Received: 14 August 2018 | Revised: 4 November 2018 | Accepted: 5 November 2018

DOI: 10.1111/1755-0998.12970

RESOURCE ARTICLE

WILEY MOLECULAR ECOLOGY RESOURCES

Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: *Crotalus*)

Christopher Blair[1,2] | Robert W. Bryson Jr[3,4] | Charles W. Linkem[3] | David Lazcano[5] | John Klicka[3] | John E. McCormack[4]

[1]Department of Biological Sciences, New York City College of Technology, The City University of New York, Brooklyn, New York

[2]Biology PhD Program, CUNY Graduate Center, New York, New York

[3]Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, Washington

[4]Moore Laboratory of Zoology, Occidental College, Los Angeles, California

[5]Laboratorio de Herpetología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Nuevo León, México

Correspondence
Christopher Blair, Department of Biological Sciences, New York City College of Technology, The City University of New York, Brooklyn, NY.
Email: cblair@citytech.cuny.edu

Funding information
Division of Environmental Biology, Grant/Award Number: 1257785, 1258205

Abstract

With the continued adoption of genome-scale data in evolutionary biology comes the challenge of adequately harnessing the information to make accurate phylogenetic inferences. Coalescent-based methods of species tree inference have become common, and concatenation has been shown in simulation to perform well, particularly when levels of incomplete lineage sorting are low. However, simulation conditions are often overly simplistic, leaving empiricists with uncertainty regarding analytical tools. We use a large ultraconserved element data set (>3,000 loci) from rattlesnakes of the *Crotalus triseriatus* group to delimit lineages and estimate species trees using concatenation and several coalescent-based methods. Unpartitioned and partitioned maximum likelihood and Bayesian analysis of the concatenated matrix yield a topology identical to coalescent analysis of a subset of the data in BPP. ASTRAL analysis on a subset of the more variable loci also results in a tree consistent with concatenation and BPP, whereas the SVDQUARTETS phylogeny differs at additional nodes. The size of the concatenated matrix has a strong effect on species tree inference using SVDQUARTETS, warranting additional investigation on optimal data characteristics for this method. Species delimitation analyses suggest up to 16 unique lineages may be present within the *C. triseriatus* group, with divergences occurring during the Neogene and Quaternary. Network analyses suggest hybridization within the group is relatively rare. Altogether, our results reaffirm the Mexican highlands as a biodiversity hotspot and suggest that coalescent-based species tree inference on data subsets can provide a strongly supported species tree consistent with concatenation of all loci with a large amount of missing data.

KEYWORDS
Bayesian, coalescence, genomics, rattlesnakes, speciation, systematics

1 | INTRODUCTION

Phylogenetics and phylogeography continue to experience a progressive transformation with the evolving power of next-generation DNA sequencing (NGS) technologies and sophisticated new analytical models (Edwards, 2009, Edwards et al., 2016; Liu, Yu, Kubatko, Pearl, & Edwards, 2009; McCormack, Hird, Zellmer, Carstens, & Brumfield, 2013; McCormack & Faircloth, 2013). Although phylogenetic studies have been conducted with whole genomes (Jarvis et al., 2014), reduced-representation approaches like restriction enzyme-based methods (e.g., RADseq, ddRADseq, GBS) and DNA sequence capture have emerged as particularly flexible methods for interrogating large eukaryotic genomes for hundreds to thousands of orthologous loci suitable for phylogenetics (Baird et al., 2008; Elshire et al., 2011; Faircloth et al., 2012; Lemmon, Emme, & Lemmon, 2012; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). The large number of loci provided by these methods can help resolve difficult nodes in species trees, particularly those involved in rapid evolutionary radiations (Giarla & Esselstyn, 2015; Leaché et al., 2015; Leaché, Banbury, Linkem, & Oca, 2016; Meiklejohn, Faircloth, Glenn, Kimball, & Braun, 2016; Wagner et al., 2013). A specific class of molecular marker used with sequence capture, ultraconserved elements (UCEs; Bejerano et al., 2004; Faircloth et al., 2012), has helped resolve long-standing evolutionary questions across taxonomic groups at multiple spatial and temporal scales (Crawford et al., 2012; Faircloth, Sorenson, Santini, & Alfaro, 2013; McCormack et al., 2012; McCormack, Harvey, Faircloth, Crawford, Glenn, & Brumfield, 2013; Leaché et al., 2016; Smith, Harvey, Faircloth, Glenn, & Brumfield, 2013; Streicher, Schulte, & Wiens, 2016). In UCEs, the conserved core regions facilitate the collection of orthologous loci among diverse organisms, whereas the flanking regions provide the sequence variability necessary for phylogenetics (Faircloth et al., 2012).

A major challenge of working with UCE and other types of large data sets is deciding on the best approaches to analyse these large, often heterogeneous data sets. Although traditional concatenation is relatively straightforward and computationally feasible, it can be statistically inconsistent in the presence of incomplete lineage sorting (Chou et al., 2015; Kubatko & Degnan, 2007; Mirarab, Bayzid, & Warnow, 2016; Roch & Steel, 2014). Meanwhile, fully probabilistic multispecies coalescent methods such as *STARBEAST2* (Ogilvie, Bouckaert, & Drummond, 2017) are attractive as they model stochastic lineage sorting and simultaneously estimate gene trees, the species tree, divergence times and substitution rates. However, likelihood-based multispecies coalescent methods are not easily scalable to large NGS data sets. To address this, gene tree reconciliation methods estimate species trees relatively quickly and directly from gene trees, instead of sequence data (Liu et al., 2009; Liu, Yu, & Edwards, 2010; Vachaspati & Warnow, 2015; Zhang, Sayyari, & Mirarab, 2017). However, poorly resolved gene trees can have a large impact on the species tree (Meiklejohn et al., 2016; Xi, Liu, & Davis, 2015; Xu & Yang, 2016). A relatively new coalescent method, *SVDQUARTETS*, bypasses the estimation of gene trees and uses singular value decomposition scores to estimate relationships among quartets prior to assembling quartets into a full species tree containing all taxa (Chifman & Kubatko, 2014, 2015). The method appears statistically consistent under the multispecies coalescent, is scalable to NGS data sets and performs well under simulations (Chou et al., 2015), but it has not been well tested with empirical data sets.

Rattlesnakes (*Crotalus* and *Sistrurus*) represent a radiation of around 45 pitvipers distributed throughout the New World from Canada to Argentina (Blair & Sánchez-Ramírez, 2016; Campbell & Lamar, 2004; Gloyd, 1940; Klauber, 1956). There continues to be broad interest in elucidating the evolutionary history of these enigmatic snakes due, in part, to their medicinal importance and contributions to yearly envenomations (Campbell & Lamar, 2004). Despite extensive effort over many decades, however, relationships among key groups remain unresolved. Early studies using morphology and protein data conflict with those using mtDNA sequences, which in turn often conflict with phylogenies generated from nuclear genes (Blair & Sánchez-Ramírez, 2016; Brattstrom, 1964; Bryson, Murphy, Lathrop, & Lazcano-Villareal, 2011; Foote & MacMahon, 1977; Gloyd, 1940; Klauber, 1956; Murphy, Fu, Lathrop, Feltham, & Kovac,2002; Reyes-Velasco, Meik, Smith, & Castoe, 2013). Part of the discrepancy seems to result from the fact that many molecular studies rely solely on mtDNA (Bryson et al., 2011; Castoe & Parkinson, 2006; Murphy et al.,2002; Parkinson, 1999; Parkinson, Campbell, Chippindale, & Schuett, 2002), which often conflict with phylogenies from other sources of data (Grummer, Bryson, & Reeder, 2014; Leaché & McGuire, 2006). An additional difficulty in resolving the rattlesnake tree is that many groups rapidly diversified during the late Miocene during periods of global climate change (Zachos, Pagani, Sloan, Thomas, & Billups, 2001) and mountain formation in Mexico (Blair & Sánchez-Ramírez, 2016; Bryson et al., 2014, 2011). Rapid radiations increase the likelihood of gene tree species tree conflicts, particularly in species with large effective population sizes (Maddison, 1997). Thus, genome-level data sets coupled with multispecies coalescent methods are often needed to provide both additional informative variation and the necessary analytical tools to resolve difficult nodes.

Rattlesnakes of the *Crotalus triseriatus* species group inhabit mountainous regions in Mexico and the south-western United States (Campbell & Lamar, 2004). Previous mtDNA work on the group indicates multiple instances of paraphyly and the presence of cryptic species (Bryson et al., 2011). A subsequent study based on multilocus and morphological data described two new species and elevated one subspecies to species status (Bryson et al., 2014). Nine species are now recognized in the group, although phylogenetic relationships among species remain uncertain, and it is likely that additional species-level diversity awaits discovery. Previous mtDNA evidence suggests that the *C. triseriatus* group rapidly diversified during formation of the Trans-Mexican Volcanic Belt (Bryson et al., 2011), and UCE data and multispecies coalescent models may provide additional insight to better resolve the divergence history of the clade.

The first objective of our study is to use a large UCE data set to estimate phylogenetic relationships and species limits within the *C. triseriatus* group. Our second goal is to estimate divergence times

and compare results to previous rattlesnake studies based on fewer loci (Blair & Sánchez-Ramírez, 2016; Bryson et al., 2011). Finally, we compare traditional concatenation and coalescent methods, including the newer SVDQUARTETS, using a large empirical data set. We hope that these comparisons facilitate additional discussion on best practices for extracting meaningful phylogenetic information from genomic data sets.

2 | MATERIALS AND METHODS

2.1 | Sampling and library preparation

We sampled 54 rattlesnakes of the *C. triseriatus* group from throughout the Mexican highlands (Figure 1; Supplementary Table S1). We included multiple exemplars of all currently described taxa to assess both species limits and phylogenetic relationships. Taxa in the group include *C. aquilus*, *C. armstrongi*, *C. campbelli*, *C. lepidus* (*C. l. lepidus*, *C. l. klauberi* and *C. l. maculosus*), *C. morulus*, *C. pusillus*, *C. ravus* (*C. r. ravus*, *C. r. brunneus* and *C. r. exiguus*), *C. tlaloci* and *C. triseriatus* (Bryson et al., 2014, 2011). We included *Agkistrodon piscivorus* as an outgroup taxon to root tree topologies (Alencar et al., 2016). Handling of animals followed animal use protocols approved by the University of Nevada at Las Vegas Animal Care Committee (R701-1105-203) and the University of Washington Institutional Animal Care and Use Committee (4209-01).

Genomic DNA was extracted from tissue samples using the Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA, USA). Aliquots of extracts were shipped to RAPiD Genomics (Gainsville, FL, USA) for UCE library preparation and sequencing following Bryson et al. (2017). Briefly, 5,472 custom-designed probes (MYbaits; MYcroarray, Inc., Ann Arbor, MI, USA) were used to enrich 5,060 UCE loci using protocols specified in Faircloth et al. (2012). Libraries were sequenced on an Illumina HiSeq 3,000 (100 paired-end) at the University of Florida ICBR Facility. Data quality filtering and assembly were performed using the PHYLUCE pipeline (Faircloth, 2016) following Bryson et al. (2017). A locus was retained in the final assembly if it was represented by >75% of taxa. Alignments were deposited in Dryad (https://doi.org/10.5061/dryad.4467407).

2.2 | Concatenated phylogenetic analysis

All phylogenetic analyses were performed using the high-performance computer cluster at the Center for Theoretical Physics at New York City College of Technology (CUNY). We performed both maximum likelihood (ML) and Bayesian phylogenetic analyses on the concatenated (75% complete) matrix. Unpartitioned ML analyses were performed using RAxML v. 8.2.10 (Stamatakis, 2014) under a GTRGAMMA substitution model. We used the -f a option to calculate bootstrap support values and infer the ML tree in a single analysis. Nodal support was assessed through 100 rapid bootstrap replicates. All trees were rooted using *A. piscivorus*.

Unpartitioned Bayesian analyses were performed using EXABAYES v. 1.5 (Aberer, Kobert, & Stamatakis, 2014). EXABAYES is a MCMC package specifically geared towards Bayesian inference of large phylogenomic data sets, providing multiple analytical solutions for high-level parallelism. All analyses utilized a general time reversible model with gamma distributed rate heterogeneity (GTRGAMMA). To make sure that chains were not getting stuck in local optima, we implemented two independent runs (-R 2) with Metropolis-Coupling (three heated and one cold chain per run [-C 4]) in parallel to better sample parameter space. Sampling proposal weights included the following: likeSpr = 4, parsimonySPR = 8, stNNI = 4, eSPR = 4, blDistGamma = 7, branchMulti = 2. All analyses were implemented on 64 cores for at least 1 million generations (sampling every 500) using default priors for all parameters. Mixing and stationarity were monitored using TRACER v. 1.6.0 (Rambaut & Drummond, 2007) with a target effective sample size (ESS) value of 200 for all parameters. The average standard deviation of split frequencies (ASDSF) was also used as a measure of convergence using the default threshold value of 5%. Extended majority-rule (MRE) consensus trees were generated through the *consense* script. The default burn-in of 25% was used for all post-run analysis. Unrooted topologies were rooted using *A. piscivorus*.

The optimal partitioning strategy for all partitioned analyses was determined using PARTITIONFINDER v. 2.1.1 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2017). Because of the large size of the data set, we used the relaxed hierarchical clustering algorithm to assign sites to different partitions (Lanfear, Calcott, Kainer, Mayer, & Stamatakis, 2014). This clustering algorithm has been shown to be nearly as accurate as the greedy algorithm implemented in PARTITIONFINDER, but allows for the scaling of up to thousands of loci. Although the strict hierarchical clustering algorithm can further improve computational performance, the relatively high error rates preclude its utility in large phylogenomic data sets (Lanfear et al., 2014). We limited substitution models to GTR + G as this is the standard model implemented for nucleotide data in RAxML and currently the only nucleotide model in EXABAYES. Initial data blocks were assigned by locus from the output provided by the PHYLUCE pipeline. Default weights (i.e., 1,0,0,0) were used to calculate similarity among partitions (subsets), which were based solely on differences in substitution rate among sets. We specified an *rcluster-max* of 1,000 and *rcluster-percent* of 10. AICc was used for all partitioning scheme selection and the *--raxml* option was executed for analysis (Stamatakis, 2014). We used the ML tree from the unpartitioned RAxML analysis as the starting tree for PARTITIONFINDER. The optimal partitioning scheme was then utilized in partitioned RAxML and EXABAYES analyses using the same parameter settings as the unpartitioned analyses.

2.3 | Coalescent-based species delimitation and species tree analysis

We first used SVDQUARTETS (Chifman & Kubatko, 2014, 2015) in PAUP* v. 4.0a152 (Swofford, 2001) for coalescent-based phylogenomic analysis. SVDQUARTETS uses site pattern frequencies in SNP or multilocus sequence data to infer singular value decomposition (SVD) scores among alternative unrooted quartet trees (lower score

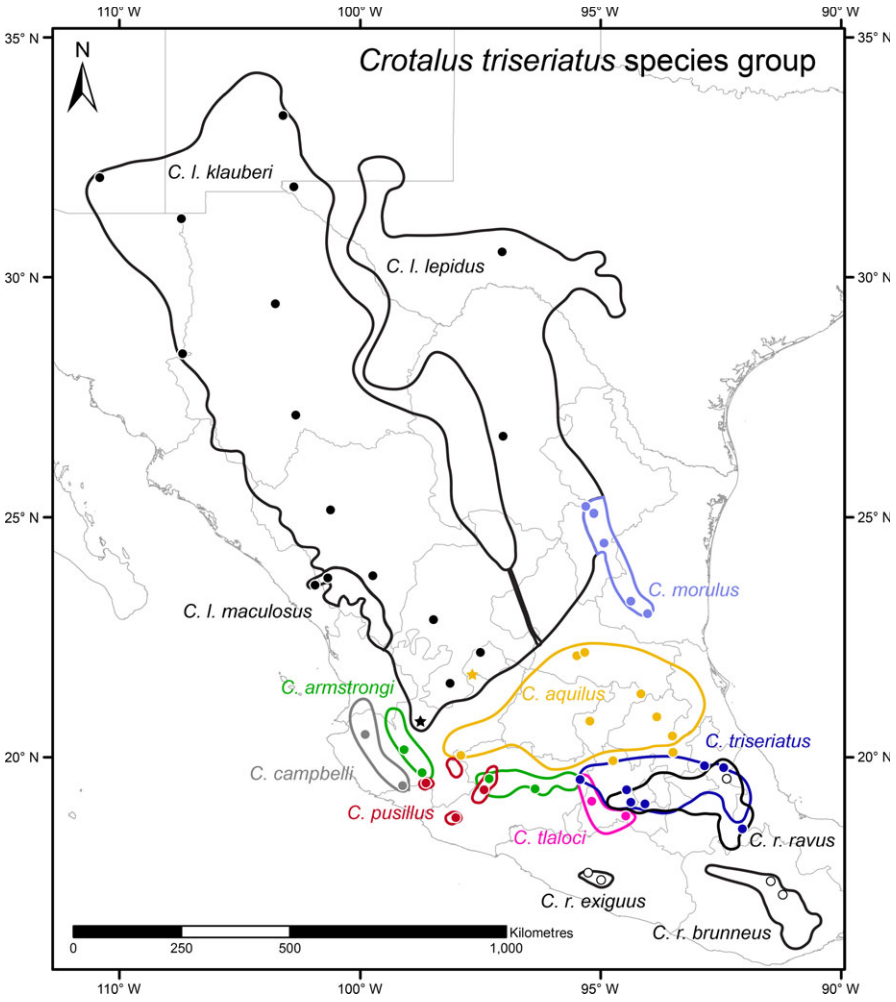

FIGURE 1 Map of all taxa used in this study. Species are colour-coded based on previous taxonomy (Bryson et al., 2014, 2011); distributions adapted from Campbell and Lamar (2004). Two samples indicated with stars were identified as probable hybrids in our study

better). A tree amalgamation step is subsequently used to assemble a full tree containing all taxa. We first ran an analysis utilizing every sequence as a terminal (i.e., no species assignments) to estimate a lineage tree statistically consistent with the multispecies coalescent. We exhaustively sampled all 341,055 quartets. Following the calculation of SVD scores for alternative quartets, *Quartet FM* (QFM; Reaz, Bayzid, & Rahman, 2014) was used in PAUP* to assemble the full tree. Support for nodes was assessed using nonparametric bootstrapping with 100 replicates. Trees were rooted with *A. piscivorus*.

We used BPP v. 3.3a (Yang & Rannala, 2010) for joint species delimitation and species tree inference (i.e., unguided species delimitation; Yang & Rannala, 2014). BPP utilizes the multispecies coalescent model in a fully probabilistic (Bayesian) framework for parameter estimation, explicitly accounting for gene tree/species tree discordance due to incomplete lineage sorting. Due to the relatively high computational burden of Bayesian methods like BPP, we performed inference on reduced data sets consisting of a subset of 100 of the most variable loci. Geneious v. 9.1.7 (Kearse et al., 2012) was used to select loci based on the percentage of variable sites. Assignment of individuals to putative species followed previous studies (Bryson et al., 2014, 2011) and the results of the concatenated phylogenetic analyses. Two samples, Ct143ClkJAL and Ct2CaqAGS, were placed in strongly supported yet conflicting relationships in concatenated and coalescent analyses (see below) in a pattern suggestive of historical introgression. Both were subsequently excluded from the analysis to meet methodological assumptions of the method (i.e., no hybridization among species; Yang & Rannala, 2010, 2014), leaving a total of 17 groupings and 53 individuals. *Crotalus lepidus* was divided into six populations consistent with geography and our phylogenetic analyses: *C. l. lepidus*, *C. l. maculosus*, *C. l. klauberi*-3 Northwest (samples from the north-eastern Sierra Madre Occidental and adjacent sky islands), *C. l. klauberi*-4 Northeast (samples from the dry slopes of the north-eastern Sierra Madre Occidental and adjacent uplands), *C. l. klauberi*-2 Durango (samples from the Sierra Madre Occidental of Durango) and *C. l. klauberi*-1 South (samples from the Sierra Madre Occidental south of the Rio Mezquital drainage). We specified gamma priors of *G*(2, 1,000) for both population sizes (θs) and the divergence time of the root (τ_0). The *locusrate* parameter was used to account for rate variation among loci using *D* (α) = 10. Uniform rooted trees were used as the species model prior. Following a burn-in of 10,000 generations, chains were sampled every five generations for a total of 20,000 samples. Multiple analyses were run using different species delimitations algorithms (Algorithms 0 and 1) to check for indications of convergence. We also tested the influence of the prior on θs and τ_0 as previous research suggests that these priors may influence species delimitation results

(Blair, Mendez de la Cruz, Law, & Murphy, 2015; Leaché & Fujita, 2010; Yang, 2015). We specifically compared results between *G*(2, 1,000) and *G*(2, 100). We also performed additional species delimitation analyses using 500 of the more variable loci to compare to the analyses using 100 loci. Parameter settings were similar to the 100 loci analyses, but in this case we used a burn-in of 20,000 and a sample frequency of 10. In total, 16 different species delimitation analyses were performed.

We also used BPP for species tree inference (Analysis 01) under the multispecies coalescent model using the 500 loci data set. Individuals were partitioned into 17 "species" following the results of the species delimitation analyses (see Results). Three independent MCMC runs were implemented with a burn-in of 10,000, a sample frequency of 5, and 20,000 total samples. We specified gamma priors of *G*(2, 1,000) for both population sizes (θs) and the divergence time of the root (τ_0). The species model prior assumed uniform rooted trees.

Based on the species delimitation results obtained from BPP, we performed a coalescent-based species tree analysis on the full concatenated matrix of 3,384 UCE loci in SVDQUARTETS. Individuals (n = 53) were assigned to 17 species, and Ct143ClkJAL and Ct2CaqAGS were excluded, matching our BPP species tree analyses. We performed a species tree search using all quartets (195,833) and used QFM to assemble the species tree. Node support was assessed using 100 nonparametric bootstrap replicates. We also performed additional species tree analyses in SVDQUARTETS using concatenated subsets of the total data (50 UCE loci, 100 UCE loci, 1,000 UCE loci, 2,000 UCE loci) to assess levels of congruence. All quartets were evaluated for each subset and 100 nonparametric bootstrap replicates were used to calculate support values.

Finally, we used the gene tree reconciliation method ASTRAL-III v. 5.5.6 (Mirarab et al., 2014; Mirarab & Warnow, 2015; Zhang et al., 2017) to infer a species tree that is statistically consistent with the multispecies coalescent. ASTRAL takes as input a set of unrooted gene trees and finds the species tree containing the maximum number of induced quartets that are consistent with the gene trees. Several studies have indicated that ASTRAL is one of the more accurate gene tree reconciliation methods currently available (e.g., Chou et al., 2015; Mirarab et al., 2014). However, caution is often warranted as poor resolution gene trees can substantially influence the accuracy of the inferred species tree (Mirarab et al., 2014; Xi et al., 2015). We began by manually inspecting the 500 loci used in BPP. We used GENEIOUS to remove individuals with all missing data (?s), the two putative hybrid samples, individuals with highly fragmented data (Hosner, Faircloth, Glenn, Braun, & Kimball, 2016) and loci without the outgroup sequence. We also removed loci with extremely low levels of variability that were unlikely to yield useful gene trees. RAxML was used with the *raxml_wrapper.pl* script (https://sco.h-its.org/exelixis/web/software/raxml/index.html) to process the 351 alignments (-f d search under GTRGAMMA). Initial runs were used to prune identical sequences from each alignment and subsequent runs searched for ML trees using unique haplotypes. All 351 Ml gene trees were used in ASTRAL, and we created a file to assign individuals to species. Support for relationships in the ASTRAL species tree was quantified using local posterior probabilities (Sayyari & Mirarab, 2016). We did not include ML bootstrap trees in analyses (i.e., multilocus bootstrapping) as recent research suggests that local posterior probabilities may be more a more accurate measure of support (Sayyari & Mirarab, 2016). The unrooted ASTRAL species tree was subsequently rooted with *A. piscivorus*.

2.4 | Divergence time estimation

Because of the large size of the data set, we estimated divergence times on the fixed BPP species tree topology using MCMCTree within PAML v. 4.9e (Rannala & Yang, 2007; Yang, 2007; Yang & Rannala, 2006). MCMCTree is particularly suitable for Bayesian estimation of divergence times under alternative clock models in large, next-generation data sets where joint estimation of tree topology and divergence times is computationally problematic (Reis & Yang, 2011). Divergence times were estimated using a pruned alignment containing a single sample per lineage. PAUP* v. 4.0a161 (Swofford, 2001) was used to test for a strict molecular clock using likelihood ratio tests, AICc, and BIC under a HKY + I + G substitution model. As the results strongly rejected the molecular clock (see Results), we used the approximate likelihood method and independent rates clock model in MCMCTree (Reis & Yang, 2011). Analyses used a HKY + G4 substitution model, default priors for kappa (*G*[6, 2]) and alpha (*G*[1, 1]), *G*(2, 2,000) for substitution rate, and *G*(1, 10) for sigma. A single alignment partition of the full 3,384 loci was used to minimize the number of parameters and ensure adequate mixing and convergence. Sampling included a burn-in of 1 million generations followed by 10,000 posterior samples drawn every 5,000 generations. We calibrated the root node using information from a recent fossil-calibrated rattlesnake phylogeny (Blair & Sánchez-Ramírez, 2016). Specifically, soft bounds between 20 and 12 Ma were placed on the MRCA of *Crotalus* and *Agkistrodon*. Convergence was assessed by running all analyses twice and monitoring effective sample sizes (ESS) in Tracer v. 1.6 (Rambaut & Drummond, 2007). We also tested the impact of the sigma prior on divergence times by re-running analyses using a *G*(1, 100) prior. Finally, MCMCTree was run without data (i.e., sampling the prior only; usedata = 0) to quantify the information content in the sequence data and make sure that the choice of priors for divergence times was reasonable.

2.5 | Network estimation

Because strictly bifurcating trees may not adequately depict evolutionary history, we used the pseudolikelihood method SNaQ (Solís-Lemus & Ané, 2016) in the PhyloNetworks package (Solís-Lemus, Bastide, & Ané, 2017) to estimate a phylogenomic network. SNaQ is a recently developed algorithm that can explicitly account for both incomplete lineage sorting and introgression/gene flow. Previous studies have suggested that a large number of loci may be needed to accurately detect hybridization events (Solís-Lemus & Ané, 2016), so we used all 3,384 loci and all 55 individuals, each of which was

assigned to species following our species tree analyses. We began by running RAxML on each of the 3,384 alignments (ML + bootstrap search). We then used the best trees to determine concordance factors at the level of individuals. This data frame was then modified to create a concordance factor table at the level of species that was used as input for SNaQ. A more comprehensive discussion of running SNaQ using multiple alleles per species can be found in the online documentation of SNaQ/PhyloNetworks. We tested hmax values of 0–4 and selected the best value by plotting hmax vs. loglik. Ten independent runs were implemented for each value of hmax. For hmax = 0, we used the ASTRAL tree as the starting topology. For all remaining hmax values, we used the best network estimated from previous hmax values as the starting topology.

3 | RESULTS

3.1 | Data set characteristics

We sequenced 3,384 UCE loci (after quality filtering) for 54 specimens of montane rattlesnakes in the *C. triseriatus* species group and the outgroup. Out of a total of 2,057,530 characters, 1,978,984 were constant, 40,907 were variable and parsimony-uninformative, and 37,639 were variable and parsimony-informative. The proportion of gaps and undetermined characters in the concatenated data matrix was 18.4%, and pairwise identity was 99.4% with 94.7% of sites identical (including the *Agkistrodon* outgroup), illustrating the high degree of conservatism in UCE loci. Also typical of UCE data, the data showed an AT-bias, with a %GC-value of 38.1.

3.2 | Concatenated analysis

Unpartitioned ML analyses of the concatenated matrix under a GTRGAMMA substitution model yielded a topology with strong support for most nodes (Figure 2). The majority of presently defined species and subspecies were monophyletic, with the exception of *C. aquilus* and *C. lepidus klauberi*, the latter of which was placed throughout a large clade in the phylogeny with strong bootstrap support. The two samples of *C. lepidus lepidus* were deeply nested within a clade of *C. lepidus klauberi*. The majority of nodes towards the base of the tree received high bootstrap support, with the newly described *C. tlaloci* placed sister to *C. pusillus*, and *C. campbelli* sister to *C. armstrongi*. *Crotalus ravus* was strongly supported as sister to the remaining sampled rattlesnakes. *Crotalus morulus* was also fully resolved and placed sister to a clade containing *C. lepidus maculosus*, *C. lepidus klauberi* and *C. lepidus lepidus*. One sample of *C. aquilus* (Ct2CaqAGS) was placed within a clade of *C. l. klauberi*.

All unpartitioned ExaBayes analyses indicated adequate convergence (ESS values >200, ASDSF < 0.05). In fact, both runs converged to the identical posterior distribution (ASDSF = 0) after 20,000 generations. However, chains were run for a full 1 million generations to adequately sample all parameters. The topology of the ExaBayes MRE tree was identical to the ML tree with all nodes fully resolved (posterior probability = 1.0).

The *rcluster* algorithm in PartitionFinder clustered the concatenated alignment into 2,005 subsets consisting of 20,156 parameters (lnL = −3,796,380.79; AICc = 7,633,472.41). This partitioning scheme was then used to perform partitioned ML and Bayesian analyses in RAxML and ExaBayes, respectively, assigning independent GTRGAMMA models to each partition. The ML tree from the partitioned analysis was identical to the unpartitioned ML and Bayesian trees, with high bootstrap support for the majority of nodes. Similar to the unpartitioned analysis, the partitioned ExaBayes runs attained excellent convergence relatively rapidly (ASDSF = 0). However, ESS values for two parameters (LnPr and TL) were low due to large fluctuations in values throughout the runs. All other parameters had excellent ESS values (>200). This indicated potential issues with the estimation of branch lengths on large partitioned ExaBayes analyses. Nevertheless, we created a MRE consensus tree from the two partitioned ExaBayes runs and the topology was identical to the unpartitioned ExaBayes analysis and the partitioned and unpartitioned RAxML analyses. However, branch length estimates were off by an order of magnitude.

3.3 | Coalescent-based species delimitation and species tree analysis

The bootstrap consensus tree from the lineage-based (without a priori species assignments) SVDquartets analysis conflicted with the inferred topology from the RAxML and ExaBayes analyses, although support for conflicting relationships was mixed (Supplementary Fig. S1). In the SVDquartets tree, following the divergence of *C. ravus*, a *C. armstrongi/C. campbelli* clade was strongly supported as sister to all remaining rattlesnakes. Similar to the concatenation (i.e., non-coalescent) analyses, *C. tlaloci* was strongly supported as sister to *C. pusillus*. Relationships within the *C. aquilus*, *C. lepidus* and *C. morulus* clade were ambiguous and contradictory to relationships inferred from the concatenated RAxML and ExaBayes analyses. However, many nodes in this clade suffered from low bootstrap support. One sample of *C. l. klauberi* (Ct143ClkJAL) was placed within a clade of *C. aquilus*.

We performed multiple coalescent-based species delimitation analyses in bpp to help determine if the genetic lineages inferred in our phylogenetic analyses may represent distinct species or reproductively isolated populations. Alternative runs, priors and search algorithms gave consistent results, indicating that chains were run for a sufficient length of time. There was overwhelming support for a species delimitation model where each of the 17 lineages (including the outgroup) represent distinct populations or species (Figure 3; Supplementary Table S2). Posterior probability values for each "species" were 1.0 for the majority of taxa in the tree in both the 100 and 500 loci data sets. Only four taxa received less than full support in our analyses: *C. lepidus klauberi*-2 (Durango), *C. lepidus klauberi*-3 (Northwest), *C. lepidus klauberi*-4 (Northeast) and *C. lepidus lepidus*. However, posterior probability values remained high, particularly for analyses based on 500 loci (Figure 3). Based on these results, we treated these taxa as distinct lineages for downstream species tree analysis.

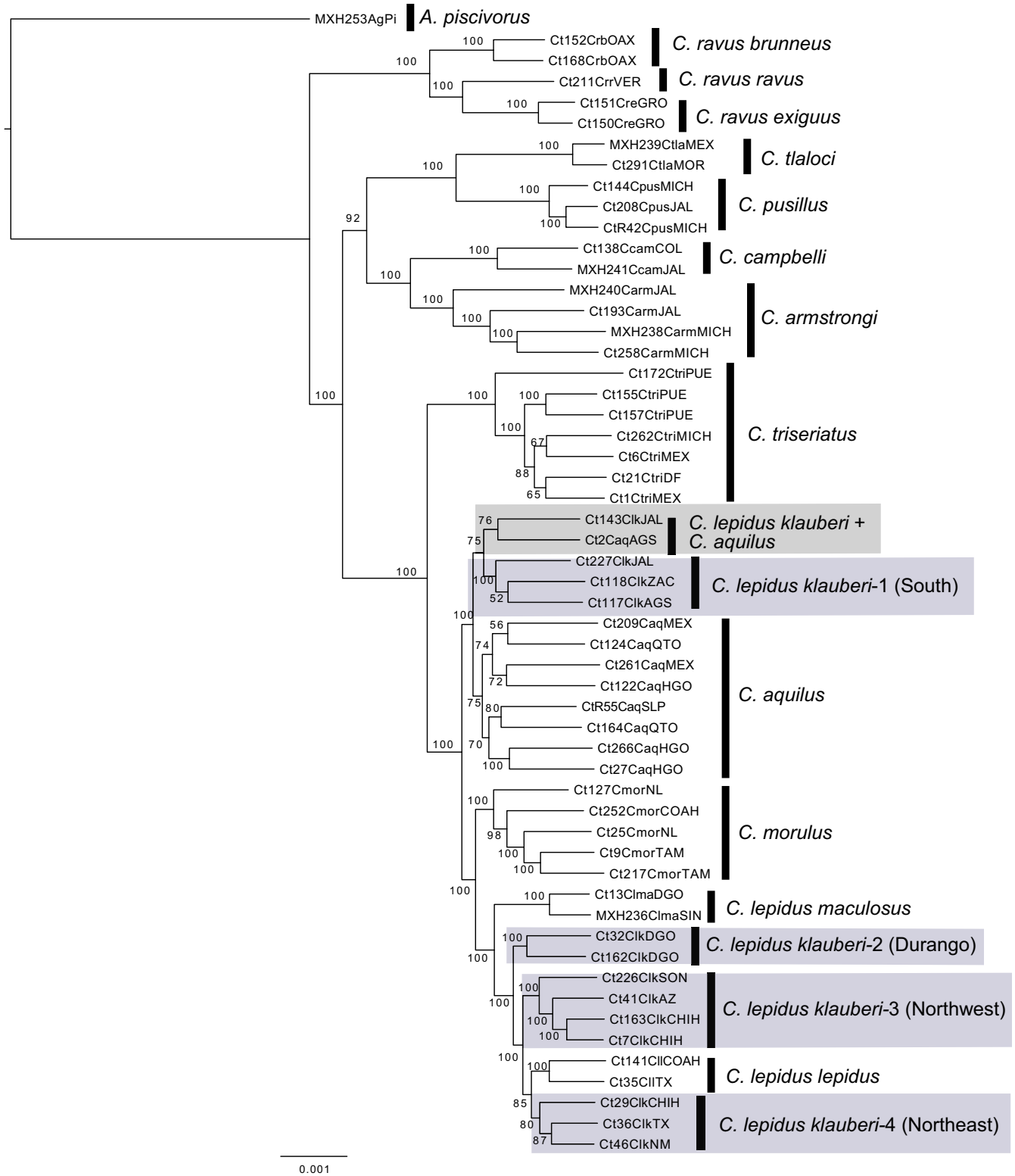

FIGURE 2 Phylogenomic relationships of montane rattlesnakes of the *Crotalus triseriatus* group based on unpartitioned concatenated maximum likelihood (ML) analysis of 3,384 UCE loci (2,057,530 bp) under a GTRGAMMA model of substitution. Values adjacent to nodes represent rapid bootstrap values from a joint bootstrap + ML search in RAxML (-f a option). Blue boxes indicate geographical lineages within *C. lepidus klauberi*; grey box indicates putative hybridization/introgression between *C. aquilus* and *C. lepidus klauberi*. MRE consensus tree from unpartitioned ExaBayes analyses was identical, with full posterior probability values (1.0) for all nodes

We used the species delimitation model with the highest posterior probability for additional coalescent-based species tree analyses. The three independent BPP runs were largely consistent with one another, suggesting strong signal in the data and that chains were run for an acceptable duration. The only topological difference among runs occurred in the clade containing *C. lepidus klauberi*-3 (Northwest), *C. lepidus klauberi*-4 (Northeast) and *C. lepidus lepidus*. Two of the three runs placed *C. lepidus lepidus* as sister to *C. lepidus klauberi*-4 (Northeast), whereas the third run placed *C. lepidus klauberi*-3 (Northwest) sister to *C. lepidus klauberi*-4 (Northeast). The tree topology of the former two runs was identical to the concatenated ML and Bayesian trees (Figures 2 and 4a).

We also used the species delimitation results from BPP for subsequent species tree analysis in SVDQUARTETS. The resulting species tree differed from the RAxML and ExaBayes trees, the BPP species tree and the SVDQUARTETS lineage tree, particularly with regard to deeper relationships (Figure 4b). For example, BPP and the concatenated ML and Bayesian analyses placed *C. campbelli* and *C. armstrongi* as sister to a clade containing *C. tlaloci* and *C. pusillus*. In contrast, the coalescent-based species tree analysis in SVDQUARTETS placed *C. campbelli* and *C. armstrongi* as sister to all remaining rattlesnakes except *C. ravus*. In addition, the concatenated analyses and BPP placed *C. aquilus* sister to one of the *C. lepidus klauberi* lineages (Figure 2 and 4a), whereas SVDQUARTETS placed *C. aquilus* as sister to a clade consisting of *C. morulus*, *C. lepidus maculosus*, *C. lepidus lepidus* and the three remaining *C. lepidus klauberi* lineages (Figure 4b). Bootstrap support was high for the majority of nodes in the SVDQUARTETS species tree (>70), except for the node uniting *C. campbelli* and *C. armstrongi* (63) and the node representing the MRCA of the majority of the species group (59).

The species tree inferred from ASTRAL analysis of 351 gene trees was largely consistent with the trees inferred from concatenated ML and Bayesian analyses and coalescent analyses in BPP (Figure 4c). The majority of differences involved nodes with low support values. *Crotalus campbelli* and *C. armstrongi* were placed sister to a clade containing *C. tlaloci* and *C. pusillus*, albeit with low local

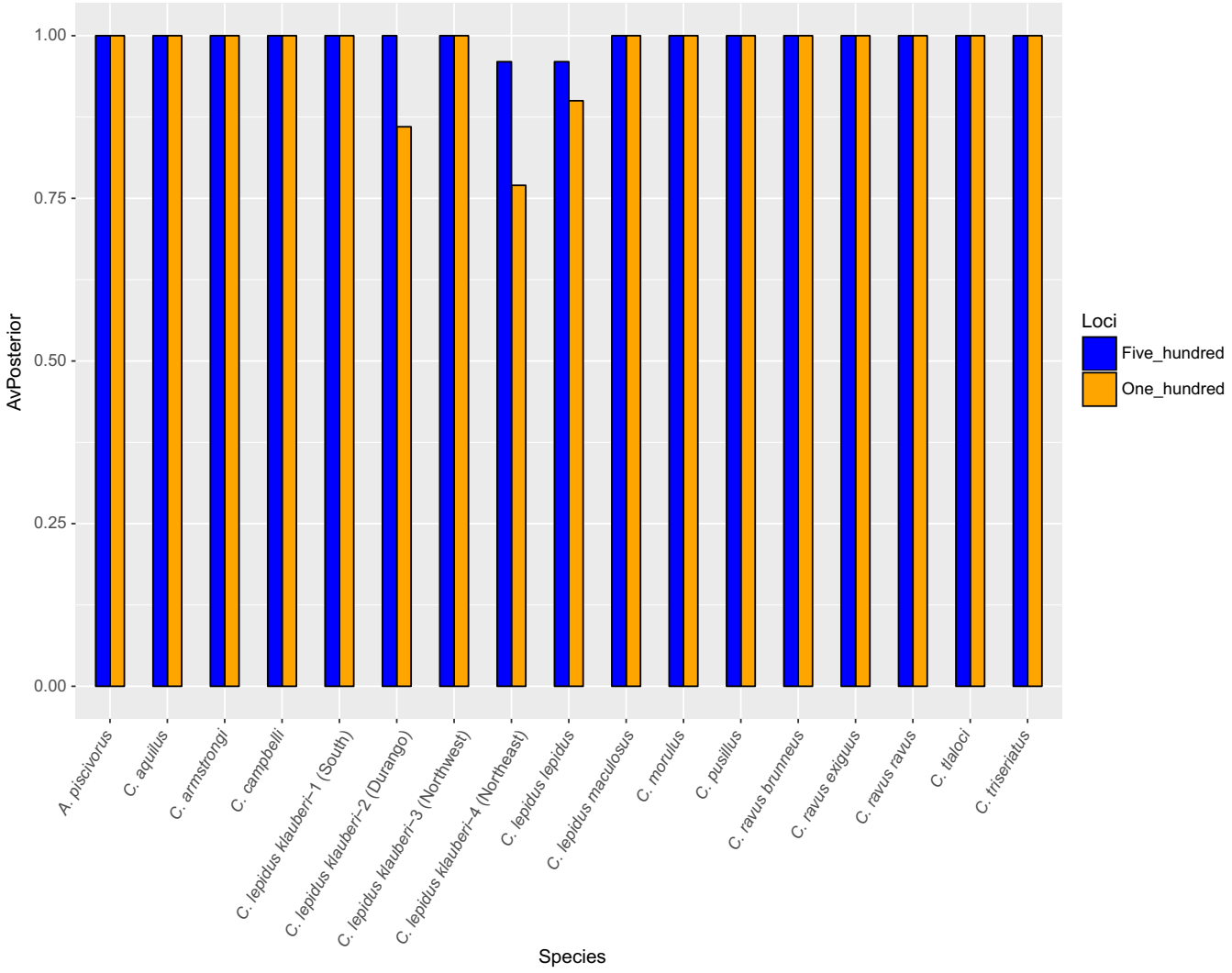

FIGURE 3 Bayesian posterior probability values for different species within the *Crotalus triseriatus* species group. Values represent averages over eight different BPP analyses using different search algorithms, priors and independent MCMC runs. Orange bars represent averages based on 100 loci and blue bars represent averages based on 500 loci

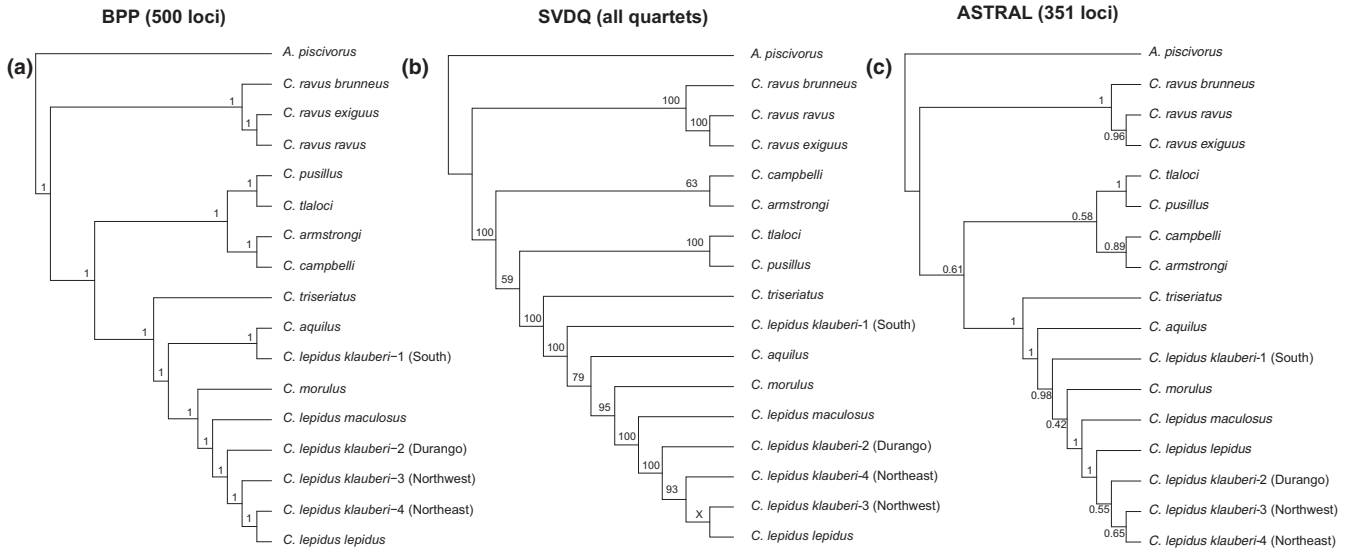

FIGURE 4 Coalescent-based species trees of rattlesnakes of the *Crotalus triseriatus* group based on alternative search algorithms. Individuals (53) were assigned to 17 species based on results from prior coalescent-based species delimitation analyses. (a) Majority-rule consensus tree from BPP analysis of 500 of the more variable loci. Values at nodes represent posterior probability values. (b) Species tree inferred from SVDQUARTETS analysis. Tree is the SVDQUARTETS tree inferred using the concatenated matrix of 2,057,530 bp. Values indicate nonparametric bootstrap support from 100 replicates. "X" indicates nodes that were not recovered in the bootstrap consensus tree. (c) Species tree inferred from ASTRAL analysis of 351 gene trees inferred under a GTRGAMMA model in RAxML. Values at nodes represent local posterior probability values

posterior probability. In the ASTRAL tree, *C. lepidus klauberi*-1 (South) was more closely related to *C. morulus* and other *C. lepidus* lineages vs. *C. aquilus*, which differed from relationships inferred using the other species tree methods utilized.

We explored how the size of the concatenated data matrix might influence species tree inference in SVDQUARTETS by inferring trees based on 50, 100, 1,000 and 2000 UCE loci. Discordant results were obtained from each analysis, although most incongruence was restricted to poorly supported nodes (Figure 5). No tree was identical to the tree inferred through the analysis of the full 3,384 loci. The 100 loci analysis had the highest number of nodes with bootstrap values <70. Thus, there appear to be additional nuances of the data other than size that influence node support.

3.4 | Divergence time estimation

A strict clock model was strongly rejected using a likelihood ratio test (difference = 184.8485; $p < 0.001$), AICc (difference = 339.6962) and BIC (difference = 151.6418). Adequate convergence was obtained in multiple runs of MCMCTree (Figure 6), with all ESS values >200. Virtually identical results were obtained using smaller gamma priors for sigma (*G*[1, 100]), and thus, we present the results using *G*(1, 10) only. The MRCA of the *C. triseriatus* group dated back to around 10 Ma. Divergence within the three subspecies of *C. ravus* occurred in the Pliocene around 4–5 Ma. The newly described *C. tlaloci* diverged from its sister taxon, *C. pusillus*, close to 5 Ma. *Crotalus campbelli* diverged from *C. armstrongi* about 5.5 Ma. Independent lineages within *C. lepidus klauberi* originated within the past 5 million years. The estimated mean rate of substitution (μ) was 3.57E-4 subs/site/My (95% HPD = 2.496E-4–4.696E-4). We also ran MCMCTree without data (i.e., sampling from the prior only) to determine the information content in the sequence data and make sure that the priors on divergence times were adequate. All divergence time priors were reasonable, yet different from posterior estimates that included the sequences (Supplementary Table S3). Thus, we concluded that the data were efficient to estimate divergence times and that our root calibration scheme was adequate.

3.5 | Network analysis

There was a sharp increase in pseudolikelihood score from hmax = 0 to hmax = 1 (−766.252 vs. −541.526) followed by a more gradual improvement with additional reticulations (Supplementary Fig. S2). Thus, a single hybridization event best represented the data. The SNaQ major tree was virtually identical to the topologies from ASTRAL, BPP and concatenation (Figure 7). Topological differences were predominantly restricted to the clade consisting of *C. lepidus lepidus* and the different populations of *C. lepidus klauberi*. The single inferred reticulation event involved the mrca of the *C. lepidus* + *C. morulus* clade contributing genes (γ = 0.194) to *C. armstrongi*.

4 | DISCUSSION

4.1 | Phylogenomic analysis

One of our main objectives was to examine levels of discordance in commonly used phylogenomic packages on a large empirical data

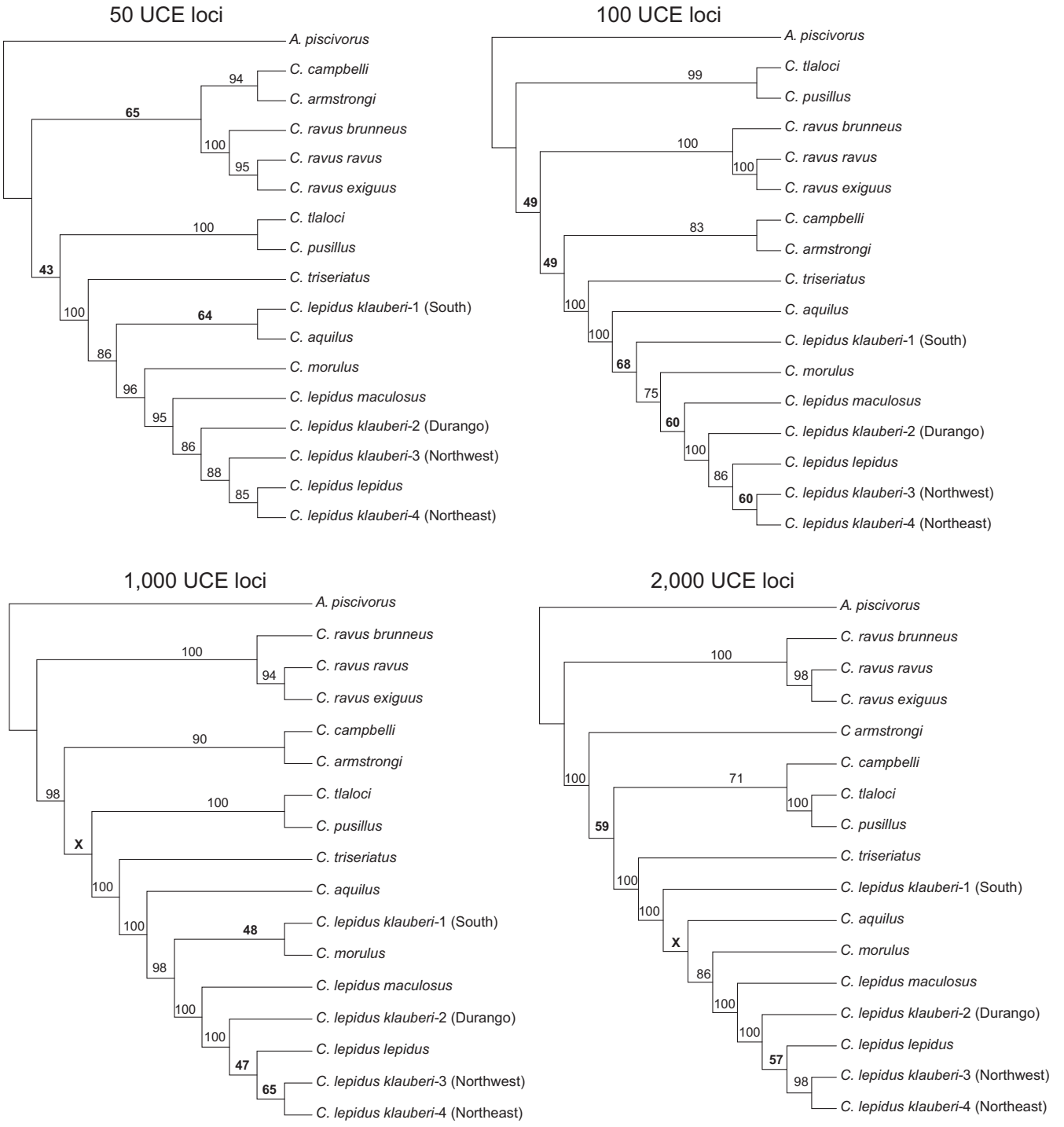

FIGURE 5 Species trees inferred by SVDQUARTETS based on different subsets of the total data set. Loci were selected randomly and concatenated for each analysis. Individual–species assignments followed analyses on the full data and were based off results from BPP. Trees shown are the SVDQUARTETS trees with bootstrap support values mapped on branches (100 replicates). Numbers in bold indicate bootstrap values below 70%. "X" indicates nodes that were not recovered in the bootstrap consensus tree

set. With the realization that gene tree topologies do not always agree with the species tree due to multiple processes including incomplete lineage sorting (Degnan & Rosenberg, 2006, 2009; Edwards, 2009; Maddison, 1997), substantial effort has been placed on the development of sophisticated algorithms to explicitly account for gene tree discordance when estimating species trees (e.g., Heled & Drummond, 2010; Liu et al., 2009; Liu et al., 2010; Mirarab et al., 2014; Vachaspati & Warnow, 2015). The species tree methods we select for comparison include concatenation, a fully probabilistic coalescent method (BPP), a gene tree reconciliation method (ASTRAL-III)

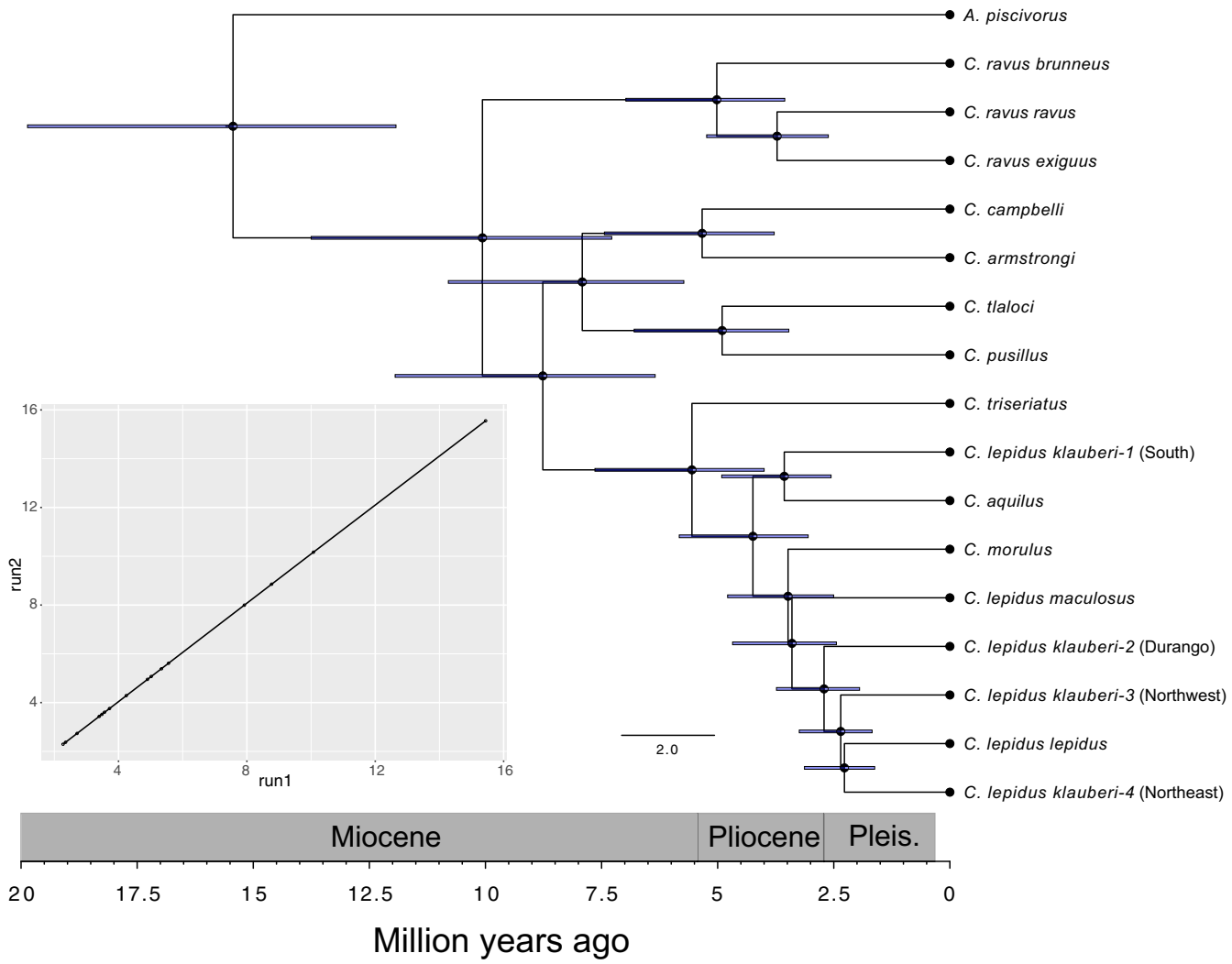

FIGURE 6 Divergence times of rattlesnakes of the *Crotalus triseriatus* group based on 2,057,530 bp. Parameter inference was based on approximate likelihood calculations in MCMCTree using the BPP species tree as a fixed topology and one individual per lineage. Horizontal bars represent 95% HPD confidence intervals for divergence times. Inset plot shows estimated mean divergence times from two independent MCMCTree runs. High levels of concordance between runs indicate that chains were run for a sufficient length of time and convergence was reached

and a site-specific technique based on algebraic statistics (SVDQUARTETS). These methods are selected based on both our data set characteristics and the relative performance of methods in simulation and empirical studies (Chifman & Kubatko, 2014; Chou et al., 2015; Mirarab et al., 2014; Mirarab & Warnow, 2015; Yang, 2015; Yang & Rannala, 2010). Both our concatenated ML and Bayesian analyses provide a well-supported phylogenomic hypothesis for rattlesnakes in the *C. triseriatus* group. Interestingly, the topology inferred from concatenation (Figure 2) is identical to the topology inferred by fully probabilistic coalescent analyses in BPP (Figure 4a). Due to the computational burden of methods like BPP, we restrict our analysis to 500 UCE loci. Our results coupled with those of previous studies suggest that BPP is likely to be an accurate estimator of species trees even with subsets of the total data (Caviedes-Solis, Bouzid, Banbury, & Leaché, 2015; Shi & Yang, 2018). We are a bit surprised that BPP is not used more often for species tree inference, although it is widely used for the purposes of species delimitation (e.g., Blair et al., 2015; Leaché & Fujita, 2010). We encourage empiricists to explore the method with additional phylogenomic data sets.

There are now myriad gene tree reconciliation techniques available to estimate species trees (Liu et al., 2010, 2009; Mirarab et al., 2014; Mirarab & Warnow, 2015; Zhang et al., 2017). These two-step procedures first use individual alignments to estimate gene trees, which are subsequently used as input data for species tree inference. We examined ASTRAL as simulation studies have shown it to be one of the more accurate techniques (Chou et al., 2015; Mirarab et al., 2014). It is also one of the few methods that can handle multiple individuals per species. To try to minimize potential errors in species tree inference, we followed the recommendations of recent studies (e.g., Hosner et al., 2016; Meiklejohn et al., 2016; Xi et al., 2015) and limited our ASTRAL analysis to 351 of the more variable UCE loci with little fragmentary data and use RAxML for gene tree

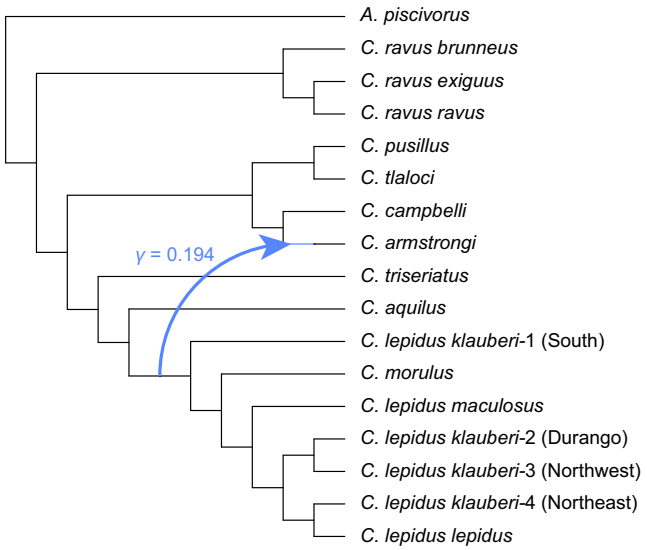

FIGURE 7 Phylogenomic network for rattlesnakes of the *Crotalus triseriatus* group based on SNaQ analyses of 3,384 loci. Individuals (55) were assigned to species (tips) based on other phylogenetic and species delimitation analyses. Concordance factors were calculated based on gene trees estimated in RAxML. Figure shows the network inferred with a single hybridization event (hmax = 1). Y = proportion of introgressed genes

reconstruction. The resulting species tree was highly congruent to the species trees inferred using concatenation and BPP. One difference that was strongly supported is the placement of *C. aquilus* and *C. lepidus klauberi*-1 (South), which were inferred as sister taxa in concatenation and BPP. Conversely, in the ASTRAL tree, *C. lepidus klauberi*-1 (South) was sister to a clade consisting of *C. morulus* and the remaining *C. lepidus* taxa. Other differences in the ASTRAL tree involved lineages within *C. lepidus klauberi*, but these relationships were weakly supported. Thus, although our ASTRAL results were consistent with other methods examined, additional studies are needed to help determine how large phylogenomic data sets should be pruned and subsampled prior to species tree inference.

Although these gene tree reconciliation methods are relatively quick, there are numerous additional considerations that researchers must be aware of when using these methods. For example, gene tree error has been shown to negatively influence the resulting species trees (Mirarab et al., 2016, 2014; Roch & Warnow, 2015; Xi et al., 2015). Unfortunately, gene tree error/uncertainty is often ubiquitous when working with loci with relatively few informative characters such as UCEs. Second, missing data can also negatively impact gene tree reconciliation techniques, most likely to a greater degree than concatenation and site-based methods (Burleigh, Kimball, & Braun, 2015; Hosner et al., 2016). Missing data can come as either an entire sequence missing for some species in an alignment or species having partial (fragmented) sequences that reduce information content. Other studies have also shown a measurable impact of the method used to reconstruct gene trees (Meiklejohn et al., 2016; Xi et al., 2015). Recently, Streicher et al. (2016) used a large UCE data set from iguanian lizards to quantify the effects of taxon and locus sampling on species tree inference. Their results based on concatenation and coalescent analyses (NJ_{st}) suggest that including a large proportion of missing data (including loci with up to 50% missing taxa) may be beneficial in phylogenomic studies. However, results and conclusions were based off branch support values in the unknown iguanian tree and differences were obtained between phylogenomic methods. In another recent study, Hosner et al. (2016) found that missing data (particularly partial/fragmented sequences) were a bigger issue for gene tree reconciliation approaches vs. concatenation and quartet-based methods like SVDQUARTETS. Their results also corroborated other recent studies (e.g., Meiklejohn et al., 2016) and suggest that gene tree reconciliation methods be restricted to a subset of the most variable loci to improve the quality of gene trees. In sum, there is still a great deal of uncertainty with how taxa and loci should be sampled in phylogenomics, particularly when using gene tree reconciliation approaches.

Because of the potential issues with gene tree reconciliation methods highlighted above, there has been substantial interest in species tree methods that work directly from the sequence data and are scalable to large phylogenomic data sets. One popular method that has emerged is the SVDQUARTETS algorithm currently implemented in PAUP*. The method is statistically consistent under the multispecies coalescent model, can work with both SNP data and concatenated UCE loci, is relatively quick, and can apparently work well with large quantities of missing data (Chifman & Kubatko, 2014; Chou et al., 2015; Leaché et al., 2016, 2015). These attributes make the methods particularly appealing to analyse large phylogenomic data sets generated through both sequence capture and RADseq-related approaches. The technique has been shown to perform well in simulation (Chou et al., 2015), but there are still little data available illustrating how SVDQUARTETS performs on empirical phylogenomic data (Hosner et al., 2016; Leaché et al., 2016, 2015). We performed multiple SVDQUARTETS analyses on different data sets to quantify discordant topological patterns. The lineage-based analysis (using individuals as terminals) differed from the concatenated ML and Bayesian analyses although bootstrap support values were relatively low. In the SVDQUARTETS lineage tree, *C. morulus* was not monophyletic and *C. campbelli* was nested with *C. armstrongi*. The SVDQUARTETS species tree based on all 3,384 loci also differed from species trees inferred from other methods. Of particular note is the early divergence of (*C. campbelli* + *C. armstrongi*) in the SVDQUARTETS tree, whereas other species tree analyses and concatenation placed these taxa as sister to (*C. pusillus* + *C. tlaloci*). However, bootstrap support for this placement was weak (59%), and it is likely that *C. campbelli* and *C. armstrongi* form a clade with *C. pusillus* and *C. tlaloci*.

To further probe how results from SVDQUARTETS may change based on data set size and characteristics, we performed additional species tree analyses on concatenated alignments consisting of 50, 100, 1,000, and 2000 UCE loci. Different topologies were recovered for each analysis, although bootstrap support at conflicting nodes was relatively weak. With only 50 UCE loci, the *C. ravus* clade was

placed sister to *C. campbelli* and *C. armstrongi* with a bootstrap value of 65%. This result disappeared in all analyses with more data. Our results also suggested an unclear relationship between data set size and bootstrap support values, indicating that there are additional nuances besides number of loci that contribute to strongly supported species trees. In general, our results are consistent with other recent studies that found lower node support values for SVDQUARTETS trees vs. trees inferred through concatenation (Leaché & Linkem, 2015; Leaché et al., 2016, 2015; Hosner et al., 2016). In addition, our results are similar to a recent study that examined phylogenomic relationships among gibbons using real and simulated data and found that species trees inferred by BPP and ASTRAL were much more consistent than those inferred by SVDQUARTETS (Shi & Yang, 2018). Although additional simulation and empirical studies are needed to examine characteristics of the SVDQUARTETS method in greater detail, it appears that relatively large data sets are needed to reach a topology and support values consistent with other methods of species tree inference, and that full likelihood-based approaches that adequately model rate heterogeneity and coalescent stochasticity may be preferable.

Strictly bifurcating trees may not adequately depict evolutionary history when gene flow and hybridization are prevalent (Solís-Lemus & Ané, 2016; Wen, Yu, Zhu, Nakhleh, & Posada, 2018; Yu & Nakhleh, 2015). Using SNaQ, we found hybridization within the sampled taxa of the *C. triseriatus* group to be relatively rare. Results provided strong support for a model consisting of a single reticulation event from the mrca of *C. lepidus* + *C. morulus* into *C. armstrongi*. The major tree inferred by SNaQ was virtually identical to the ASTRAL tree and very similar to the trees inferred by BPP and concatenation. Interestingly, SNaQ detected no introgression between *C. aquilus* and *C. lepidus klauberi*-1 (South) even though this was suggested in our other phylogenetic analyses. It is possible that introgression between these two taxa was limited to two individuals in our study (Ct143ClkJAL and Ct2CaqAGS), and the inclusion of larger number of samples from each species masked the signal of hybridization. Additional sampling of both species from near where their ranges meet should further elucidate the propensity for introgression in these snakes.

Interest in phylogenomic networks continues to increase as more evidence accumulates suggesting that both incomplete lineage sorting and reticulation/hybridization may be a common component of the evolutionary history of several groups (Solís-Lemus & Ané, 2016; Solís-Lemus et al., 2017; Wen et al., 2018). Unfortunately, many of these explicit network methods are computationally demanding (Than, Ruths, & Nakhleh, 2008), necessitating the use of some type of approximation. New pseudolikelihood methods such as SNaQ and similar techniques in the PhyloNet package (Yu & Nakhleh, 2015; Wen et al., 2018) show promise for analysing large phylogenomic data sets. Combining tree-based and network-based phylogenomic methods, each with their own strengths and weaknesses, will likely provide the most robust conclusions regarding evolutionary history. Further, although our inferred divergence times are broadly concordant with previous studies (e.g., Blair & Sánchez-Ramírez, 2016; Bryson et al., 2011), more research is needed to determine how reticulated histories influence inferred divergence times.

4.2 | Systematics and biogeography of the *C. triseriatus* group

The Mexican highlands are a known biodiversity hotspot (Mittermeier et al., 2017), and multiple phylogeographic studies have documented cryptic diversity distributed throughout high-elevation biomes within Mexico's major mountain systems (e.g., Bryson, García-Vázquez, & Riddle, 2012; León-Paniagua, Navarro-Sigüenza, Hernández-Baños, & Morales, 2007; McCormack, Peterson, Bonaccorso, & Smith, 2008; Navarro-Sigüenza, Townsend Peterson, Nyari, García-Deras, & García-Moreno, 2008; Puebla-Olivares et al., 2008). Based on our analyses, it is clear that *C. ravus* is divergent from the remainder of the group. The three currently recognized subspecies of *C. ravus* are allopatric (Campbell & Lamar, 2004), morphologically differentiated (Campbell & Armstrong, 1979), form distinct lineages (Bryson et al., 2011), and are supported as distinct species by our species delimitation models using phylogenomic data. Based on these combined lines of evidence, we recommend that the three subspecies be elevated to full species status: *C. ravus*, *C. brunneus* and *C. exiguus*.

Nearly all species trees place *C. tlaloci* sister to *C. pusillus* and *C. campbelli* sister to *C. armstrongi*. These relationships are consistent with geography, with *C. pusillus* inhabiting the western regions of the Trans-Mexican Volcanic Belt and adjacent Sierra de Coalcomán, and *C. tlaloci* found along the central part of the Trans-Mexican Volcanic Belt. *Crotalus campbelli* is distributed along the western margin of the Trans-Mexican Volcanic Belt and is separated by low-elevation valleys from *C. armstrongi* inhabiting regions of the Trans-Mexican Volcanic Belt to the east (Bryson et al., 2014). None of these species are known to occur sympatrically, although *C. pusillus* and *C. armstrongi* are both found on Cerro Tancítaro in western Michoacán.

Crotalus triseriatus is the sister species to the large clade containing *C. aquilus*, *C. morulus* and *C. lepidus*. *Crotalus aquilus* and southern *C. l. klauberi* (*C. l. klauberi*-1 South) are strongly supported as sister taxa in several of our analyses. Previous results based on mtDNA suggested *C. aquilus* formed a closer relationship with *C. morulus* and placed southern *C. l. klauberi* as sister to the remaining *C. lepidus* (Bryson et al., 2011). Our species delimitation analyses indicate *C. lepidus* contains as many as six distinct lineages, all consistent with geography. These same lineages were inferred with mtDNA (Bryson et al., 2011). Although our sampling spanned the geographical distribution of most subspecies, we lack samples from critical regions where subspecies converge (Campbell & Lamar, 2004). We therefore hold off on formal taxonomic changes to *C. lepidus* until further range-wide sampling combined with a detailed examination of morphology can be made. Additional sampling and integrative approaches to species delimitation will also help alleviate concerns of overestimating the number of true species when using BPP and other multispecies coalescent programmes (Sukumaran & Knowles, 2017).

Our results based on UCE data and results based on mtDNA (Bryson et al., 2011) present evidence for limited hybridization of *C. aquilus* with southern *C. lepidus* and *C. morulus* across the northern region of the Central Mexican Plateau from San Luis Potosí to Jalisco. The patchy islands of suitable rocky habitat above 2,000 m across this region, similar to "sky islands" in northern Mexico and Arizona (McCormack, Huang, & Knowles, 2009), suggest populations in this area likely harbour genetic signatures of a former Pleistocene hybrid zone. Current gene flow at a low rate may also be possible in those few areas with contiguous suitable habitat. Divergence time estimates here and elsewhere (Bryson et al., 2011) indicate *C. aquilus*, southern *C. l. klauberi* and *C. morulus* likely diverged from each other prior to onset of the Pleistocene 2.6 Ma. After these divergences, montane woodlands covered much of the Central Mexican Plateau (McDonald, 1993; Metcalfe, 2006; Gugger, González-Rodríguez, Rodríguez-Correa, Sugita, & Cavender-Bares, 2011; Van Devender, 1990), and as a result, the distributions of these taxa probably extensively overlapped during most of the Pleistocene. However, within the core distributions of *C. aquilus* and *C. morulus*, no evidence of introgression has been uncovered and each species maintains phenotypic cohesion, suggesting these two species have maintained their distinctiveness throughout the past several million years. It is clear, however, that southern *C. l. klauberi* has a closer relationship to *C. aquilus* than to other *C. lepidus*, at least based on genomic and potentially phenotypic data. While mtDNA clearly supports the distinctiveness of southern *C. l. klauberi*, consistent with our species delimitation models, it unites this lineage with other *C. lepidus* (Bryson et al., 2011). This finding suggests that a history of introgression between southern *C. l. klauberi* and *C. aquilus* has been maintained in the nuclear genome, but not the mitochondrial genome. Finer-scale population-level sampling across the northern Central Mexican Plateau is needed to fully characterize potential hybrid-zone dynamics within the species group and help determine if southern *C. l. klauberi* deserves recognition as a distinct species.

ACKNOWLEDGEMENTS

We thank the following curators, institutions and individuals for providing tissue samples: R.W. Murphy and A. Lathrop (Royal Ontario Museum), O. Flores-Villela and A. Nieto-Montes (Universidad Nacional Autónoma de México), J.A. Campbell, C. Franklin and E.N. Smith (University of Texas at Arlington), J. Alvarado-Díaz (Universidad Michoacana de San Nicolás de Hidalgo, Michoacán), A. Kardon (San Antonio Zoo), U.O. García-Vázquez, J. Lemos-Espinal, L. Canseco-Marquez, G. Swinford, E. Mociño-Deloya, K. Setser and G. Quintero-Díaz. Funding was provided in part by NSF (DEB-1257785 and DEB-1258205). Collecting was conducted under permits granted by SEMARNAT to RWB, the late F. Mendoza-Quijano, and UNAM. For additional support and advice, we thank members of the McCormack and Klicka Labs, J. Jones, C. Grunwald, and the many volunteers that helped with field work. The authors are grateful to the Center for Theoretical Physics of New York City College of Technology (CUNY) for providing computational resources. Finally, we thank A. Stamatakis and A. Aberer for assistance with ExaBayes and C. Ané for assistance with SNaQ.

AUTHOR CONTRIBUTIONS

C.B., R.W.B., J.K. and J.E.M. developed the conceptual framework for the project; R.W.B. and D.L. performed fieldwork and contributed samples; R.W.B. generated the data; C.B., R.W.B. and C.W.L. analysed the data; C.B. and R.W.B. led the writing, and all authors contributed to and approved the final manuscript.

DATA ACCESSIBILITY

Alignment and tree files are available from the Dryad Digital Repository (https://doi.org/10.5061/dryad.4467407).

ORCID

Christopher Blair https://orcid.org/0000-0002-4782-6788

REFERENCES

Aberer, A. J., Kobert, K., & Stamatakis, A. (2014). ExaBayes: Massively parallel Bayesian tree inference for the whole-genome era. *Molecular Biology and Evolution*, *31*, 2553–2556.

Alencar, L. R. V., Quental, T. B., Grazziotin, F. G., Alfaro, M. L., Martins, M., Venzon, M., & Zaher, H. (2016). Diversification in vipers: Phylogenetic relationships, time of divergence and shifts in speciation rates. *Molecular Phylogenetics and Evolution*, *105*, 50–62.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*, e3376.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S., & Haussler, D. (2004). Ultraconserved elements in the human genome. *Science*, *304*, 1321–1325.

Blair, C., Mendez de la Cruz, F. R., Law, C., & Murphy, R. W. (2015). Molecular phylogenetics and species delimitation of leaf-toed geckos (Phyllodactylidae: *Phyllodactylus*) throughout the Mexican tropical dry forest. *Molecular Phylogenetics and Evolution*, *84*, 254–265.

Blair, C., & Sánchez-Ramírez, S. (2016). Diversity-dependent cladogenesis throughout western Mexico: Evolutionary biogeography of rattlesnakes (Viperidae: Crotalinae: *Crotalus* and *Sistrurus*). *Molecular Phylogenetics and Evolution*, *97*, 145–154.

Brattstrom, B. H. (1964). Evolution of the pit vipers. *Transactions of the San Diego Society of Natural History*, *13*, 185–268.

Bryson, R. W., García-Vázquez, U. O., & Riddle, B. R. (2012). Diversification in the Mexican horned lizard *Phrynosoma orbiculare* across a dynamic landscape. *Molecular Phylogenetics and Evolution*, *62*, 87–96.

Bryson, R. W., Linkem, C. W., Dorcas, M. E., Lathrop, A., Jones, J. M., Alvarado-Díaz, J., … Murphy, R. W. (2014). Multilocus species delimitation in the *Crotalus triseriatus* species group (Serpentes: Viperidae: Crotalinae), with the description of two new species. *Zootaxa*, *3826*, 475–496.

Bryson, R. W., Linkem, C. W., Pavón-Vázquez, C. J., Nieto-Montes de Oca, A., Klicka, J., & McCormack, J. E. (2017). A phylogenomic perspective on the biogeography of skinks in the *Plestiodon brevirostris* group inferred from target enrichment of ultraconserved elements. *Journal of Biogeography*, *44*, 2033–2044.

Bryson, R. W., Murphy, R. W., Lathrop, A., & Lazcano-Villareal, D. (2011). Evolutionary drivers of phylogeographical diversity in the highlands

of Mexico: A case study of the *Crotalus triseriatus* species group of montane rattlesnakes. *Journal of Biogeography*, *38*, 697–710.

Burleigh, J. G., Kimball, R. T., & Braun, E. L. (2015). Building the avian tree of life using a large-scale, sparse supermatrix. *Molecular Phylogenetics and Evolution*, *84*, 53–63.

Campbell, J. A., & Armstrong, B. L. (1979). Geographic variation in the Mexican Pygmy Rattlesnake, *Sistrurus ravus*, with the description of a new subspecies. *Herpetologica*, *35*, 304–317.

Campbell, J. A., & Lamar, W. W. (2004). *The venomous reptiles of the western hemisphere*. Ithaca, NY:Comstock Pub. Associates Ithaca.

Castoe, T. A., & Parkinson, C. L. (2006). Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). *Molecular Phylogenetics and Evolution*, *39*, 91–110.

Caviedes-Solis, I. W., Bouzid, N. M., Banbury, B. L., & Leaché, A. D. (2015). Uprooting phylogenetic uncertainty in coalescent species delimitation: A meta-analysis of empirical studies. *Current Zoology*, *61*, 866–873.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*, 3317–3324.

Chifman, J., & Kubatko, L. (2015). Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of Theoretical Biology*, *374*, 35–47.

Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., & Warnow, T. (2015). A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics*, *16*, S2.

Crawford, N. G., Faircloth, B. C., McCormack, J. E., Brumfield, R. T., Winker, K., & Glenn, T. C. (2012). More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. *Biology Letters*, *8*, 783–786.

Degnan, J. H., & Rosenberg, N. A. (2006). Discordance of species trees with their most likely gene trees. *PLoS Genetics*, *2*, e68.

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*, 332–340.

Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution International Journal of Organic Evolution*, *63*, 1–19.

Edwards, S. V., Xi, Z., Janke, A., Faircloth, B. C., McCormack, J. E., Glenn, T. C., … Davis, C. C. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, *94*, 447–462.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, *6*, e19379.

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788.

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*, 717–726.

Faircloth, B. C., Sorenson, L., Santini, F., & Alfaro, M. E. (2013). A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). *PLoS One*, *8*, e65923.

Foote, R., & MacMahon, J. A. (1977). Electrophoretic studies of rattlesnake (*Crotalus* & *Sistrurus*) venom: Taxonomic implications. *Comparative Biochemistry and Physiology B*, *57*, 235–241.

Giarla, T. C., & Esselstyn, J. A. (2015). The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. *Systematic Biology*, *64*, 727–740.

Gloyd, H. K. (1940). The rattlesnakes, genera *Sistrurus* and *Crotalus*.: A study in zoogeography and evolution. *Chicago*. Retrieved from https://catalog.hathitrust.org/Record/001501205

Grummer, J. A., Bryson, R. W., & Reeder, T. W. (2014). Species delimitation using Bayes factors: Simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology*, *63*, 119–133.

Gugger, P. F., González-Rodríguez, A., Rodríguez-Correa, H., Sugita, S., & Cavender-Bares, J. (2011). Southward Pleistocene migration of Douglas-fir into Mexico: Phylogeography, ecological niche modeling, and conservation of 'rear edge' populations. *New Phytologist*, *189*, 1185–1199.

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*, 570–580.

Hosner, P. A., Faircloth, B. C., Glenn, T. C., Braun, E. L., & Kimball, R. T. (2016). Avoiding missing data biases in phylogenomic inference: An empirical study in the landfowl (Aves: Galliformes). *Molecular Biology and Evolution*, *33*, 1110–1125.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., … Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, *346*, 1320–1331.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649.

Klauber, L. M. (1956). *Rattlesnakes: Their Habits, Life Histories, and Influence on Mankind*. Berkeley: University of California Press.

Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, *56*, 17–24.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*, 82.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*, 772–773.

Leaché, A. D., Banbury, B. L., Linkem, C. W., & de Oca, A. N.-M. (2016). Phylogenomics of a rapid radiation: Is chromosomal evolution linked to increased diversification in North American spiny lizards (Genus *Sceloporus*)? *BMC Evolutionary Biology*, *16*, 63.

Leaché, A. D., Chavez, A. S., Jones, L. N., Grummer, J. A., Gottscho, A. D., & Linkem, C. W. (2015). Phylogenomics of phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biology and Evolution*, *7*, 706–719.

Leaché, A. D., & Fujita, M. K. (2010). Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B-Biological Sciences*, *277*, 3071–3077.

Leaché, A. D., & Linkem, C. W. (2015). Phylogenomics of horned lizards (Genus: *Phrynosoma*) using targeted sequence capture data. *Copeia*, *103*, 586–594.

Leaché, A. D., & McGuire, J. A. (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.

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, *61*, 727–744.

León-Paniagua, L., Navarro-Sigüenza, A. G., Hernández-Baños, B. E., & Morales, J. C. (2007). Diversification of the arboreal mice of the genus *Habromys* (Rodentia: Cricetidae: Neotominae) in the Mesoamerican highlands. *Molecular Phylogenetics and Evolution*, *42*, 653–664.

Liu, L., Yu, L., & Edwards, S. V. (2010). A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evolutionary Biology*, *10*, 302.

Liu, L., Yu, L., Kubatko, L., Pearl, D. K., & Edwards, S. V. (2009). Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics and Evolution*, *53*, 320–328.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, *46*, 523–536.

McCormack, J. E., & Faircloth, B. C. (2013). Next-generation phylogenetics takes root. *Molecular Ecology*, *22*, 19–21.

McCormack, J. E., Faircloth, B. C., Crawford, N. G., Gowaty, P. A., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species tree analysis. *Genome Research*, *22*, 746–754.

McCormack, J. E., Harvey, M. G., Faircloth, B. C., Crawford, N. G., Glenn, T. C., & Brumfield, R. T. (2013). A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS One*, *8*, e54848.

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, *66*, 526–538.

McCormack, J. E., Peterson, A. T., Bonaccorso, E., & Smith, T. B. (2008). Speciation in the highlands of Mexico: Genetic and phenotypic divergence in the Mexican jay (*Aphelocoma ultramarina*). *Molecular Ecology*, *17*, 2505–2521.

McCormack, J. E., Huang, H., & Knowles, L. L. (2009). Sky islands. In R. G. Gillespie, & D. Clague (Eds.), *Encyclopedia of Islands* (pp. 839–843). Berkeley, CA: University of California Press.

McDonald, J. A. (1993). Phytogeography and history of the alpine–subalpine flora of northeastern Mexico. In T. P. Ramamoorthy, R. Bye, A. Lot, & J. Fa (Ed.), *Biological diversity in Mexico: Origins and distribution* (pp. 681–703). New York: Oxford University Press.

Meiklejohn, K. A., Faircloth, B. C., Glenn, T. C., Kimball, R. T., & Braun, E. L. (2016). Analysis of a rapid evolutionary radiation using ultraconserved elements: Evidence for a bias in some multispecies coalescent methods. *Systematic Biology*, *65*, 612–627.

Metcalfe, S. E. (2006). Late Quaternary environments of the northern deserts and central Transvolcanic Belt of Mexico. *Annals of the Missouri Botanical Garden*, *93*, 258–273.

Mirarab, S., Bayzid, M. S., & Warnow, T. (2016). Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Systematic Biology*, *65*, 366–380.

Mirarab, S., Reaz, R., Bayzid, M. S., Zimmermann, T., Swenson, M. S., & Warnow, T. (2014). ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics*, *30*, i541–548.

Mirarab, S., & Warnow, T. (2015). ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, *31*, i44–i52.

Mittermeier, R. A., Gil, P. G., Hoffman, M., Pilgrim, J., Brooks, T. M., Mittermeier, C. G., … Da Fonseca, G.A.B. (2017). Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions. Retrieved from https://wedocs.unep.org/handle/20.500.11822/15160

Murphy, R. W., Fu, J., Lathrop, A., Feltham, J. V., & Kovac, V. (2002). Phylogeny of the rattlesnakes (Crotalus and Sistrurus) inferred from sequences of five mitochondrial DNA genes. In Gordon W. Schuett, Matts Hoggren, & Michael E. Douglas (Eds.), *Biology of the Vipers*, (pp. 69–92). Eagle Mountain, UT: Eagle Mountain Publishing LC.

Navarro-Sigüenza, A. G., Townsend Peterson, A., Nyari, A., García-Deras, G. M., & García-Moreno, J. (2008). Phylogeography of the Buarremon brush-finch complex (Aves, *Emberizidae*) in Mesoamerica. *Molecular Phylogenetics and Evolution*, *47*, 21–35.

Ogilvie, H. A., Bouckaert, R. R., & Drummond, A. J. (2017). StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution*, *34*, 2101–2114.

Parkinson, C. L. (1999). Molecular systematics and biogeographical history of pitvipers as determined by mitochondrial ribosomal DNA sequences. *Copeia*, *3*, 576–586.

Parkinson, C. L., Campbell, J. A., Chippindale, P. T., & Schuett, G. (2002). Multigene phylogenetic analysis of pitvipers, with comments on their biogeography. *Biology of the Vipers*, *9*, 3–110.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*, e37135.

Puebla-Olivares, F., Bonaccorso, E., De Los Monteros, A. E., Omland, K. E., Llorente-Bousquets, J. E., Peterson, A. T., & Navarro-Sigüenza, A. G. (2008). Speciation in the Emerald Toucanet (*Aulacorhynchus prasinus*) Complex. *The Auk*, *125*, 39–50.

Rambaut, A., & Drummond, A. (2007). Tracer v1. 4.

Rannala, B., & Yang, Z. (2007). Inferring speciation times under an episodic molecular clock. *Systematic Biology*, *56*, 453–466.

Reaz, R., Bayzid, M. S., & Rahman, M. S. (2014). Accurate phylogenetic tree reconstruction from quartets: A heuristic approach. *PLoS One*, *9*, e104008.

dos Reis, M., & Yang, Z. (2011). Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Molecular Biology and Evolution*, *28*, 2161–2172.

Reyes-Velasco, J., Meik, J. M., Smith, E. N., & Castoe, T. A. (2013). Phylogenetic relationships of the enigmatic longtailed rattlesnakes (*Crotalus ericsmithi*, *C. lannomi*, and *C. stejnegeri*). *Molecular Phylogenetics and Evolution*, *69*, 524–534.

Roch, S., & Steel, M. (2014). Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. *Theoretical Population Biology*, *100C*, 56–62.

Roch, S., & Warnow, T. (2015). On the robustness to gene tree estimation error (or lack thereof) of coalescent-based species tree methods. *Systematic Biology*, *64*, 663–676.

Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, *33*, 1654–1668.

Shi, C.-M., & Yang, Z. (2018). Coalescent-based analyses of genomic sequence data provide a robust resolution of phylogenetic relationships among major groups of gibbons. *Molecular Biology and Evolution*, *35*, 159–179.

Smith, B. T., Harvey, M. G., Faircloth, B. C., Glenn, T. C., & Brumfield, R. T. (2013). Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. *Systematic Biology*, *63*, 83–95.

Solís-Lemus, C., & Ané, C. (2016). Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics*, *12*, e1005896.

Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, *34*, 3292–3298.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313.

Streicher, J. W., Schulte, J. A., & Wiens, J. J. (2016). How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. *Systematic Biology*, *65*, 128–145.

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*, 1607–1612.

Swofford, D. L. (2001). PAUP*: Phylogenetic analysis using parsimony (and other methods) 4.0.b5.

Than, C., Ruths, D., & Nakhleh, L. (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, *9*, 322.

Vachaspati, P., & Warnow, T. (2015). ASTRID: Accurate species trees from internode distances. *BMC Genomics*, *16*, S3.

Van Devender, T. R. (1990). *Late Quaternary vegetation and climate of the Chihuahuan Desert, United States and Mexico*. Tucson, AZ: University of Arizona Press.

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., … Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, *22*, 787–798.

Wen, D., Yu, Y., Zhu, J., Nakhleh, L., & Posada, D. (2018). Inferring phylogenetic networks using PhyloNet. *Systematic Biology*, *67*, 735–740.

Xi, Z., Liu, L., & Davis, C. C. (2015). Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. *Molecular Phylogenetics and Evolution*, *92*, 63–71.

Xu, B., & Yang, Z. (2016). Challenges in species tree estimation under the multispecies coalescent model. *Genetics*, *204*, 1353–1368.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*, 1586–1591.

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*, 854–865.

Yang, Z., & Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution*, *23*, 212–226.

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *107*, 9264–9269.

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135.

Yu, Y., & Nakhleh, L. (2015). A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics*, *16*, S10.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, *292*, 686–693.

Zhang, C., Sayyari, E., & Mirarab, S. (2017). ASTRAL-III: Increased scalability and impacts of contracting low support branches. In *Comparative genomics. Lecture notes in computer science* (pp. 53–75). Cham: Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-67979-2_4

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Blair C, Bryson RW Jr, Linkem CW, Lazcano D, Klicka J, McCormack JE. Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: *Crotalus*). *Mol Ecol Resour*. 2018;00:1–17. https://doi.org/10.1111/1755-0998.12970