2018

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
CUNY New York City College of Technology

Robert W. Bryson Jr
University of Washington

Charles W. Linkem
University of Washington

David Lazcano
Universidad Autónoma de Nuevo León

John Klicka
University of Washington

See next page for additional authors

RH: Phylogenomics of montane rattlesnakes

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], and John E. McCormack[4]

[1]Department of Biological Sciences, New York City College of Technology, The City University of New York, 300 Jay Street, Brooklyn, NY, 11201, USA.

[2]Biology PhD Program, CUNY Graduate Center, 365 5th Avenue, New York, NY 10016.

[3]Department of Biology and Burke Museum of Natural History and Culture, University of Washington, 4331 Memorial Way Northeast, Seattle, WA, 98195, USA.

[4]Moore Laboratory of Zoology, Occidental College, Los Angeles, CA, 90041, USA.

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

Correspondence: cblair@citytech.cuny.edu; 718-260-5342 (ph)

Key Words: Bayesian, coalescence, genomics, rattlesnakes, speciation, systematics

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 (UCE) data set (>3000 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 result 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.

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; Liu et al. 2009; McCormack et al. 2013; McCormack and Faircloth 2013; Edwards et al. 2016). 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 et al. 2012; Peterson et al. 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 (Wagner et al. 2013; Giarla and Esselstyn 2015; Leaché et al. 2015, 2016; Meiklejohn et al. 2016). A specific class of molecular marker used with sequence capture, ultraconserved elements (UCEs; Bejerano et al. 2004; Faircloth et al. 2012), has helped resolve longstanding evolutionary questions across taxonomic groups at multiple spatial and temporal scales (Crawford et al. 2012; McCormack et al. 2012; Faircloth et al. 2013; McCormack et al. 2013; Smith et al. 2013; Leaché et al. 2016; Streicher et al. 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 analyze 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 (Kubatko and Degnan 2007; Roch and Steel 2014; Chou et al. 2015; Mirarab et al. 2016). Meanwhile, fully probabilistic

multispecies coalescent methods such as STARBEAST2 (Ogilvie et al. 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 et al. 2010; Vachaspati and Warnow 2015; Zhang et al. 2017). However, poorly resolved gene trees can have a large impact on the species tree (Xi et al. 2015; Meiklejohn et al. 2016; Xu and 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 and Kubatko 2014; Chifman and Kubatko 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 (Gloyd 1940; Klauber 1956; Campbell and Lamar 2004; Blair and Sánchez-Ramírez 2016). 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 and 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 (Gloyd 1940; Klauber 1956; Brattstrom 1964; Foote and MacMahon 1977; Murphy et al. 2002; Bryson et al. 2011; Reyes-Velasco et al. 2013; Blair and Sánchez-Ramírez 2016). Part of the discrepancy seems to result from the fact that many molecular studies rely solely on mtDNA (Parkinson 1999; Murphy et al. 2002; Parkinson et al. 2002; Castoe and Parkinson 2006;

Bryson et al. 2011), which often conflict with phylogenies from other sources of data (Leaché and McGuire 2006; Grummer et al. 2014). 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 et al. 2001) and mountain formation in Mexico (Bryson et al. 2011; Bryson et al. 2014; Blair and Sánchez-Ramírez 2016). 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 southwestern US (Campbell and 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 (Bryson et al. 2011; Blair and Sánchez-Ramírez 2016). 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.

Materials and Methods

Sampling and library preparation

We sampled 54 rattlesnakes of the *C. triseriatus* group from throughout the Mexican highlands (Fig. 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. 2011; Bryson et al. 2014). 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 3000 (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 (doi:10.5061/dryad.4467407).

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 et al. 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 and 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 et al. 2017). Because of the large size of the data set, we used the relaxed hierarchical clustering algorithm to assign sites to different partitions (Lanfear

et al. 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 1000 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.

Coalescent-based species delimitation and species-tree analysis

We first used SVDQUARTETS (Chifman and Kubatko 2014; Chifman and Kubatko 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 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 et al. 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 and Rannala 2010) for joint species delimitation and species-tree inference (i.e., unguided species delimitation; Yang and 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. 2011; Bryson et al. 2014) 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 and 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 northeastern Sierra Madre Occidental and adjacent sky islands), *C. l. klauberi*-4 Northeast (samples from the dry slopes of the northeastern 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 $G(2,1000)$ 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 $D(\alpha) = 10$. Uniform rooted trees was used as the species model prior. Following a burn-in of 10,000 generations, chains were sampled every 5 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 (Leaché and Fujita 2010; Blair et al. 2015; Yang 2015). We specifically compared results between *G*(2,1000) 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*(2,1000) 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, 1000 UCE loci, 2000 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 and 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., Mirarab et al. 2014; Chou et al. 2015). 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 et al. 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 and 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 and Mirarab 2016). The unrooted ASTRAL species tree was subsequently rooted with *A. piscivorus*.

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 (Yang and Rannala 2006; Rannala

and Yang 2007; Yang 2007). 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 and 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 and Yang 2011). Analyses used a HKY+G4 substitution model, default priors for kappa (G[6,2]) and alpha (G[1,1]), G(2,2000) 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 and Sánchez-Ramírez 2016). Specifically, soft bounds between 20-12 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 and 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.

Network estimation

Because strictly bifurcating trees may not adequately depict evolutionary history, we used the pseudolikelihood method SNaQ (Solís-Lemus and Ané 2016) in the PhyloNetworks package (Solís-Lemus et al. 2017) to estimate a phylogenomic network. SNaQ is 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 and 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 versus 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.

Results

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.

Concatenated analysis

Unpartitioned ML analyses of the concatenated matrix under a GTRGAMMA substitution model yielded a topology with strong support for most nodes (Fig. 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.

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 (Fig. 3; Supplementary Table S2). Posterior probability values for each 'species' was 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,

386 particularly for analyses based on 500 loci (Fig. 3). Based on these results, we treated these
387 taxa as distinct lineages for downstream species tree analysis.

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

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

410 The species tree inferred from ASTRAL analysis of 351 gene trees was largely
411 consistent with the trees inferred from concatenated ML and Bayesian analyses and coalescent

analyses in BPP (Fig. 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 posterior probability. In the ASTRAL tree *C. lepidus klauberi*-1 (South) was more closely related to *C. morulus* and other *C. lepidus* lineages versus *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, 1000, and 2000 UCE loci. Discordant results were obtained from each analysis, although most incongruence was restricted to poorly supported nodes (Fig. 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 appears to be additional nuances of the data other than size that influence node support.

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 (Fig. 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.

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 (Fig. 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 ($\gamma = 0.194$) to *C. armstrongi*.

Discussion

Phylogenomic analysis

One of our main objectives was to examine levels of discordance in commonly used phylogenomic packages on a large empirical data set. With the realization that gene-tree topologies do not always agree with the species tree due to multiple processes including incomplete lineage sorting (Maddison 1997; Degnan and Rosenberg 2006; Degnan and Rosenberg 2009; Edwards 2009), substantial effort has been placed on the development of sophisticated algorithms to explicitly account for gene-tree discordance when estimating species trees (e.g., Liu et al. 2009, 2010; Heled and Drummond 2010; Mirarab et al. 2014; Vachaspati and 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),

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 (Yang and Rannala 2010; Chifman and Kubatko 2014; Mirarab et al. 2014; Chou et al. 2015; Mirarab and Warnow 2015; Yang 2015). 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 (Fig. 2) is identical to the topology inferred by fully probabilistic coalescent analyses in BPP (Fig. 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 et al. 2015; Shi and 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., Leaché and Fujita 2010; Blair et al. 2015). 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. 2009; Liu et al. 2010; Mirarab et al. 2014; Mirarab and 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 (Mirarab et al. 2014; Chou et al. 2015). 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., Xi et al. 2015; Hosner et al. 2016; Meiklejohn et al. 2016) and limited our ASTRAL analysis to 351 of the more variable UCE loci with little fragmentary data and use RAxML for gene tree 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. 2014, 2016; Roch and 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 et al. 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 (Xi et al. 2015; Meiklejohn et al. 2016). 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 versus 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 and Kubatko 2014; Chou et al. 2015; Leaché et al. 2015; Leaché et al. 2016). These attributes make the methods particularly appealing to analyze 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 is still little data available illustrating how SVDQUARTETS performs on empirical phylogenomic data (Leaché et al. 2015; Hosner et al. 2016; Leaché et al. 2016). 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*.

542 To further probe how results from SVDQUARTETS may change based on data set size
543 and characteristics, we performed additional species tree analyses on concatenated alignments
544 consisting of 50, 100, 1000, and 2000 UCE loci. Different topologies were recovered for each
545 analysis, although bootstrap support at conflicting nodes was relatively weak. With only 50 UCE
546 loci, the *C. ravus* clade was placed sister to *C. campbelli* and *C. armstrongi* with a bootstrap
547 value of 65%. This result disappeared in all analyses with more data. Our results also
548 suggested an unclear relationship between data set size and bootstrap support values,
549 indicating that there are additional nuances besides number of loci that contribute to strongly
550 supported species trees. In general, our results are consistent with other recent studies that
551 found lower node support values for SVDQUARTETS trees versus trees inferred through
552 concatenation (Leaché and Linkem 2015; Leaché et al. 2015, 2016; Hosner et al. 2016). In
553 addition, our results are similar to a recent study that examined phylogenomic relationships
554 among gibbons using real and simulated data and found that species trees inferred by BPP and
555 ASTRAL were much more consistent than those inferred by SVDQUARTETS (Shi and Yang
556 2018). Although additional simulation and empirical studies are needed to examine
557 characteristics of the SVDQUARTETS method in greater detail, it appears that relatively large data
558 sets are needed to reach a topology and support values consistent with other methods of
559 species tree inference, and that full likelihood-based approaches that adequately model rate
560 heterogeneity and coalescent stochasticity may be preferable.
561 Strictly bifurcating trees may not adequately depict evolutionary history when gene flow
562 and hybridization are prevalent (Yu and Nakhleh 2015; Solís-Lemus and Ané 2016; Wen et al.
563 2018). Using SNaQ, we found hybridization within the sampled taxa of the *C. triseriatus* group to
564 be relatively rare. Results provided strong support for a model consisting of a single reticulation
565 event from the mrca of *C. lepidus* + *C. morulus* into *C. armstrongi*. The major tree inferred by
566 SNaQ was virtually identical to the ASTRAL tree and very similar to the trees inferred by BPP
567 and concatenation. Interestingly, SNaQ detected no introgression between *C. aquilus* and *C.*

568 *lepidus klauberi*-1 (South) even though this was suggested in our other phylogenetic analyses. It
569 is possible that introgression between these two taxa was limited to two individuals in our study
570 (Ct143ClkJAL and Ct2CaqAGS), and the inclusion of larger number of samples from each
571 species masked the signal of hybridization. Additional sampling of both species from near where
572 their ranges meet should further elucidate the propensity for introgression in these snakes.
573 Interest in phylogenomic networks continues to increase as more evidence accumulates
574 suggesting that both incomplete lineage sorting and reticulation/hybridization may be a common
575 component of the evolutionary history of several groups (Solís-Lemus and Ané 2016; Solís-
576 Lemus et al. 2017; Wen et al. 2018). Unfortunately, many of these explicit network methods are
577 computationally demanding (Than et al. 2008), necessitating the use of some type of
578 approximation. New pseudolikelihood methods such as SNaQ and similar techniques in the
579 PhyloNet package (Yu and Nakhleh 2015; Wen et al. 2018) show promise for analyzing large
580 phylogenomic data sets. Combining tree-based and network-based phylogenomic methods,
581 each with their own strengths and weaknesses, will likely provide the most robust conclusions
582 regarding evolutionary history. Further, although our inferred divergence times are broadly
583 concordant with previous studies (e.g., Bryson et al. 2011; Blair and Sánchez-Ramírez 2016),
584 more research is needed to determine how reticulated histories influence inferred divergence
585 times.
586
587 *Systematics and biogeography of the* C. triseriatus *group*
588 The Mexican highlands are a known biodiversity hotspot (Mittermeier et al. 2017), and multiple
589 phylogeographic studies have documented cryptic diversity distributed throughout high-
590 elevation biomes within Mexico's major mountain systems (e.g., León-Paniagua et al. 2007;
591 McCormack et al. 2008; Navarro-Sigüenza et al. 2008; Puebla-Olivares et al. 2008; Bryson et
592 al. 2012). Based on our analyses, it is clear that *C. ravus* is divergent from the remainder of the
593 group. The three currently recognized subspecies of *C. ravus* are allopatric (Campbell and

Lamar 2004), morphologically differentiated (Campbell and 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 geographic distribution of most subspecies, we lack samples from critical regions where subspecies converge (Campbell and 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 programs (Sukumaran and 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 2000 m across this region, similar to 'sky islands' in northern Mexico and Arizona (McCormack et al. 2009), suggest populations in this area likely harbor 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 (Van Devender, 1990; McDonald, 1993; Metcalfe, 2006; Gugger et al. 2011), 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.

References

Aberer AJ, Kobert K, Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31:2553–2556.

Alencar LRV, Quental TB, Grazziotin FG, Alfaro ML, Martins M, Venzon M, Zaher H. 2016. Diversification in vipers: phylogenetic relationships, time of divergence and shifts in speciation rates. Mol. Phylogenet. Evol. 105:50–62.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3:e3376.

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. 2004. Ultraconserved elements in the human genome. Science 304:1321–1325.

674 Blair C, Mendez de la Cruz FR, Law C, Murphy RW. 2015. Molecular phylogenetics and species
675 delimitation of leaf-toed geckos (Phyllodactylidae: *Phyllodactylus*) throughout the
676 Mexican tropical dry forest. Mol Phylogenet Evol 84:254–265.

677 Blair C, Sánchez-Ramírez S. 2016. Diversity-dependent cladogenesis throughout western
678 Mexico: evolutionary biogeography of rattlesnakes (Viperidae: Crotalinae: *Crotalus* and
679 *Sistrurus*). Mol. Phylogenet. Evol. 97:145–154.

680 Brattstrom BH. 1964. Evolution of the pit vipers. Smith II.

681 Bryson RW, García-Vázquez UO, Riddle BR. 2012. Diversification in the Mexican horned lizard
682 *Phrynosoma orbiculare* across a dynamic landscape. Mol. Phylogenet. Evol. 62:87–96.

683 Bryson RW, Linkem CW, Dorcas ME, Lathrop A, Jones JM, Alvarado-Díaz J, Grünwald CI,
684 Murphy RW. 2014. Multilocus species delimitation in the *Crotalus triseriatus* species
685 group (Serpentes: Viperidae: Crotalinae), with the description of two new species.
686 Zootaxa 3826:475–496.

687 Bryson RW, Linkem CW, Pavón-Vázquez CJ, Nieto-Montes de Oca A, Klicka J, McCormack JE.
688 2017. A phylogenomic perspective on the biogeography of skinks in the *Plestiodon*
689 *brevirostris* group inferred from target enrichment of ultraconserved elements. J.
690 Biogeogr. 44:2033–2044.

691 Bryson RW, Murphy RW, Lathrop A, Lazcano-Villareal D. 2011. Evolutionary drivers of
692 phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus*
693 *triseriatus* species group of montane rattlesnakes. J. Biogeogr. 38:697–710.

694 Burleigh JG, Kimball RT, Braun EL. 2015. Building the avian tree of life using a large-scale,
695 sparse supermatrix. Mol. Phylogenet. Evol. 84:53–63.

696 Campbell JA, Armstrong BL. 1979. Geographic variation in the Mexican Pygmy Rattlesnake,
697 *Sistrurus ravus*, with the description of a new subspecies. Herpetologica 35:304–317.

698 Campbell JA, Lamar WW. 2004. The venomous reptiles of the western hemisphere. Comstock
699 Pub. Associates Ithaca.

700 Castoe TA, Parkinson CL. 2006. Bayesian mixed models and the phylogeny of pitvipers
701 (Viperidae: Serpentes). Mol. Phylogenet. Evol. 39:91–110.

702 Caviedes-Solis IW, Bouzid NM, Banbury BL, Leaché AD. 2015. Uprooting phylogenetic
703 uncertainty in coalescent species delimitation: a meta-analysis of empirical studies. Curr.
704 Zool. 61:866–873.

705 Chifman J, Kubatko L. 2014. Quartet inference from SNP data under the coalescent model.
706 Bioinformatics 30:3317–3324.

707 Chifman J, Kubatko L. 2015. Identifiability of the unrooted species tree topology under the
708 coalescent model with time-reversible substitution processes, site-specific rate variation,
709 and invariable sites. J. Theor. Biol. 374:35–47.

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

713 Crawford NG, Faircloth BC, McCormack JE, Brumfield RT, Winker K, Glenn TC. 2012. More
714 than 1000 ultraconserved elements provide evidence that turtles are the sister group of
715 archosaurs. Biol. Lett. 8:783–786.

716 Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene
717 trees. PLoS Genet. 2:e68.

718 Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the
719 multispecies coalescent. Trends Ecol. Evol. 24:332–340.

720 Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? Evol. Int.
721 J. Org. Evol. 63:1–19.

722 Edwards SV, Xi Z, Janke A, Faircloth BC, McCormack JE, Glenn TC, Zhong B, Wu S, Lemmon
723 EM, Lemmon AR, et al. 2016. Implementing and testing the multispecies coalescent
724 model: a valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. 94:447–462.

725 Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A
726 robust, simple genotyping-by-sequencing (GBS) approach for high diversity species.
727 PloS One 6:e19379.

728 Faircloth BC. 2016. PHYLUCE is a software package for the analysis of conserved genomic
729 loci. Bioinforma. Oxf. Engl. 32:786–788.

730 Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012.
731 Ultraconserved elements anchor thousands of genetic markers spanning multiple
732 evolutionary timescales. Syst. Biol. 61:717–726.

733 Faircloth BC, Sorenson L, Santini F, Alfaro ME. 2013. A phylogenomic perspective on the
734 radiation of ray-finned fishes based upon targeted sequencing of ultraconserved
735 elements (UCEs). PLoS One 8:e65923.

736 Foote R, MacMahon JA. 1977. Electrophoretic studies of rattlesnake (*Crotalus* & *Sistrurus*)
737 venom: taxonomic implications. Comp. Biochem. Physiol. Part B Comp. Biochem.
738 57:235–241.

739 Giarla TC, Esselstyn JA. 2015. The challenges of resolving a rapid, recent radiation: empirical
740 and simulated phylogenomics of Philippine shrews. Syst. Biol. 64:727–740.

741 Gloyd HK. 1940. The rattlesnakes, genera *Sistrurus* and *Crotalus*.: a study in zoogeography and
742 evolution. Chicago Available from: https://catalog.hathitrust.org/Record/001501205.

743 Grummer JA, Bryson RW, Reeder TW. 2014. Species delimitation using Bayes factors:
744 simulations and application to the *Sceloporus scalaris* species group (Squamata:
745 Phrynosomatidae). Syst. Biol. 63:119–133.

746 Gugger PF, González-Rodríguez A, Rodríguez-Correa H, Sugita S, Cavender-Bares J (2011)
747 Southward Pleistocene migration of Douglas-fir into Mexico: phylogeography, ecological
748 niche modeling, and conservation of 'rear edge' populations. New Phytol. 189:1185–
749 1199.

750 Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. Mol.
751 Biol. Evol. 27:570–580.

752 Hosner PA, Faircloth BC, Glenn TC, Braun EL, Kimball RT. 2016. Avoiding missing data biases
753 in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol.
754 Biol. Evol. 33:1110–1125.

755 Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B,
756 Howard JT, et al. 2014. Whole-genome analyses resolve early branches in the tree of
757 life of modern birds. Science 346:1320–1331.

758 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
759 Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable
760 desktop software platform for the organization and analysis of sequence data.
761 Bioinformatics 28:1647–1649.

762 Klauber LM. 1956. Rattlesnakes. Univ of California Press

763 Kubatko LS, Degnan JH. 2007. Inconsistency of phylogenetic estimates from concatenated data
764 under coalescence. Syst Biol 56:17–24.

765 Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014. Selecting optimal partitioning
766 schemes for phylogenomic datasets. BMC Evol. Biol. 14:82.

767 Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new
768 methods for selecting partitioned models of evolution for molecular and morphological
769 phylogenetic analyses. Mol. Biol. Evol. 34:772–773.

770 Leaché AD, Banbury BL, Linkem CW, de Oca AN-M. 2016. Phylogenomics of a rapid radiation:
771 is chromosomal evolution linked to increased diversification in North American spiny
772 lizards (Genus *Sceloporus*)? BMC Evol. Biol. 16:63.

773 Leaché AD, Chavez AS, Jones LN, Grummer JA, Gottscho AD, Linkem CW. 2015.
774 Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture
775 versus restriction site associated DNA sequencing. Genome Biol. Evol. 7:706–719.

776 Leaché AD, Fujita MK. 2010. Bayesian species delimitation in West African forest geckos
777 (*Hemidactylus fasciatus*). Proc. R. Soc. B Biol. Sci. 277:3071–3077.

778 Leaché AD, Linkem CW. 2015. Phylogenomics of horned lizards (Genus: *Phrynosoma*) using
779 targeted sequence capture data. Copeia 103:586–594.

780 Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*)
781 based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene
782 tree. Mol. Phylogenet. Evol. 39:628–644.

783 Lemmon AR, Emme SA, Lemmon EM. 2012. Anchored hybrid enrichment for massively high-
784 throughput phylogenomics. Syst. Biol. 61:727–744.

785 León-Paniagua L, Navarro-Sigüenza AG, Hernández-Baños BE, Morales JC. 2007.
786 Diversification of the arboreal mice of the genus *Habromys* (Rodentia: Cricetidae:
787 Neotominae) in the Mesoamerican highlands. Mol. Phylogenet. Evol. 42:653–664.

788 Liu L, Yu L, Edwards SV. 2010. A maximum pseudo-likelihood approach for estimating species
789 trees under the coalescent model. BMC Evol. Biol. 10:302.

790 Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV. 2009. Coalescent methods for estimating
791 phylogenetic trees. Mol. Phylogenet. Evol. 53:320–328.

792 Maddison WP. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

793 McCormack JE, Faircloth BC. 2013. Next-generation phylogenetics takes root. Mol. Ecol.
794 22:19–21.

795 McCormack JE, Faircloth BC, Crawford NG, Gowaty PA, Brumfield RT, Glenn TC. 2012.
796 Ultraconserved elements are novel phylogenomic markers that resolve placental
797 mammal phylogeny when combined with species-tree analysis. Genome Res. 22:746–
798 754.

799 McCormack JE, Harvey MG, Faircloth BC, Crawford NG, Glenn TC, Brumfield RT. 2013. A
800 phylogeny of birds based on over 1,500 loci collected by target enrichment and high-
801 throughput sequencing. PLoS One 8:e54848.

802 McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-
803 generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol.
804 66:526–538.

805 McCormack JE, Peterson AT, Bonaccorso E, Smith TB. 2008. Speciation in the highlands of
806 Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma*
807 *ultramarina*). Mol. Ecol. 17:2505–2521.

808 McCormack JE, Huang H, Knowles LL. 2009. Sky islands. Encyclopedia of Islands (ed.
809 Gillespie R.G., Clague D.) pp. 839–843. University of California Press, Berkeley, CA.

810 McDonald JA. 1993. Phytogeography and history of the alpine–subalpine flora of northeastern
811 Mexico. Biological diversity in Mexico: origins and distribution (ed. by T.P.
812 Ramamoorthy, R. Bye, A. Lot and J. Fa,), pp. 681–703. Oxford University Press, New
813 York.

814 Meiklejohn KA, Faircloth BC, Glenn TC, Kimball RT, Braun EL. 2016. Analysis of a rapid
815 evolutionary radiation using ultraconserved elements: evidence for a bias in some
816 multispecies coalescent methods. Syst. Biol. 65:612–627.

817 Metcalfe SE. 2006. Late Quaternary environments of the northern deserts and central
818 Transvolcanic Belt of Mexico. Ann. Mo. Bot. Gard. 93:258–273.

819 Mirarab S, Bayzid MS, Warnow T. 2016. Evaluating summary methods for multilocus species
820 tree estimation in the presence of incomplete lineage sorting. Syst. Biol. 65:366–380.

821 Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014. ASTRAL:
822 genome-scale coalescent-based species tree estimation. Bioinforma. Oxf. Engl. 30:i541-
823 548.

824 Mirarab S, Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many
825 hundreds of taxa and thousands of genes. Bioinforma. Oxf. Engl. 31:i44-52.

826 Mittermeier RA, Gil PG, Hoffman M, Pilgrim J, Brooks TM, Mittermeier CG, Lamoreux J, Da
827 Fonseca G a. B. 2017. Hotspots Revisited: Earth's Biologically Richest and Most
828 Endangered Terrestrial Ecoregions. Available from:
829 https://wedocs.unep.org/handle/20.500.11822/15160

830 Murphy RW, Fu J, Lathrop A, Feltham JV, Kovac V. 2002. Phylogeny of the rattlesnakes
831 (*Crotalus* and *Sistrurus*) inferred from sequences of five mitochondrial DNA genes. Biol.
832 Vipers:69–92.

833 Navarro-Sigüenza AG, Townsend Peterson A, Nyari A, García-Deras GM, García-Moreno J.
834 2008. Phylogeography of the Buarremon brush-finch complex (Aves, *Emberizidae*) in
835 Mesoamerica. Mol. Phylogenet. Evol. 47:21–35.

836 Ogilvie HA, Bouckaert RR, Drummond AJ. 2017. StarBEAST2 brings faster species tree
837 inference and accurate estimates of substitution rates. Mol. Biol. Evol. 34:2101–2114.

838 Parkinson CL. 1999. Molecular systematics and biogeographical history of pitvipers as
839 determined by mitochondrial ribosomal DNA sequences. Copeia:576–586.

840 Parkinson CL, Campbell JA, Chippindale PT, Schuett G. 2002. Multigene phylogenetic analysis
841 of pitvipers, with comments on their biogeography. Biol. Vipers 9:3–110.

842 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an
843 inexpensive method for de novo SNP discovery and genotyping in model and non-model
844 species. PloS One 7:e37135.

845 Puebla-Olivares F, Bonaccorso E, De Los Monteros AE, Omland KE, Llorente-Bousquets JE,
846 Peterson AT, Navarro-Sigüenza AG. 2008. Speciation in the Emerald Toucanet
847 (*Aulacorhynchus prasinus*) Complex. The Auk 125:39–50.

848 Rambaut A, Drummond A. 2007. Tracer v1. 4.

849 Rannala B, Yang Z. 2007. Inferring speciation times under an episodic molecular clock. Syst.
850 Biol. 56:453–466.

851 Reaz R, Bayzid MS, Rahman MS. 2014. Accurate phylogenetic tree reconstruction from
852 quartets: a heuristic approach. PLOS ONE 9:e104008.

853 Reis M dos, Yang Z. 2011. Approximate likelihood calculation on a phylogeny for Bayesian
854 estimation of divergence times. Mol. Biol. Evol. 28:2161–2172.

855 Reyes-Velasco J, Meik JM, Smith EN, Castoe TA. 2013. Phylogenetic relationships of the
856 enigmatic longtailed rattlesnakes (*Crotalus ericsmithi*, *C. lannomi*, and *C. stejnegeri*).
857 Mol. Phylogenet. Evol. 69:524–534.

858 Roch S, Steel M. 2014. Likelihood-based tree reconstruction on a concatenation of aligned
859 sequence data sets can be statistically inconsistent. Theor. Popul. Biol. 100C:56–62.

860 Roch S, Warnow T. 2015. On the robustness to gene tree estimation error (or lack thereof) of
861 coalescent-based species tree methods. Syst. Biol. 64:663–676.

862 Sayyari E, Mirarab S. 2016. Fast coalescent-based computation of local branch support from
863 quartet frequencies. Mol. Biol. Evol. 33:1654–1668.

864 Shi C-M, Yang Z. 2018. Coalescent-based analyses of genomic sequence data provide a robust
865 resolution of phylogenetic relationships among major groups of gibbons. Mol. Biol. Evol.
866 35:159–179.

867 Smith BT, Harvey MG, Faircloth BC, Glenn TC, Brumfield RT. 2013. Target capture and
868 massively parallel sequencing of ultraconserved elements for comparative studies at
869 shallow evolutionary time scales. Syst. Biol. 63:83–95.

870 Solís-Lemus C, Ané C. 2016. Inferring phylogenetic networks with maximum pseudolikelihood
871 under incomplete lineage sorting. PLOS Genet. 12:e1005896.

872 Solís-Lemus C, Bastide P, Ané C. 2017. PhyloNetworks: a package for phylogenetic networks.
873 Mol. Biol. Evol. 34:3292–3298.

874 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
875 large phylogenies. Bioinformatics 30:1312–1313.

876 Streicher JW, Schulte JA, Wiens JJ. 2016. How should genes and taxa be sampled for
877 phylogenomic analyses with missing data? An empirical study in iguanian lizards. Syst.
878 Biol. 65:128–145.

879 Sukumaran J, Knowles LL. 2017. Multispecies coalescent delimits structure, not species. Proc.
880 Natl. Acad. Sci. 114:1607–1612.

881 Swofford DL. 2001. PAUP*: Phylogenetic Analysis Using Parsimony (and other methods)
882 4.0.b5.

883 Than C, Ruths D, Nakhleh L. 2008. PhyloNet: a software package for analyzing and
884 reconstructing reticulate evolutionary relationships. BMC Bioinformatics 9:322.

885 Vachaspati P, Warnow T. 2015. ASTRID: Accurate Species TRees from Internode Distances.
886 BMC Genomics 16:S3.

887 Van Devender TR. 1990. Late Quaternary vegetation and climate of the Chihuahuan Desert,
888 United States and Mexico. University of Arizona Press Tucson, AZ.

889 Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O.
890 2013. Genome-wide RAD sequence data provide unprecedented resolution of species

boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22:787–798.

Wen D, Yu Y, Zhu J, Nakhleh L, Posada D. 2018. Inferring phylogenetic networks using PhyloNet. Syst. Biol. 67:735–740.

Xi Z, Liu L, Davis CC. 2015. Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. Mol. Phylogenet. Evol. 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. Mol. Biol. Evol. 24:1586–1591.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 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. Mol. Biol. Evol. 23:212–226.

Yang Z, Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. 107:9264–9269.

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 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. Springer, Cham. p. 53–75. Available from: https://link.springer.com/chapter/10.1007/978-3-319-67979-2_4

Data Accessibility

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

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.

931 **Figures and Figure Legends**

932

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

936

937 **Fig. 2.** Phylogenomic relationships of montane rattlesnakes of the *Crotalus triseriatus* group
938 based on unpartitioned concatenated maximum likelihood (ML) analysis of 3384 UCE loci
939 (2,057,530 bp) under a GTRGAMMA model of substitution. Values adjacent to nodes represent
940 rapid bootstrap values from a joint bootstrap + ML search in RAxML (-f a option). Blue boxes
941 indicate geographical lineages within *C. lepidus klauberi*; gray box indicates putative
942 hybridization/introgression between *C. aquilus* and *C. lepidus klauberi*. MRE consensus tree
943 from unpartitioned ExaBayes analyses was identical, with full posterior probability values (1.0)
944 for all nodes.

945

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

950

951

952

Fig. 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.

963 **Fig. 5.** Species trees inferred by SVDQUARTETS based on different subsets of the total data set.
964 Loci were selected randomly and concatenated for each analysis. Individual – species
965 assignments followed analyses on the full data and were based off results from BPP. Trees
966 shown are the SVDQUARTETS trees with bootstrap support values mapped on branches (100
967 replicates). Numbers in bold indicate bootstrap values below 70%. 'X' indicates nodes that were
968 not recovered in the bootstrap consensus tree.
969
970
971

Fig. 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 indicates that chains were run for a sufficient length of time and convergence was reached.

Fig. 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.