 1 (Thermo Scientific model of Rohland and Reich (2012) modified with 0.05 μL Exonuclease 1 (Thermo Scientific) and 4 µL of the DNA sample. Following amplification, all amplicons were purified using the Serapead bead extraction protocol of Rohland and Reich (2012) modified by Faircloth (2014).

| Locus information and optimal model for all loci utilized in this study to infer the evolutionary history of the nine genera within Lampropeltini. For each 30 µL PCR reaction, there was an initial denaturation step of 94 °C for 3 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing (see below) for 30 s, extension at 72 °C for 60 s and then a final extension for 15 min at 72 °C. An * indicates loci downloaded from GenBank that were generated in previous studies.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Model in RAxML</th>
<th>Alignment length (bp)</th>
<th>Model in BEAST</th>
<th>Ancestral temp. (°C)</th>
<th>Reverse primer</th>
<th>Forward primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S-rRNA</td>
<td>125L1</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>GTRGAMMA 481 Kocher et al. (1989)</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>ND2</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1032 de Queiroz et al. 2002</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND3</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1117 Burbrink et al. (2000)</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>COI*</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 680 Townsend et al. (2008)</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND1*</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1010 Groth and Barrowclough (1999)</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND2*</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 626 GTRGAMMA 569</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND4*</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>12S</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 964 GTRGAMMA 964</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>COI</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND1</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>ND4</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>RAG-1</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>RAG-1</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>RAG-1</td>
<td>216L1</td>
<td>5′-GCTGCGCCTGTTTATCAAAAACAT-3′</td>
<td>5′-GCGGGTTTGGACAAACGGCC-3′</td>
<td>GTRGAMMA 1068 GTRGAMMA 1068</td>
<td>53</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 1
starting best model for downstream phylogenetic analyses. We generated the programs BEAST2 v 2.4.5 (Bouckaert et al., 2014) and RAxML v 8.2.8 (Stamatakis, 2014), respectively. For all BEAST2 analyses, .xml files were analyzed with LogCombiner v 2.4.5 (Bouckaert et al., 2014) and subsampled for a final file of 9000 tree states. We used the combined .trees file to generate the final phylogeny in FigTree v 1.4.3 (Rambaut, 2016). The same dataset was used to infer an ML tree in RAxML using the GTR-GAMMA model (-m GTRGAMMA) using the rapid bootstrap analysis and search (-f a) with the autoMRE flag to identify the extended majority rules tree quickly (-# autoMRE; Pattengale et al., 2009). For BI and ML, nodes ≥ 0.95 and ≥75% respectively, were considered to be significantly supported.

2.5. Fossil-calibrated tree

To estimate the age of the divergences within Lampropeltini, we used a fossil-calibrated analysis in BEAST2. The clock and tree models were linked among all loci and the site models were the same as those used in the BI concatenated phylogeny. We used a relaxed lognormal clock for the analysis and a Birth Death Model tree prior (Heath et al., 2014). The initial prior for the relative death rate was set to 0.5 and the birth diff rate was set to 1.0 and then both parameters were estimated and set to conditional on root (Heath et al., 2014). We calibrated our phylogeny using the same calibrations as Pyron and Burbrik (2009a) based on data in Holman (2000). All five calibration points were set to a log normal distribution and monophyly was enforced for each calibration. These points were modeled so that 95% of the prior weight fell within the estimated age of the fossils. We ran the analysis four independent times for 250 million generations sampling every 25 thousand generations. The four runs were evaluated in Tracer to ensure convergence, stationarity, and that all ESS values were ≥200. The first 25% of each run was discarded as burn-in and the remainder was combined into maximum clade credibility tree with the ages of the nodes displayed.

2.6. *BEAST coalescent tree

Our analysis to estimate a coalescent genus-level tree was done in BEAST2 using the software package StarBEAST2 v 0.13.2 (Ogilvie et al., 2017). The input file was created in BEAUti using the StarBEAST2 template. For this analysis, all mitochondrial loci were linked for both the clock and tree models because the mitochondrial DNA is inherited as a unit. The two nuclear loci Vim4 and Vim5 were also linked together for the clock and tree models because they are the introns of the same gene and are in the same linkage group. Site models were the same as in the concatenated and fossil-calibrated runs. All species from within a genus were grouped into a “Species/Population”, e.g. *Cemophora* grouped by two described species and *Senticolis* grouped by two sampled regions. The gene ploidy for the mitochondrial tree was set to 0.5 and all nuclear loci were set to 2.0. The population model used for the analysis was set to estimate constant population size with a starting prior of 1.0. The clock rate was estimated independently for all trees under the strict clock model. The species tree was run under a Yule model not conditional on root with an estimated birth diff rate and the initial prior of 180 for this parameter based on recommendations for StarBEAST2 (Ogilvie et al., 2017). Each analysis was run four independent times for 1 billion generations and sampled every 100 thousand generations. Based on the lnL values, the first 10% was discarded as burn-in. The log files were analyzed with Tracer to verify convergence and that all parameters reached an ESS ≥200. We displayed the phylogeny using FigTree and DensiTree v 2.2.5 (Bouckaert and Heled 2014).

3. Results

3.1. Locus diversity

The mean within-group uncorrected p-distance within each genus of Lampropeltini ranged between 0.011–0.085 for mitochondrial loci and 0.001–0.004 for nuclear loci (Table 2). Estimates using the Tamura-Nei model were similar to the uncorrected p-distance values when visualized the final phylogeny in FigTree v 1.4.3 (Rambaut, 2016).

In order to assess levels of genetic diversity within our sequence data at the generic level, we first grouped the sequences for each locus by genus within Lampropeltini. For loci that had at least two individuals per genus, we measured intragenic genetic distance using Tamura-Nei genetic distance (Tn; Tamura and Nei, 1993) and uncorrected p-distance in MEGA 7.0.14 (Kumar et al., 2016). For both, we left the default parameters and set the site coverage cutoff to 95%. The average mitochondrial genetic distance was calculated using ND4, 16S, cytB, ND2, and COI and the average nuclear genetic distance was calculated using NT3, RAG-1, and TATA. We excluded ND1, 12S, vim4, vim5, SPTBN1, and c-mos because in these loci some genera were not represented by at least two sequences.

We used PartitionFinder 2.1.1 (Lanfear et al., 2017) to determine the best-fit model of nucleotide evolution for each locus (Table 1). For all schemes, we used the “greedy” search algorithm with linked branch lengths and the Bayesian Information Criterion (BIC) to determine the best model for downstream phylogenetic analyses. We generated the starting tree in PhyML v 3.0 (Guindon et al., 2010). We performed model searches using each locus individually and all loci combined. For these, we tested models with protein coding genes partitioned by codon position and as one unit. All PartitionFinder2 analyses were run on the Stokes High Performance Computer (SHPC) through the University of Central Florida Advanced Research Computing Center (UCFARCC).

We reconstructed phylogenetic trees from our dataset using both Bayesian inference (BI) and Maximum Likelihood (ML) techniques in the programs BEAST2 v 2.4.5 (Bouckaert et al., 2014) and RAxML v 8.2.8 (Stamatakis, 2014), respectively. For all BEAST2 analyses, .xml files were created with BEAUti v 2.4.5 (Bouckaert et al., 2014) and site models were chosen based on the results from PartitionFinder2 (Lanfear et al., 2017). First, individual gene trees were estimated in BEAST2 to assess discordance among loci. Following this assessment, we performed four simultaneous, independent BEAST2 runs with Beagle 2.1.2 (Ayres et al., 2012) in a concatenated analysis for all loci. These were grouped by two samplings. The two nuclear loci Vim4 and Vim5 were also linked together for the clock and tree models because they are the introns of the same gene and are in the same linkage group. Site models were the same as in the concatenated and fossil-calibrated runs. All species from within a genus were grouped into a “Species/Population”, e.g. *Cemophora* grouped by two described species and *Senticolis* grouped by two sampled regions. The gene ploidy for the mitochondrial tree was set to 0.5 and all nuclear loci were set to 2.0. The population model used for the analysis was set to estimate constant population size with a starting prior of 1.0. The clock rate was estimated independently for all trees under the strict clock model. The species tree was run under a Yule model not conditional on root with an estimated birth diff rate and the initial prior of 180 for this parameter based on recommendations for StarBEAST2 (Ogilvie et al., 2017). Each analysis was run four independent times for 1 billion generations and sampled every 100 thousand generations. Based on the lnL values, the first 10% was discarded as burn-in. The log files were analyzed with Tracer to verify convergence and that all parameters reached an ESS ≥200. We displayed the phylogeny using FigTree and DensiTree v 2.2.5 (Bouckaert and Heled 2014).

We reconstructed phylogenetic trees from our dataset using both Bayesian inference (BI) and Maximum Likelihood (ML) techniques in the programs BEAST2 v 2.4.5 (Bouckaert et al., 2014) and RAxML v 8.2.8 (Stamatakis, 2014), respectively. For all BEAST2 analyses, .xml files were created with BEAUti v 2.4.5 (Bouckaert et al., 2014) and site models were chosen based on the results from PartitionFinder2 (Lanfear et al., 2017). First, individual gene trees were estimated in BEAST2 to assess discordance among loci. Following this assessment, we performed four simultaneous, independent BEAST2 runs with Beagle 2.1.2 (Ayres et al., 2012) in a concatenated analysis for all loci. These were grouped by two samplings. The two nuclear loci Vim4 and Vim5 were also linked together for the clock and tree models because they are the introns of the same gene and are in the same linkage group. Site models were the same as in the concatenated and fossil-calibrated runs. All species from within a genus were grouped into a “Species/Population”, e.g. *Cemophora* grouped by two described species and *Senticolis* grouped by two sampled regions. The gene ploidy for the mitochondrial tree was set to 0.5 and all nuclear loci were set to 2.0. The population model used for the analysis was set to estimate constant population size with a starting prior of 1.0. The clock rate was estimated independently for all trees under the strict clock model. The species tree was run under a Yule model not conditional on root with an estimated birth diff rate and the initial prior of 180 for this parameter based on recommendations for StarBEAST2 (Ogilvie et al., 2017). Each analysis was run four independent times for 1 billion generations and sampled every 100 thousand generations. Based on the lnL values, the first 10% was discarded as burn-in. The log files were analyzed with Tracer to verify convergence and that all parameters reached an ESS ≥200. We displayed the phylogeny using FigTree and DensiTree v 2.2.5 (Bouckaert and Heled 2014).

3. Results

3.1. Locus diversity

The mean within-group uncorrected p-distance within each genus of Lampropeltini ranged between 0.011–0.085 for mitochondrial loci and 0.001–0.004 for nuclear loci (Table 2). Estimates using the Tamura-Nei model were similar to the uncorrected p-distance values when...
chimeric taxa are one individual of the same species and geographic area (Table S1). This mitochondrial trees to determine the in
genera within Lampropeltini were particularly volatile. Additionally, taxa when sequences di
concatenated dataset. Some sequences were combined into chimeric (57 bp) that
regions leaves the possibility of clinal variation (Fig. 2).

3.2. Concatenated phylogeny

The final molecular data matrix consisted of 260 terminals and 14 loci with 52% coverage. Excluding a short region of 16S (57 bp) that could not be aligned, a total of 11,696 base pairs were included in our concatenated dataset. Some sequences were combined into chimeric taxa when sequences of different loci were available from different individuals of the same species and geographic area (Table S1). This was done mostly in non-focal genera when few sequences were available from the same individuals. Within the depauperate genera, the chimeric taxa are one A. elegans, one B. rosaliae, two B. subocularis, one P. flavirufa, and two S. triaspis.

Of the sites incorporated in the molecular data matrix, 8035 (69%) were found to be invariant within Lampropeltini. PartitionFinder2 identified the best-fitting partitioning scheme for each locus (Table 2). Protein-coding loci were not partitioned by codon position in favor of site models that account for rate heterogeneity among sites via gamma distributions. In all analyses we linked the site models in the following configuration: 12S, 16S, and ND4rrNA; c-mos and RAG-1; cytb, ND1, and ND4; SPTBNI, vim4, vim5, and TATA.

We examined the gene trees as well as the concatenated mitochondrial trees to determine the influence of each on the overall topology. Gene trees generated by our analyses indicated significant phylogenetic discordance among loci. Some of these topologies resembled previously published phylogenetic estimates of the group. Among the different gene trees, the positions of the six depauperate genera within Lampropeltini were particularly volatile. Additionally, Senticolis in some gene trees is recovered outside of Lampropeltini, such as in the gene tree of cytb where Senticolis is sister to a Lampropeltini-Elaphe-Coronella-Orthriophis-Zamenis clade (Figs. S3–16).

It is clear when comparing our nuclear and mitochondrial trees that there is greater discordance among the nuclear loci. The mitochondrial topology is most similar to the full concatenated tree, thus, contributing the strongest signal. The nuclear gene trees had low posterior probability and the genera were not always monophyletic. Additionally, the positioning of Arizona and Rhinocelus as sister taxa in a clade sister to Lampropeltis and Cemophora and positioning of Pseudelaphe and Senticolis were inconsistent (Figs. S3–16). This is unlikely to be caused by missing data, as those individuals with only mitochondrial data were few in number and interspersed with those that did. Specifically, there are 30 total lacking nuclear data in this set: two A. elegans, two A. pa-
cata, nine B. subocularis, three outgroup individuals, two P. flavirufa, three R. lecontei, and five S. triaspis (Table S1).

Despite observed gene tree discordance, each of the four independent runs of our Bayesian concatenated phylogenetic analysis in BEAST concluded with an ESS value greater than 4000, indicating that stationarity was reached within the 250 million generations allotted. These runs all inferred similar topologies to the combined run. The BI tree (Fig. 1; Fig. S1) generated a similar topology to our maximum likelihood analysis in RAxML (Fig. S2) from the same data matrix. At nodes where the topology differs, both analyses had relatively low support values. Nodes at which both trees agree are those with high support values. The monophyly of Lampropeltini, as well as that of all genera within was strongly supported by both our Bayesian inference (Fig. 1; Fig. S1) and maximum likelihood analyses (Fig. S2). Some higher-order relationships indicated in this tree are not consistent with those previously published. We recovered a sister relationship between Bogertophs and Pseudelaphe and notably, Rhinocelus and Arizona were not inferred as sister genera as in (Pyron et al., 2013). Aside from those nodes that are not well-resolved and the recovery of L. calligaster as sister to Cemophora, this topology was also recently recovered by Chen et al. (2017) using next-generation data for fewer individuals.

Increased sampling, particularly within Arizona, Rhinocelus, and Senticolis, also revealed previously unexamined substructure within these genera. Arizona and Rhinocelus both appear to possess eastern and western clades with boundaries geographically centered on the Cochise Filter Barrier, a common phylogeographic break (Zink and Blackwell, 1998; Riddle et al., 2000; Castoe et al., 2007). Senticolis also exhibits a significant level of sequence divergence between two major clades, but the geographic separation of the two associated sampling regions leaves the possibility of clinal variation (Fig. 2).

3.3. Fossil-calibrated

Our four runs converged and all ESS values were > 200. We found that all genera arose at approximately the same time period in the mid Miocene (Fig. 3). These divergence dates are similar to previous studies in the group (Pyron and Burbrink, 2009a; Chen et al., 2017). The timing of speciation events differed within the genera and corresponds to the level of genetic diversity found. Rhinocelus and Cemophora had the most recent accumulation of divergent lineages at ~ 4 MYA during the Pliocene whereas the two species within Bogertophis and the most distantly related species within Lampropeltis diverged ~ 12 MYA in the late Miocene.

3.4. Species coalescence

This study’s wider sampling within Senticolis, Pseudelaphe, Arizona, and Rhinocelus allowed us to test their intergeneric relationships in a coalescent framework. Our StarBEAST2 analysis recovered a topology consistent with the current hypotheses of the relationships among genera within Lampropeltini (Fig. 4). Using the DensiTree visualization of all trees sampled, there is clear evidence for incomplete lineage sorting and/or reticulate events among the genera (Fig. 4). This resulted in the backbone of the topology to not be well supported and the trees sampled show competing topologies based on the loci used and has been well documented in Lampropeltis (Burbrink and Gehara, 2018). For example, there were many trees that supported an Arizona + Rhinocelus sister relationship as seen in our concatenated phylogeny (Fig. 1).
Fig. 1. Concatenated phylogeny of the nine genera within Lampropeltini generated in BEAST. Nodes with strong support (≥0.95 posterior probability) are denoted with a black dot. Node values with ≤0.50 posterior probability are omitted. Abbreviations after taxon names refer to location as follows: AZ – Arizona, USA; BS – Baja California Sur, Mexico; CA – California, USA; CM – Campeche, Mexico; CR – Costa Rica; GT – Guatemala; HN – Honduras; KS – Kansas, USA; NM – New Mexico, USA; OA – Oaxaca, Mexico; QR – Quintana Roo, Mexico; SO – Sonora, Mexico; TM – Tamaulipas, Mexico; TX – Texas, USA; C. MX – central Mexico, S. MX – southern Mexico.
4. Discussion

4.1. Diversity within Lampropeltini

We included the best representation of the depauperate genera within Lampropeltini to date and documented previously unassessed diversity within these genera. *Senticolis*, *Pseudelaphe*, *Rhinocheilus*, and *Arizona* all had much more genetic diversity than expected when compared to the more speciose *Pituophis*, *Pantherophis*, and *Lampropeltis* (Table 2; Fig. 2). The apparent mismatch of expected diversity within these genera compared to number of species diagnosed may reflect disparity in the level of scrutiny to each genus, the sampling effort, or in the taxonomic dogma applied therein.

Overall, our analyses highlight the historical difficulties of determining the relationships among genera within Lampropeltini. Though our increased sampling and loci coverage resolved most nodes of interest with strong support, our concatenated, fossil-calibrated, and species trees each have different topologies and poor support at select problematic nodes. Additionally, the concatenated phylogeny was strongly influenced by the mitochondrial loci as in previous studies comparing nuclear and mitochondrial discordance (Wiens et al., 2010; Fontenot et al., 2011; Toews and Brelsford, 2012). The mitochondrial loci in our concatenated analysis heavily influenced unusual relationships in particular nodes. For example, *Lampropeltis calligaster* fell outside of *Lampropeltis* (Fig. 1). This is due to competing topology in the mtDNA. The loci cytb, ND1, and COI all suggested a monophyletic *Lampropeltis*. However, ND4, which is the most complete dataset in our analysis, and ND2 drove the non-monophyletic relationships. In some studies, the node placing *L. calligaster* in the most basal position to the remaining *Lampropeltis* appears with low support likely due to the competing signals (Pyron et al., 2011; Pyron et al., 2013), but the overwhelming evidence is that *Lampropeltis* is monophyletic (Pyron and Burbrink, 2009a; Pyron et al., 2011; Ruane et al., 2014). However, by increasing taxon sampling, adding additional molecular sequences, and using the available data on Genbank in a coalescent framework, we recovered a topology in agreement with that of Chen et al. (2017) who used many more characters (Fig. 4).

Chen et al. (2017) found that the diversity within Lampropeltini accumulated at the ancestral node that includes *Pituophis*, *Pantherophis*, *Arizona*, *Rhinocheilus*, *Cemophora*, and *Lampropeltis*, but not *Bogertophis*. 

Fig. 2. Sample distributions compared to phylogeny of the nine genera within Lampropeltini with historically poor sampling. Base maps are colored based on elevation (lighter color = higher elevation) as well as hill shading. Species ranges are sourced from IUCN with modifications based on current literature. Major clades within each genus are assigned a separate color of sample point. Phylogenies were adapted from Fig. 1.
Pseudelaphe, and Senticolis. If the change in diversification rate coincided with the invasion of colubrids into the New World, it would appear at the base of Lampropeltini. The authors discuss the possibility that unsampled lineages could be contributing to their conclusion. Given the diversity that we found in Pseudelaphe and Senticolis (discussed below), we agree that taxon sampling will need to be addressed to determine if it was the invasion of the New World when diversification started or sometime after (Chen et al., 2017). As Chen et al. (2017) indicated, if their analysis had more samples included of Senticolis and Pseudelaphe the conclusion might be more in line with other snake groups that similarly invaded the New World and rapidly diversified (Wüster et al., 2008; Brandley et al., 2011; Guo et al., 2012).

We were able to sample both species and all eight subspecies within Arizona, both species of Pseudelaphe, and included more Bogertophis and Senticolis samples than any molecular study to date. Of the putative diversity not represented in our phylogenetic analysis, most represent rare species/subspecies isolated to small distributions. This includes Rhinocheilus etheridgei which is only known from Isla Cerralvo and no tissue samples are currently available (Grismer, 1990). We were unable to include Bogertophis subocularis amplinotus nor two subspecies of Pseudelaphe flavirufa which are only known from their two type localities in Chiapas, Mexico (Dowling, 1960b). Finally, we are missing Senticolis triaspis triaspis which occurs in the Yucatan Peninsula (Dowling, 1960). Including this subspecies might further increase the amount of diversity within Senticolis.

4.2. Depauperate genera in space and time

In addition to the assessment of undocumented genetic diversity within the depauperate genera of Lampropeltini, we also assessed the distribution of that diversity geographically and temporally. For most of the depauperate genera sampled, the largest divisions observed in the substructure of the genera were spatially associated with known biogeographic barriers or significant distance between disjoint sections of species’ ranges. Temporally, our recovered node dates were slightly older than those recovered by Chen et al. (2017), likely due to differences associated with the size of the two datasets (Heath et al., 2014).

Our confidence intervals overlapped those of Chen et al. (2017) and suggest that the higher number of loci yielded a more specific time of divergence. We found that generic-level diversity within the group arose from the late-Oligocene to mid-Miocene and most species-level diversity and major substructure within the genera arose in the late-Miocene and early-Pliocene (Fig. 3). This further corroborates the long-established hypothesis put forth by other investigations and lines of evidence—including fossil evidence (reviewed in Holman, 2000) and other molecular analyses (Pyron et al., 2011; Pyron et al., 2013; Chen et al., 2017)—that colubrids underwent a rapid radiation in North America during the Miocene (Underwood, 1967; Williams et al., 1967; Keogh, 1996). Incomplete lineage sorting and reticulation has further complicated the evolutionary history of at least Lampropeltis and likely all of Lampropeltini (Burbrink and Gehara, 2018). The addition of our findings adds new insight into the diversity of the depauperate genera and paint a more detailed picture of the tribe’s biogeographic history during and after the colubrid radiation.

The evolutionary history of two particular genera, Arizona and Rhinocheilus, in this tribe are notable because of the close co-distribution of their species (Fig. 2). The divergences associated with two major clades of Arizona elegans and two major clades of Rhinocheilus lecontei correspond to the Cochise filter barrier, a proposed barrier to gene flow in many other desert taxa (Liebherr, 1986; Castoe et al., 2007; Pyron and Burbrink, 2009b). The dates associated with these diversifications in our analysis were very different from each other (9.78 MYA for Arizona, and 3.97 MYA for Rhinocheilus), and older than expected. Other east-west divergences at this barrier are often associated with the Pleistocene glacial cycles (Pyron and Burbrink, 2010). With a single-locus dataset, Myers et al. (2017) explored the potential contributing factors to this asynchronous diversification across the Cochise filter barrier in these two species and others. They recovered generally much younger divergence dates than either our analysis or Chen et al. (2017), but concluded that Rhinocheilus did diverge more recently than Arizona. Unlike previous investigations, our approach combined increased sampling of individuals in these genera with increased numbers of loci, also yielding a different topology for substructure within each genus.
We find the putative species *Arizona pacata* sister to the major western clade of *Arizona elegans* while maintaining reciprocal monophyly with strong support values (Fig. 1). Further tests for gene flow and population structure between our other recovered lineages in *Arizona* and *Rhinocheilus* are needed to determine if these clades warrant taxonomic revision.

The phylogeography of the genus *Cemophora* was recently reviewed by Weinell and Austin (2017), wherein they recommend the elevation of the subspecies *C. coccinea lineri* to full species status. Our results indicated that *Cemophora* contains two major clades, one with individuals in the disjoint Texas Coastal Bend area of the species range (*C. lineri*), and the other clade containing the remaining individuals spread throughout the southeastern U.S. with little differentiation. Thus, our findings were largely consistent with the recovered phylogeny of Weinell and Austin (2017) and support the elevation of *C. lineri* to full species. The split between *C. lineri* and *C. coccinea* is relatively recent at 4.1 MYA (Fig. 3), but is older than several divergences between proposed species pairs in *Lampropeltis* (Ruane et al., 2014).

Our increased sampling of *Pseudelaphe* compared to previous studies reveals substructure within *P. flavirufa* (Fig. 1). We recovered substantial genetic differentiation between the different areas sampled (Fig. 2). The putative species *P. phaescens* of the northeastern Yucatan (identified by geography) was found to be basal to *P. flavirufa*. As the location of a boundary or intergrade zone between these taxa has not been investigated, and the sampling of this genus is not as representative as would be ideal. Clinal variation between the sampling localities of the two species is a possibility, as the habitat and climate of the dry northeastern Yucatan is markedly different than the humid south, and this is reflected in the distribution of other reptile species (Lee, 1996). Dowling (1952b) first established *P. flavirufa phaescens* based on differentiation in pattern and scalé count. Liner (1994) reported the subspecies as elevated to full species without further explanation available at the time of this writing, and no additional assessment has been made since the initial description of *P. phaescens*.

Two major clades within *Senticolis* correspond geographically with the Transvolcanic Belt of Mexico, though there is considerable distance in the central part of the *Senticolis triquis* range where we could not sample. The pattern observed is similar to that seen in other taxa with similar distributions (Card et al., 2016). Further study will be necessary to elucidate the contributors to the phylogeographic substructure we observed within *Senticolis*, whether it be a geographic barrier reducing gene flow, clinal variation via isolation-by-distance, or simply a longer span of time since *Senticolis* became established compared to the other genera of Lampropeltini (Endler, 1977). Increased geographic sampling will be particularly important for this group if all extant lineages are to be represented. If the pattern of diversification within the Yucatan peninsula demonstrated by *Pseudelaphe* holds in *Senticolis*, then we would expect to find a distinct lineage in the northern area of the peninsula. This would further increase observed diversity within *Senticolis*. This pattern of divergence has appeared in other taxa as well. A known environmental transition within the Yucatan peninsula gives it a high degree of endemism (Köhler, 1996; Vázquez-Domínguez and Arita, 2010; Morales-Mávil et al., 2016).

For many of the depauperate genera we sampled here—namely *Arizona*, *Rhinocheilus*, *Pseudelaphe* and *Senticolis*—it is clear that more fine-scaled phylogeographic and species delimitation investigations are warranted and will be a beneficial next step in understanding the diversity, biogeography, and evolutionary history of Lampropeltinines. A more thorough understanding of lineage diversity within Lampropeltinines will also allow for reassessment of previous hypotheses tested under a taxonomic sampling scheme. We find that a purely taxonomic sampling scheme of Lampropeltini would not be representative of existing lineage diversity within the tribe and would likely introduce a sampling bias towards *Lampropeltis*, *Pantherophis*, and *Pituophis*. 5. Conclusions

The snake tribe Lampropeltini is one of the most heavily studied snake clades in the world and, over the last 30 years, the number of potential species in this tribe has increased from 23 to 51 with the majority of these being in *Lampropeltis* (Pyron and Burbrick, 2009b; Myers et al., 2013; Ruane et al., 2014). Additionally, many papers have been published that have explored historical biogeography (Rodriguez-Robles and De Jesus-Escobar, 1999, Burbrick and Lawson, 2007, Hansen and Salmon, 2017), morphological evolution (Keogh, 1996; Manier, 2004) and character evolution such as mimicry (Davis Rabosky et al., 2016), oviparity (Pyron and Burbrick, 2014), and diet (Rodriguez-Robles and De Jesus-Escobar, 1999). However, these were all tested under the assumption that *Arizona, Bogertophis, Cemophora, Pseudelaphe, Rhinocheilus*, and *Senticolis* were depauperate and contained little genetic diversity. Our study clearly illustrates the importance of representing putative diversity in phylogenies so the information can be used to fully inform other hypothesis. Given the disparity in the number of species per genus, this group would be a model system for testing species delimitation techniques. The genera cover the spectrum from highly morphologically/ecologically divergent without genetic divergence to highly genetically divergent but little morphological/ecological divergence (Barley et al., 2013). Within the tribe Lampropeltini both ancient adaptive radiation and under-sampled diversity may combine to create a uniquely challenging situation for phylogenetic techniques, requiring further analysis before major conclusions can be drawn that rely on robust taxonomy and phylogeny.

Acknowledgments

For providing tissues, we gratefully acknowledge D. Kizirian (American Museum of Natural History), J. Vindum (California Academy of Sciences), R. Brumfield and D. Dittmann (Louisiana State University Museum of Natural History), J. McGuire and S. Bouzid (Museum of Vertebrate Zoology), A. Wynn (Smithsonian National Museum of Natural History), N. Camacho (Natural History Museum of Los Angeles), J. Braun and B. Coyner (Sam Noble Oklahoma Museum of Natural History), B. Hollingsworth (San Diego State University Museum of Biodiversity), T. Taggart and C. Schmidt (Sternberg Museum of Natural History), T. Hibbitts (Texas A&M Biodiversity Research and Teaching Collections), C. Franklin (University of Texas at Arlington Amphibian and Reptile Diversity Research Center), D. Cannatella and T. LaDuc (University of Texas Natural History Collections), G. Watkins-Colwell (Yale Peabody Museum of Natural History), J. Fauth, O. Flores-Villela, and P. Lindsay. For providing photographs used in Fig. 1 and the graphical abstract we thank Matthew Cage (Pseudelaphe), David Deem (Senticolis), Travis Fisher (Arizona, Pituophis, Rhinocheilus), and Erik McCormick (Bogertophis, Cemophora). Support and assistance in the field was provided by J. Aceves, J. Adams, B. Bloom, T. Burkhardt, K. Carman, J. Castaneda, B. Chambers, C. Constantinos, T. Dimler, B. Eaton, R. Engelsdorf, M. Feldner, T. Fisher, R. Govreau, J. Houck, T. Jones, M. Linsalata, J. Lock, C. Mallery, C. May, S. May, R. Mayerhofer, E. McCormick, C. McMartin, J. McNally, D. Ortiz, A. Osmanski, A. Owens, T. Petty, M. Price, G. Salmon, J. Slone, D. Speckin, S. Stevens, K. Swanson, R. Swanson, B. Townsend, C. Trumbower, K. VanSooy, C. Vratil, D. Weber, W. Wüster and participants of SnakeDays in Sannder, Texas. We thank D. Greenspan, N. Lucas, J. Schnaitter, and P. Wiegand (UCF ARCC) for assistance with Stokes HPC. Comments and suggestions were given by A. Mason and E. Hofmann which greatly improved the quality of this manuscript. Financial support was generously provided by the Chicago Herpetological Society (HD), Explorer Club Youth Activity Fund (HD), Operation Wallacea (TJC), Prairie Biotic Research Inc. (JS), Sigma Xi Grants-in-aid-of-research (HD and JS), SnakeDays Research Grant (JS), the Southwestern Association of Naturalists (HD and JS), the UCF Academic Advancement Programs Research and Mentoring Program (RAMP; AO and HD), the UCF OUR.


