Biological Journal of the Linnean Society, 2023, **XX**, 1–20. With 5 figures.

Phylogenomic data reveal cryptic diversity and deep phylogeographical structure within the common chuckwalla, *Sauromalus ater* (Squamata: Iguanidae)

ALEXANDRA SUMARLI[1,4,*], BRADFORD D. HOLLINGSWORTH[2], JORGE H. VALDEZ–VILLAVICENCIO[3] and TOD W. REEDER[4]

[1]*Department of Evolution, Ecology and Organismal Biology, University of California, Riverside, Riverside, CA 92521, USA*
[2]*Department of Herpetology, San Diego Natural History Museum, PO Box 121390, San Diego, CA 92112, USA*
[3]*Conservación de Fauna de Noroeste, Ensenada, Baja California 22785, México*
[4]*Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA*

Received 20 April 2023; revised 6 July 2023; accepted for publication 6 July 2023

Understanding how historical geological processes drive diversification and shape the contemporary distribution of species is fundamental to phylogeography. We take a genomic approach in order to elucidate the deep phylogeographical history and species limits of chuckwallas (*Sauromalus*), a conspicuous group of lizards of the arid lands of south-western North America. Phylogenetic and population genomic analyses of double digest restriction site-associated DNA sequencing data confirm the presence of at least two major lineages, peninsular and continental groups, within the widespread and morphologically variable common chuckwalla (*Sauromalus ater*). These lineages diversified in the vicinity of the head of the Gulf of California in north-eastern Baja California in the early Pliocene to late Miocene, during the formation of the northern gulf. The peninsular lineage of *S. ater* subsequently gave rise to the four insular endemic species of *Sauromalus* associated with the Baja California peninsula. Genomic analyses strongly support the continued recognition of the insular gigantics *Sauromalus varius* and *Sauromalus hispidus* as distinct species, although their relationship as sister species remains unresolved. Weaker phylogenetic signal for the insular species *Sauromalus slevini* and *Sauromalus klauberi* is provided by the genomic data; thus, it is advocated to continue recognizing these species until additional data can be analysed to evaluate their distinctiveness.

ADDITIONAL KEYWORDS: Baja California – biogeography – genomics – phylogeography – reptiles – species delimitation.

INTRODUCTION

A central goal of phylogeographical studies in Baja California and the broader Sonoran Desert has been to link concordant phylogenetic breaks with the timing of known geomorphological processes (Riddle *et al.*, 2000; Dolby *et al.*, 2015). Among taxonomic groups spanning opposite sides of the Gulf of California, speciation has largely been attributed to the formation of the Baja California peninsula and subsequent creation of the modern gulf. This is because of populations with continental origins becoming geographically isolated on the peninsula, resulting in restricted gene flow between populations over time. More than two decades ago, Murphy (1983) and Grismer (1994) presented vicariance-based biogeographical scenarios that have served as a framework to explain the contemporary distribution and phylogenetic relationships of squamate reptiles around the Gulf of California. The first scenario is hypothesized to have occurred during the early formation of the gulf, as the peninsula separated from mainland Mexico ~5.5 Mya, in the late Miocene. This southern gulf vicariance scenario postulates that lineages formed during this time should occur in southern Baja California and have a sister taxon located in mainland Mexico, ranging south of Isla Tiburon. In contrast, the Northern gulf vicariance scenario is hypothesized to have occurred more recently, during the flooding of the northern

*Corresponding author. E-mail: sumarli.alex@gmail.com

Gulf of California ~3 Mya, in the late Miocene–early Pliocene. Lineages that diversified under this scenario should exhibit a circum-gulf distribution and sister-species relationship at the vicinity of the head of the gulf, where they might come into secondary contact. Other biogeographical scenarios, based on climate-driven dispersal (Savage, 1960), overwater dispersal (de Queiroz & Lawson, 2008; Wood *et al.*, 2008) and the presence of transpeninsular seaways across Baja California (Upton & Murphy, 1997), have also been proposed as possible drivers of speciation. Regardless of the wealth of studies in the region, the phylogeographical history of the fauna of Baja California remains largely understudied from a coalescent phylogenetic perspective.

Given the high levels of unique biodiversity among squamate reptiles in the region, including ≥ 40 insular endemic reptile species (Grismer, 2002; Lovich *et al.*, 2009), it is imperative that the species limits and underlying speciation patterns of these taxa are well understood, in order to inform conservation management strategies adequately. Historically, phylogenetic inferences of squamates of Baja California and the gulf islands were based exclusively on morphological characteristics (e.g. Klauber, 1947; Montanucci, 1987; Weins, 1993; Hollingsworth, 1998) or mitochondrial DNA (mtDNA) (e.g. Upton & Murphy, 1997; McGuire *et al.*, 2007; Wood *et al.*, 2008; Davy *et al.*, 2011). In more recent years, studies have incorporated microsatellites (Valdivia-Carrillo *et al.*, 2017) and multilocus nuclear data (Leaché & Mulcahy, 2007; Mulcahy & Macey, 2009; Leavitt *et al.*, 2017, 2020), and an even smaller but growing number of studies have begun to use high-throughput sequencing techniques for genomic perspectives (Gottscho *et al.*, 2017; Harrington *et al.*, 2017; Meik *et al.*, 2018). Genomic studies have also begun to adopt more sophisticated and rigorous analytical approaches under the multispecies coalescent (MSC) model to test and determine species limits (Leaché & Oaks, 2017; Flouri *et al.*, 2018). Although they have limitations (Carstens *et al.*, 2013; Sukumaran & Knowles, 2017; Chambers & Hillis, 2020), species delimitation analyses under the MSC model have several advantages over more traditional single-gene and concatenated phylogenetic analyses. These advantages include estimating species tree topologies while accommodating gene tree heterogeneity and estimating demographic parameters, such as ancestral population sizes and population divergence times, that are important in the speciation process (Fujita *et al.*, 2012; Edwards *et al.*, 2016; Leaché *et al.*, 2019). Taken together, multilocus genome-level data coupled with adequate sampling offer the ability to inform biogeographical hypotheses and interrogate species boundaries previously delimited using morphological or single-locus molecular data.

Here, we sought to examine the phylogeography and early diversification history of the chuckwalla (*Sauromalus*), a geographically widespread, herbivorous member of the Iguanidae. *Sauromalus* is morphologically divergent from other iguanas and adapted to retreating to rock crevices. Within the genus, they are also morphologically diverse in adult male coloration, scalation and body size. They inhabit rocky outcroppings throughout the south-western deserts of the USA, north-western Mexico and islands of the Gulf of California. Four species (*Sauromalus varius*, *Sauromalus hispidus*, *Sauromalus slevini* and *Sauromalus klauberi*) are insular endemics, occurring nowhere else in mainland Mexico or the Baja California peninsula (Hollingsworth, 1998). There have been several hypotheses about the historical biogeography of *Sauromalus* and the evolution of body size between continental and peninsular forms (Murphy, 1983; Grismer, 1994, 2002; Grismer *et al.*, 1995; Petren & Case, 1997, 2002; Hollingsworth, 1998; Welsh, 1988). However, such questions about the biogeography and evolutionary diversification of *Sauromalus* have yet to be investigated with genomic data and recent phylogenetic systematic methods.

Within *Sauromalus*, the greatest uncertainty in species limits lies within the geographically widespread *Sauromalus ater*. The species was once partitioned among three morphologically variable species: *Sauromalus obesus*, *S. ater* and *Sauromalus australis*. *Sauromalus obesus* was further divided into four subspecies (*S. o. obesus*, *S. o. multiforminatus*, *S. o. tumidus* and *S. o. townsendi*) and restricted primarily to the continental USA, Sonora and the northern Baja California peninsula. *Sauromalus ater* (with its two subspecies: *S. a. shawi* and *S. a. ater*) was restricted to the islands of the southern Gulf of California. *Sauromalus australis*, being the least morphologically variable of the three species, was restricted to the middle to southern portion of the Baja California peninsula (Fig. 1). The insular *S. slevini* and *S. klauberi* have been treated differentially as subspecies of *S. ater*, without discussion (Soulé & Sloan, 1966; Robinson, 1974; Case, 1982). In the last taxonomic revision, Hollingsworth (1998) demonstrated that the three species, with their subspecies, exhibited morphological overlap explained by clinal variation, namely from north to south, in scalation traditionally used in *Sauromalus* taxonomy. Based on a lack of fixed diagnosable morphological characters, Hollingsworth (1998) hypothesized that this widespread continental complex represented a single polymorphic species and synonymized *S. obesus* and *S. australis* with *S. ater*, the oldest name available in the genus. For the insular endemics, Hollingsworth (1998) recognized four species: *S. slevini*, *S. klauberi*, *S. varius* and *S. hispidus*.

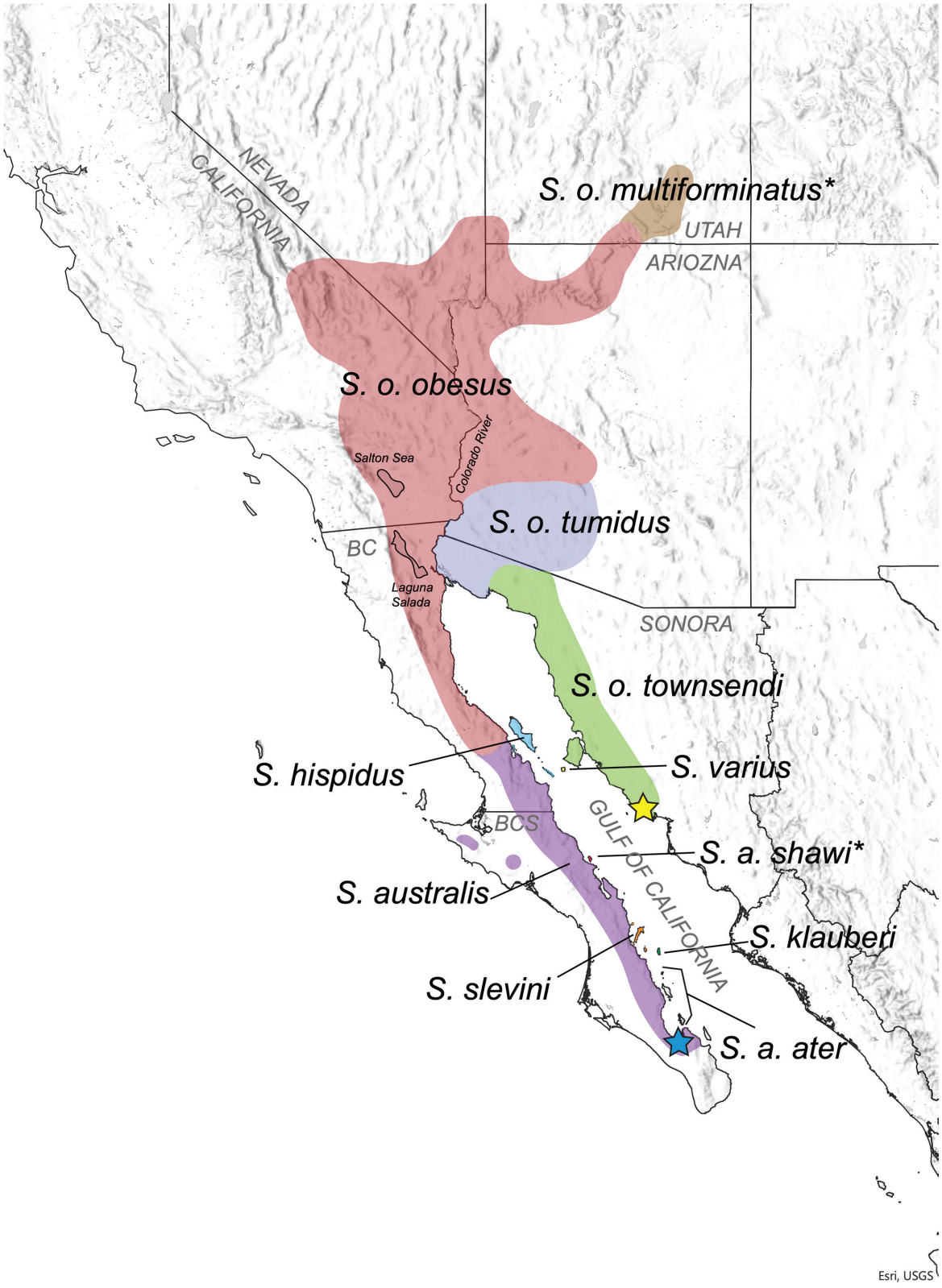

Figure 1. Distribution map of all the named taxa within *Sauromalus* following Shaw (1945), with some modifications (*). A blue star represents the type locality of *Sauromalus interbrachialis* Dickerson (1919); a yellow star represents the restricted type locality of *Sauromalus ater* Duméril (1856).

Although Hollingsworth (1998) recognized a single geographically widespread *S. ater*, phylogenetic results based on morphology (Petren & Case, 1997, 2002) demonstrated substantial cryptic diversity within *S. ater* using mtDNA [cytochrome *b* (*Cytb*)]. Furthermore, the inferred phylogenetic relationships (gene tree) were incongruent with a single *S. ater*. Specifically, the results provided by Petren & Case (1997, 2002) supported four major mtDNA clades that are incongruent with the previous taxonomies of *Sauromalus* based on morphology that fail to explain the geographical morphological variation within *S. ater*; these include: (1) an eastern/northern (EN) clade, including *S. o. tumidus*, *S. o. townsendi* and portions of *S. o. obesus*; (2) a southern peninsular (SP) clade, including *S. australis* and insular *S. ater*; (3) a northern peninsular (NP) clade, including a portion of *S. o. obesus*; and (4) an island gigantics (IG) clade, including *S. hispidus* and *S. varius*. The mtDNA results support the existence of divergent genetic lineages within *S. ater s.l.*, thus leading one to question whether *S. ater* represents a single morphologically variable species. Despite its sparse sampling at crucial contact zones, the mitochondrial study by Petren & Case (2002) remains the most comprehensive molecular systematic study on *Sauromalus* to date.

The specific objectives of this study were to explore the phylogeographical structure and phylogenetic relationships of *Sauromalus* from a genomic perspective. With extensive geographical sampling, we applied various coalescent-based species delimitation approaches to test whether *S. ater* is a single species or represents multiple independent species as suggested by previous mtDNA studies. We included individuals from other currently recognized insular species of *Sauromalus* (*S. varius*, *S. hispidus*, *S. klauberi* and *S. slevini*) and, to a lesser extent, we evaluate their validity as species. Finally, we explore the historical biogeography and early diversification history of major lineages of *Sauromalus* using both concatenated and coalescent-based analyses for divergence dating.

MATERIAL AND METHODS

Taxon sampling

Genomic DNA sequence data were generated from a total of 93 individuals across the geographical range of *S. ater*, and other species of *Sauromalus* (65 *S. ater*, 5 *S. hispidus*, five *S. varius*, 4 *S. klauberi* and 2 *S. slevini*), in addition to two individuals of *Iguana iguana* as an outgroup taxon (Supporting Information, Table S1). *Iguana iguana* was chosen as an outgroup to *Sauromalus* based on its close relationship in previous phylogenetic studies, in comparison to other species within *Iguanidae* (Wiens & Hollingsworth, 2000; Pyron *et al.*, 2013; Malone *et al.*, 2017). We included individuals from across the range of *S. ater* and targeted sampling from potential contact zones (i.e. where mtDNA and/or traditional subspecies boundaries meet), in addition to including samples from type localities or reasonably close proxies.

All work on live animals was conducted under IACUC approval (APF 17-05-005R) at San Diego State University. We obtained permits to collect and/or sample from wild animals from the Arizona Department of Game and Fish (LIC #SP620177), Bureau of Land Management (Palm Springs, South Coast Field Office), California Department of Fish and Wildlife (SCP 13684), National Park Service (ORPI-2017-SCI-0010) and Secretariat of Environment and Natural Resources of Mexico (DGVS–SEMARNAT 2017-09012).

Genomic data collection and bioinformatics

Genomic DNA was extracted from liver, muscle, tail, nail and/or blood samples. Sequence data were collected using the double-digest restriction site-associated DNA sequencing (ddRADseq) method of Peterson *et al.* (2012). We prepared three separate libraries using the same restriction enzymes (SbfI and MspI) and methods as Gottscho *et al.* (2017). Each library was sequenced on three separate NextSeq v.2.5 mid-output flow cell lanes at the Institute for Integrative Genome Biology Core Instrumentation Facility (University of California, Riverside) to generate 2 × 75 bp reads. Raw Illumina reads were demultiplexed, then combined, if from separate sequencing lanes, using the merging function in IPYRAD v.0.7.25 (Eaton & Overcast, 2020). Although we recognize that multiple library preparation and sequencing efforts were not ideal, we were limited by logistical difficulties in acquiring permits and tissues during the course of data collection. All data processing and computationally intensive analyses were performed on the UCR Biocluster (University of California, Riverside, Institute for Integrative Genome Biology).

DNA sequence reads were processed further in IPYRAD to produce *de novo* loci using the following parameters: clustering threshold for *de novo* assembly set to 0.88 and a minimum depth of ten for statistical base calling and majority-rule base calling. Multiple assemblies of the final datasets were constructed with different numbers of individuals and taxa for use in different downstream analyses. The majority of dataset assembly analyses were performed without the outgroup taxon to avoid issues with allelic drop-out (Arnold *et al.*, 2013; Leaché *et al.*, 2015) when constructing RADseq assemblies with distantly related species. We also used different missing data thresholds (i.e. setting the minimum number of samples at a

given locus parameter in IPYRAD) because of a lack of consensus about how to accommodate for missing data in phylogenomic analyses (Huang & Knowles, 2016; Xi *et al.*, 2016), acquisition biases when constructing phylogenies with higher missing data (Leaché *et al.*, 2015) and decreased efficiency in certain programs that tolerate less missing data (i.e. SNAPP; Bryant *et al.*, 2012).

Clustering analyses and concatenated phylogenetic inference

Genomic variation among individuals was visualized using principal component analysis (PCA). Then individuals were assigned to groups using ADMIXTURE (Alexander *et al.*, 2009), a maximum likelihood (ML) population clustering method that operates under Hardy–Weinberg equilibrium assumptions, and two methods of concatenated phylogenetic inferences. Initially, the R program SNPRELATE (Zheng *et al.*, 2012) was used to perform PCA on the variant call format (VCF) file output from IPYRAD of all 91 ingroup individuals (i.e. *Sauromalus*). In ADMIXTURE, the optimal number of populations was determined using cross-validation, testing *K* values from 1 to 5, then selecting the lowest cross-validation score as the optimal number of populations. Maximum likelihood phylogenetic inference was performed in IQ-TREE v.1.68 (Nguyen *et al.*, 2015). The best-fitting molecular evolutionary model and partitioning scheme was inferred using MODELFINDER (Kalyaanamoorthy *et al.*, 2017; GTR+F+R2 based on the Bayesian information criterion). Branch support was accessed via 500 bootstrap replicates using ultrafast bootstrap (UFB) approximation (Hoang *et al.*, 2018), with UFB values ≥ 95 being indicative of strongly supported clades (Minh *et al.*, 2013; Hoang *et al.*, 2018). To explore phylogenetic relationships within lineages of *Sauromalus* and to identify distinct populations without enforcing a strict bifurcating topology, phylogenetic networks were inferred using the NEIGHBORNET algorithm in SPLITSTREE v.4.15.1 (Hudson & Bryant, 2006). This approach has been used in similar studies to identify population clusters and to visualize phylogenetic conflict and uncertainty attributable to potential reticulated evolution (e.g. Barley *et al.*, 2019).

To estimate divergence dates across *Sauromalus*, we constructed two time-calibrated relaxed clock phylogenies in BEAST 2 v.2.6.3 (Bouckaert *et al.*, 2014) using reduced sets of individuals. Previous attempts at divergence dating using all 91 ingroup individuals consistently failed to converge with a lognormal relaxed clock under HKY- and GTR-based substitution models. The first reduced time calibrated phylogeny was based on a dataset including only 20 *Sauromalus* individuals and 1451 loci. The other phylogeny was based on a reduced dataset including the same 20 *Sauromalus* individuals, two outgroup *I. iguana* individuals and 1276 loci. Although fossil calibrations are available for more distant relatives within Iguanidae (Norell, 1989; Norell & de Queiroz, 1991), there are no fossils ancestral to *Sauromalus* or the divergence between *Iguana* and *Sauromalus*. We assumed a mutation rate of 0.00077 per site/Myr based on neutral substitution rate estimates calculated for lizard genomes (Perry *et al.*, 2018). We used a GTR+G substitution model and selected the Yule model as the tree prior. For the reduced time-calibrated phylogeny including only ingroup taxa, two chains were run for 150 million generations, sampling every 5000 generations. The reduced time-calibrated phylogeny including *I. iguana* was run with two chains for 500 million generations, sampling every 5000 generations. Adequate convergence between runs was assessed in TRACER v.1.6 to ensure that effective sample size (ESS) values were > 200. A maximum clade credibility tree was constructed in TREEANNOTATER, discarding the first 10% as burn-in.

Finally, we inferred a dated phylogeny of 36 existing *Cytb* sequences [28 from the studies by Petren & Case (1997, 2002) and 8 from other studies] of *Sauromalus* as a comparison to our dated ddRADseq phylogenies (Supporting Information, Table S2). Sequences were acquired from GenBank, aligned using MUSCLE (Edgar, 2004; Madeira *et al.*, 2019) and partitioned by codon position. A Yule model tree prior and lognormal relaxed clock were selected, and each partition was assigned an HKY+G model because previous runs using a GTR model failed to reach adequate convergence. This mtDNA phylogeny was calibrated with a known mtDNA squamate mutation rate (Barley *et al.*, 2014), using a 95% normal distribution prior, with the mean set to 0.00895 and SD set to 0.0025. Two chains were run for 150 million generations, sampling 1000 generations, then assessed for convergence and combined as above.

Species tree inference and species delimitation

Two MSC species tree methods, SVDQUARTETS (Chifman & Kubatko, 2014) and SNAPP, were used to infer the phylogenetic relationships among divergent lineages of *Sauromalus* identified from the previous clustering analyses. Using PAUP v.4 (Swofford, 2003), SVDQUARTETS species trees were inferred using the dataset that included the outgroup taxon *I. iguana*. This species tree estimation consisted of 1000 bootstrap replicates, and all possible quartets were evaluated. These SVDQUARTETS species tree analyses inferred the relationships among six divergent lineages (putative species) of *Sauromalus* [i.e. continental (C)

group, peninsular (P) group, *S. varius* (V), *S. hispidus* (H), *S. slevini* and *S. klauberi*].

The Bayesian species tree method SNAPP estimates species trees directly from single nucleotide polymorphisms (SNPs) without sampling from gene trees at each locus. This method avoids the computational and statistical difficulties of sampling gene trees from hundreds of loci for phylogenomic species tree inference (Bryant *et al.*, 2012). However, it tolerates only small amounts of missing data (Schmidt-Lebuhn *et al.*, 2017) and is still very computationally intensive. The VCF output from IPYRAD was converted to unlinked biallelic SNP data using VCFTOOLS v.4.2 (Danecek *et al.*, 2011). Owing to the computational intensiveness of SNAPP, we constructed a reduced dataset of 489 unlinked biallelic SNPs of 27 individuals, with at least two individuals representing the four main putative species within *Sauromalus*: a continental (C) group, peninsular (P) group, *S. varius* (V) and *S. hispidus* (H). A species tree was inferred from SNAPP using default priors and only ingroup taxa (i.e. *Sauromalus*). Attempts to generate a species tree also including individuals of *S. slevini* and *S. klauberi* as putative species failed to reach adequate convergence. Thus, we chose to lump them with the peninsular group for this analysis to save computational time. Two Markov chain Monte Carlo (MCMC) chains were run for 5 000 000 generations (10 000 pre-burn-in), and convergence was accessed in TRACER v.1.6. Independent runs were combined in LOGCOMBINER v.2.5.2, discarding the first 10% of trees as burn-in for each run, and subsequently, a maximum clade credibility tree was constructed in TREEANNOTATOR v.2.5.2. The posterior distribution of trees was then visualized in DENSITREE v.2.5.2 (Bouckaert, 2010). An additional SNAPP tree that included the outgroup *I. iguana* was also inferred. This dataset included the same 27 ingroup *Sauromalus*, 2 *I. iguana* and 490 unlinked biallelic SNPs. The analysis was run following the same parameters as the SNAPP analysis of only ingroup taxa.

To validate putative species identified in the previous analyses and those recognized in the current taxonomy, Bayes factor delimitation (BFD*; Grummer *et al.*, 2014) was implemented in SNAPP within BEAST v.2.6.1 using the same dataset as the previous SNAPP Bayesian species tree inference. Three species delimitation models (Table 1) of different numbers of putative species were evaluated. Putative species were chosen based on the current taxonomy and our clustering and phylogenetic analyses. Marginal likelihoods for each model were estimated using 48 path sampling steps, an α-value of 0.3 and a MCMC chain length of 5 000 000 generations (10 000 pre-burn-in) (Leaché *et al.*, 2014). The strength of each alternative species delimitation model was determined by calculating and comparing the Bayes factors (BFs) of each model using the following equation: BF = 2 × (marginal likelihood of model 1 minus marginal likelihood of model 2), where model 1 is the more species rich of the two. Positive BF values indicate support for model 1, whereas negative values indicate support for model 2.

COALESCENT DIVERGENCE DATING ANALYSES

Finally, we used G-PHOCS v.1.3 (generalized phylogenetic coalescent sampler; Gronau *et al.*, 2011) to estimate divergence times between major divergent lineages of *Sauromalus*. G-PHOCS is a Bayesian MSC method that allows for the joint estimation of divergence times, migration rates and effective population size. To reduce computational time and intensiveness, we constructed a reduced dataset of 850 SNPs of 20 individuals, with five individuals representing each of the four major lineages within *Sauromalus*: a continental (C) group, peninsular (P) group, *S. varius* (V) and *S. hispidus* (H). Given that the focus of this study was to estimate the timing of deep divergences within *S. ater*, particularly between peninsular and continental lineages, we chose not to include *S. klauberi* and *S. slevini* in these analyses. Based on the results of our clustering analyses and phylogenetic inferences, individuals located near the contact zone between peninsular and continental lineages were chosen as representatives for these respective groups for this analysis. Divergence dates were calculated for each model using the fixed species topology of the SNAPP tree as a guide. Model A estimated parameters between the continental group and all three peninsular lineages lumped together (= VHP). This model is consistent with all phylogenetic analyses that provide strong support for the continental group being the sister taxon to all remaining *Sauromalus*. Model B estimated parameters for three species: the peninsular group, *S. varius* and

Table 1. Summary of the three Bayes factor delimitation species delimitation models

Number of species	Model	Marginal Likelihood	Bayes factor	Rank
3	(H)(V)(*ater*)	−6254.4699	2270.4452	3
5	(H)(V)(*ater*) (*klaub*) (*slev*)*	−5940.6221	1642.7496	2
4	(C)(V)(H)(P)	−5489.8066	741.1186	1

Abbreviations: C, continental group; H, *Sauromalus hispidus*; *klaub*, *Sauromalus klauberi*; P, peninsular group; *slev*, *Sauromalus slevini*; V, *Sauromalus varius*.
*Current taxonomy (Hollingsworth, 1998).

S. hispidus, with *S. varius* and *S. hispidus* designated as sister species. Owing to uncertainty in the sister-species relationship of *S. varius* and *S. hispidus*, model C estimated parameters for the same species as model B, except that the peninsular group and *S. hispidus* were designated as sister species. All three models were run with and without migration bands between species to compare the effects of potential migration on inferred divergence dates. We assumed a gamma distribution on the priors for each model. Default values of $\alpha = 1$, $\beta = 10\,000$ were used for τ and θ, and default values of $\alpha = 0.002$, $\beta = 0.00001$ were used for migration rates. For each model, two replicate analyses were run for 1 000 000 generations, and the first 10% of samples were discarded as burn-in. Runs were combined in LogCombiner v.2.5.2, and Tracer v.1.6 was used to ensure that runs had converged adequately and that all ESS values were > 200.

To convert coalescent units from G-PhoCS to absolute time, we used the equation: $T = (\tau g)/\mu$. As in the previous time-calibrated phylogeny, a mutation rate of 0.00077 per site/Myr was assumed. We used the earliest age of reproductive maturity as a proxy for generation time (g). To account for uncertainty and variation in growth rates between populations and species, generation times of 2 and 3 years were used, based on life-history observations on wild *S. ater* (Berry, 1974; Abts, 1987; but see Tracy, 1999) and *S. varius* (Sylber, 1985). Finally, we calculated the proportion of migrants per generation by multiplying raw migration rate(m) by the per-generation mutation rate (μ).

RESULTS

Double-digest restriction site-associated DNA sequencing data

Processing of DNA sequence reads in ipyrad resulted in two final alignments, which were modified further in downstream analyses. The total number of reads per individual from the first alignment, including all *Sauromalus* and *I. iguana*, ranged from 119 923 to 2 363 563, with an average of 673 769 raw reads per individual and a total of 32 621 prefiltered loci. After filtering, the alignment contained 1122 loci present in a minimum of 41 out of 93 individuals, 154 880 bp and 10 092 SNPs. The median coverage per sample after filtering was 777 loci. The total number of reads per individual from the second alignment, including only *Sauromalus*, had an identical range to the first alignment, with an average of 672 385 raw reads per individual and 30 082 prefiltered loci. After filtering, the alignment contained 1146 loci present in a minimum of 40 out of 91 individuals, 157 670 bp and 9051 SNPs. The median coverage per sample after filtering was 790 loci.

The first and second alignments had 31.94% and 31.68% of missing data, respectively. Previous studies have shown that RADseq datasets with moderate to high levels of missing data still produce robust phylogenetic inferences because they often contain a large and significant number of parsimony-informative SNPs and are not biased towards slowly evolving sites (Huang & Knowles, 2016; Tripp *et al.*, 2017; Crotti *et al.*, 2019). We acknowledge that missing data in our study might have been caused by technical inconsistencies between our three separate library preparation and sequencing efforts, despite using identical protocols; however, this is a common concern for all studies based on the combining of genomic data from multiple libraries. Furthermore, although *I. iguana* represents the closest outgroup to *Sauromalus*, their more distant relationship might have been a source of allelic drop-out leading to missing data.

Phylogeographical structure and phylogenetic relationships

The ML phylogenetic inferences of the concatenated dataset of *Sauromalus* strongly supported at least four major groups within *Sauromalus* that corresponded broadly to a continental group consisting of *S. ater* (in part) and a peninsular group containing *S. ater* (in part) and the four insular endemics (Fig. 2A). Within the peninsular group, there was a gradual progression from the northern to southern peninsula, in addition to the monophyly of *S. klauberi* and *S. slevini*, although they were nested within *S. ater* samples and were not each other's closest relative. *Sauromalus slevini* was sister to an individual from Isla San Cosme (SDF 4040), whereas *S. klauberi* was closely related to southern peninsular samples. The topology of the ML phylogeny strongly supported the basal positions of *S. varius* and *S. hispidus* to the peninsular group, with *S. hispidus* being sister to the peninsular group. This differs from the conclusions of the mtDNA studies of Petren & Case (1997, 2002), which placed the insular gigantics nested within peninsular *S. ater*, although this result was not strongly supported (bootstrap < 70%). Within continental *S. ater*, the ML phylogeny yielded roughly four geographically structured groups [eastern Sonoran (east of the Salton Sea), western Sonoran, Mojave and southern Sonoran (south of Caborca, Mexico and along the coast)], with moderate to strong support. Phylogenetic network results from SplitsTree of all *Sauromalus* individuals were largely concordant with the results from the ML phylogeny and displayed a deep divergence between the continental and peninsular groups of *S. ater* (Fig. 2C).

In the PCA (Fig. 3A), the first principal component axis (PC1) clearly distinguished continental *S. ater*

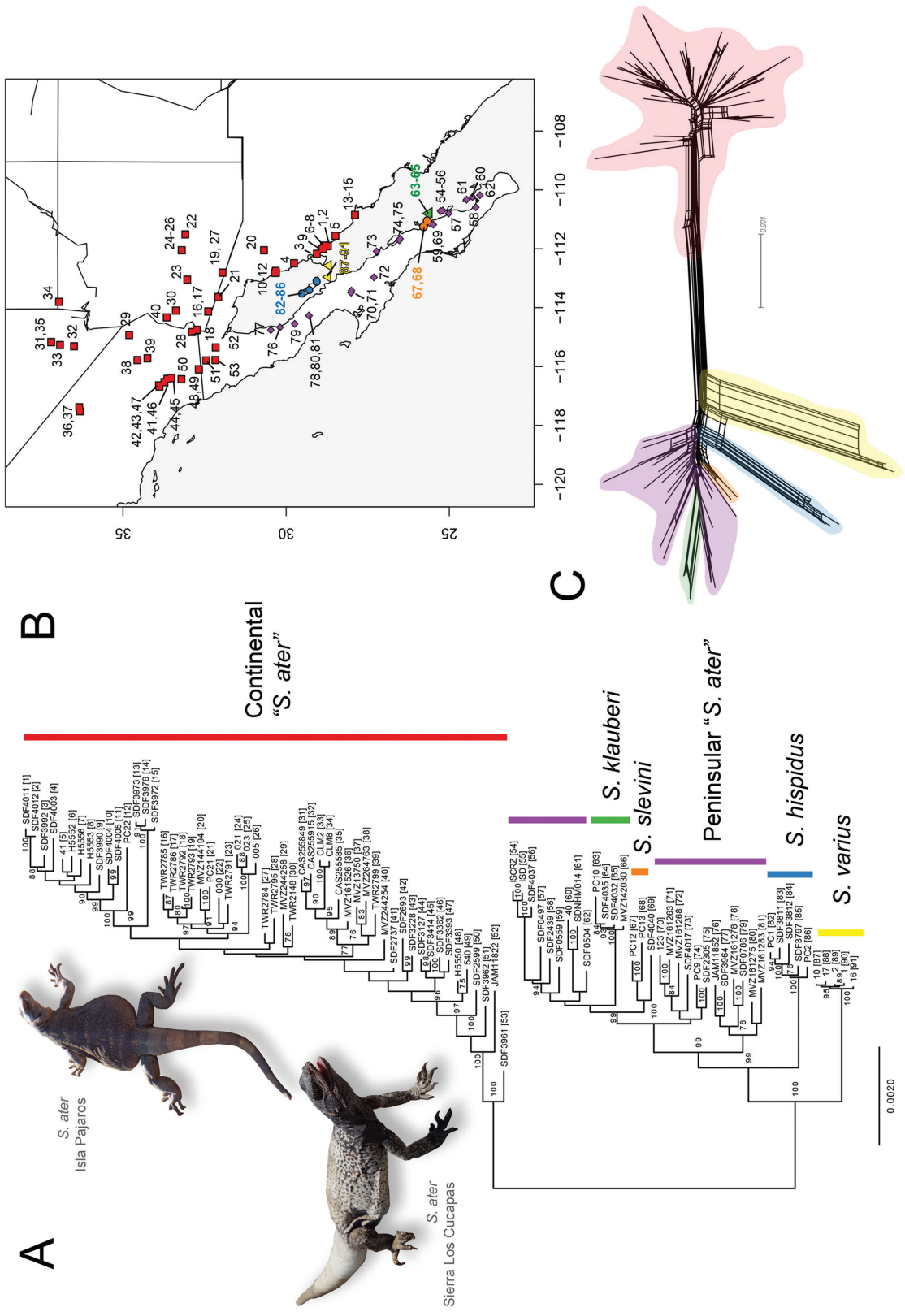

Figure 2. A, maximum likelihood (ML) phylogeny of all *Sauromalus* based on concatenated double-digest restriction site-associated DNA sequencing data. Numbers at nodes indicate bootstrap support. Major groups and currently recognized taxa are labelled. Pictures taken by A. Sumarli. B, sampling localities of the 91 individuals used in this study. Numbers

from the rest of *Sauromalus*. The second (PC2) distinguished *S. varius* as a distinct group from peninsular *Sauromalus*. Separation of *S. hispidus* from peninsular *Sauromalus* was present along PC2 but was not as distinct. The third principal component axis (PC3) clearly distinguished *S. hispidus* from all *Sauromalus* and moderately separated the continental group from the rest of peninsular *Sauromalus*, including *S. varius*, *S. slevini* and *S. klauberi*. Along the fourth principal component axis (PC4), the continental group was spread along a north-to-south axis, with individuals from Isla Pajáros, the southernmost extent of *S. ater* in Sonora, Mexico, forming a distinct cluster from the continental group. *Sauromalus slevini* and *S. klauberi* did not form distinct clusters along any PC axes and were grouped within peninsular *S. ater*.

ADMIXTURE (Fig. 3B) supported $K = 4$ (out of 5) as the optimal number of populations, and these populations corresponded to similar groupings as the previous clustering and phylogenetic analyses. However, under $K = 4$ the continental group was split roughly into north-western and south-eastern groups, and there appeared to be admixture between individuals in south-western Arizona. *Sauromalus varius* and *S. hispidus* were clustered into a single distinct group, whereas the other insular endemics were grouped within peninsular *S. ater*. There appeared to be slight admixture between continental and peninsular groups near the vicinity of the contact zone in northern Baja California. Interestingly, there also appeared to be slight admixture between some continental and peninsular individuals and the island gigantics.

Coalescent species tree inference and species delimitation

Both Bayesian species trees inferred in SNAPP of the higher-level species relationships of *Sauromalus* strongly supported the basal position of the continental group of *S. ater* relative to the rest of *Sauromalus* (Fig. 4). However, the SNAPP tree of only ingroup taxa (Fig. 4A) strongly supported the sister relationship between *S. varius* and S. *hispidus* (posterior probability = 0.97), similar to the mtDNA studies by Petren & Case (1997, 2002). The SNAPP tree including an outgroup (Fig. 4B) strongly supported *I. iguana* as sister to *Sauromalus* and inferred a similar topology to the concatenated ML and Bayesian phylogenies. *Sauromalus varius* was weakly supported as sister to a clade including peninsular *S. ater* and *S. hispidus*. The sister relationship of *S. hispidus* and peninsular *S. ater* was also weakly supported.

The SVDQUARTETS species tree of all *Sauromalus* individuals and the outgroup taxon *I. iguana* was inferred using six a priori putative species, which were identified in the clustering analyses and represent the current taxonomy. This resulted in a similar topology to both concatenated ddRAD phylogenies and the SNAPP species tree that included an outgroup. However, most of the SVDQUARTETS inferred relationships appeared to be weakly supported (Supporting Information, Fig. S1). This analysis also inferred the continental group of *S. ater* to be the sister lineage to the rest of *Sauromalus*. Within the remaining *Sauromalus*, *S. varius* was weakly supported as sister to the more exclusive clade containing *S. hispidus* and the peninsular group of *S. ater*. The sister relationship of *S. slevini* and *S. klauberi* was moderately supported. Furthermore, BFD* analyses supported the continental and peninsular groups as a distinct lineage (Table 1). Models hypothesizing that all *S. ater* populations (i.e. the continental and peninsular groups) should be grouped within a single wide-ranging species were the poorly supported.

Divergence dating analyses

The Bayesian relaxed clock time trees using reduced sets of individuals broadly estimated the split between the continental and peninsular groups to be 7.1 Mya [95% highest posterior density (HPD): 3.57–11.28 Mya] in the phylogeny including *I. iguana* (Supporting Information, Fig. S2) and 5.42 Mya (95% HPD: 2.98–8.24 Mya) in the phylogeny including only ingroup taxa (Supporting Information, Fig. S3). In both time trees, *S. varius* diverged from *S. hispidus* and peninsular *Sauromalus* in the early Pleistocene to late Miocene. These dates corresponded to 4.38 Mya (95% HPD: 2.07–7.12 Mya) in the phylogeny including *I. iguana* and 3.66 Mya (95% HPD: 1.95–5.66 Mya) in the phylogeny with only ingroup taxa. The timing of the split of *S. hispidus* from peninsular *Sauromalus* ranged from the mid-Pleistocene to early Pliocene and late Miocene. These dates corresponded to 3.22 Mya (95% HPD:1.48–5.46 Mya) in the phylogeny including *I. iguana* and 2.66 Mya (95% HPD: 1.42–4.14 Mya) in the phylogeny with only ingroup taxa. The time tree including an outgroup (Supporting Information, Fig. S2) also confirmed the sister relationship of

correspond to individual field numbers labelled on the ML phylogeny, and colours correspond to currently recognized taxa or major groups. C, phylogenetic network of all *Sauromalus* based on concatenated double-digest restriction site-associated DNA sequencing data. Numbers indicate bootstrap proportions. Major lineages and currently recognized species are labelled. Colours correspond to groups identified in the ML phylogeny.

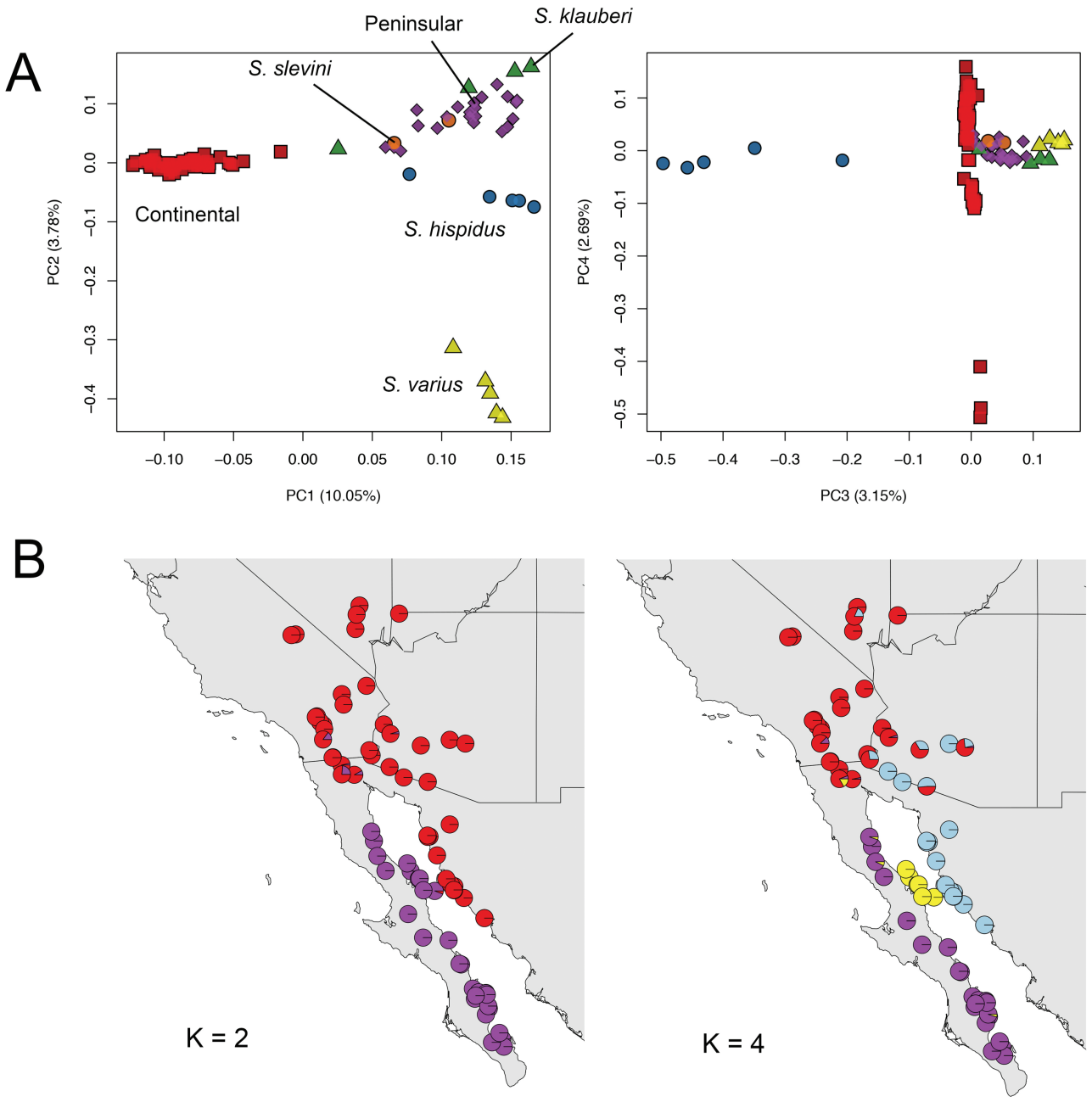

Figure 3. A, principal component analysis of all *Sauromalus*. B, ADMIXTURE results for all *Sauromalus*. Pie chart sections represent the proportion of ancestry for each individual at a site. Results are shown for $K = 2$ and the optimal value of $K = 4$. Colours correspond to the following populations: red, continental; purple, peninsular; light blue, south-eastern continental; yellow, *Sauromalus varius* and *Sauromalus hispidus*.

Sauromalus and *Iguana* and suggest that the two genera diverged ~13.85 Mya (95% HPD: 6.96–21.1 Mya), similar to dates inferred by Malone *et al.* (2017). The dated *Cytb* phylogeny (Supporting Information, Fig. S4) estimated that continental and peninsular groups diverged in the late Pliocene to late Miocene (5.92 Mya; 95% HPD: 2.98–9.72 Mya).

Divergence times and demographic parameters estimated by G-PHOCS are summarized in Figure 5 and Supporting Information, Table S3, respectively. Overall, models run with migration produced older dates than those run under an isolation model, especially in model A. Furthermore, dates inferred using a generation time of 3 years were older than

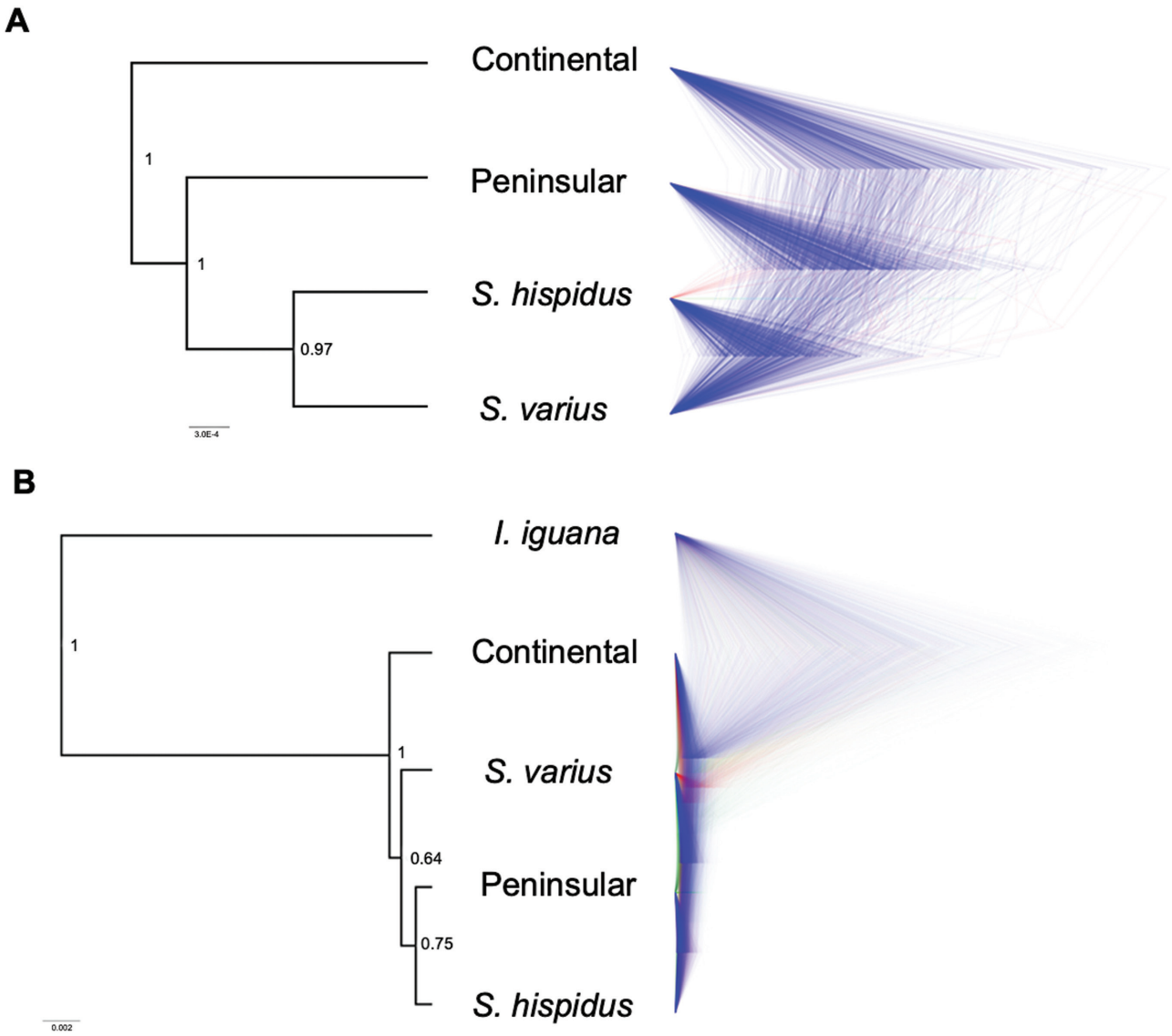

Figure 4. Bayesian species trees based on double-digest restriction site-associated DNA sequencing single nucleotide polymorphism data showing the higher-level species relationships of *Sauromalus* of only ingroup taxa (A) and higher-level species relationships of *Sauromalus* including the outgroup taxon, *Iguana iguana* (B). To the right of the putative species are the posterior distributions visualized with DENSITREE (blue represents the dominant topology, and red/green represent alternative topologies), and on the left are the maximum clade credibility trees.

those using 2 years. Divergence dates across all models were mostly within the margin of error or similar to those estimated in BEAST, regardless of the generation time used. For model A the dates were within the early Pliocene to early Miocene, and for models B and C the dates were within the early Pleistocene to late Pliocene. However, dates inferred under model A using a generation time of 3 years and run with migration produced the oldest range of dates, within the mid-Miocene. Interestingly, model C, which used the topology of the concatenated ddRADseq phylogeny, SVDQUARTETS species tree and SNAPP species tree including *I. iguana*, produced slightly older dates than model B, which used the topology of the SNAPP species tree of only ingroup taxa and *Cytb* phylogeny. G-PHOCS detected little to no gene flow across all three models [effective population migration rate (2NM) < 1; Supporting Information, Table S3].

DISCUSSION

Several previous studies have proposed the phylogenetic relationships and historical biogeography of *Sauromalus*. However, those studies have been limited by their primary reliance

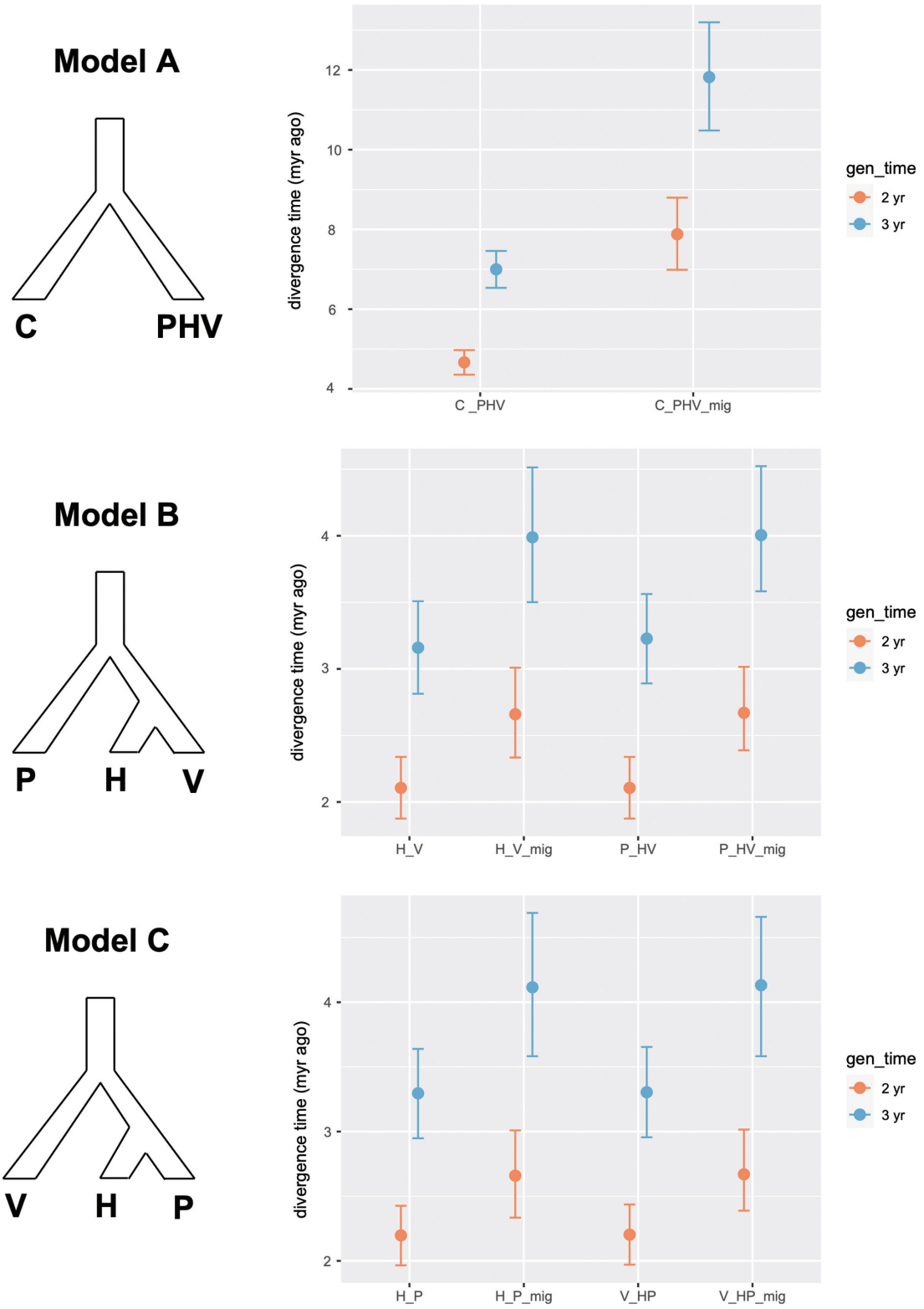

Figure 5. Divergence times inferred by G-PHOCS. On the right are comparisons of divergence times inferred with and without migration and with a generation time of 2 or 3 years. The *x*-axis indicates the species hypothesis being tested, and the *y*-axis indicates millions of years. Schematic diagrams of each model tested are shown on the left. Abbreviations for species are as follows: C, continental; H, *Sauromalus hispidus*; P, peninsular; V, *Sauromalus varius*. 'mig' indicates that the model was run with migration.

on morphology (Hollingsworth, 1998), restriction site polymorphisms of mtDNA (Lamb *et al.*, 1991), single-locus mitochondrial data, limited geographical sampling and less sophisticated phylogenetic methods (i.e. unweighted maximum parsimony; Petren & Case, 1997, 2002). This study represents the first phylogenetic study on *Sauromalus* using genomic data (ddRADSeq), coalescent model-based phylogenetic methods and improved geographical sampling. It is also the most comprehensive molecular phylogenetic study on *Sauromalus* to date and the first to generate genomic species-level phylogenetic hypotheses. Although our study is focused explicitly on the evolutionary history of what is currently recognized as the widespread and morphologically variable *S. ater*, it also includes individuals from the other currently recognized species: *S. varius*, *S. hispidus*, *S. klauberi* and *S. slevini*.

Higher-level species relationships within *Sauromalus*

All concatenated and species tree analyses moderately to strongly supported the phylogenetic position of the continental group of *S. ater* as the sister lineage to the remaining *Sauromalus*, which includes the peninsular group of *S. ater*, the insular chuckwallas *S. klauberi* and *S. slevini*, and the 'islands gigantics' *S. hispidus* and *S. varius*. The relatively basal positions of *S. varius* and *S. hispidus* to the rest of the peninsular group were strongly supported in the concatenated ML phylogeny and weakly supported in the SVDQUARTETS species tree and SNAPP species tree including *I. iguana*. However, similar to the mtDNA study of Petren & Case (1997, 2002) and unlike the other genomic phylogenetic inferences presented in this study, the species tree from SNAPP of only ingroup taxa strongly supported the sister relationship of *S. varius* and *S. hispidus* and placed this clade as sister to the remaining peninsular group. In SNAPP, species relationships appeared to be influenced by the inclusion or exclusion of *I. iguana*. Excluding *I. iguana* resulted in stronger support for the sister relationship of *S. varius* and *S. hispidus*. The competing relationships and varying amounts of support at the base of peninsular *Sauromalus* suggest that the divergence time between peninsular *S. ater*, *S. varius* and *S. hispidus* might have been short in comparison to the deeper split between continental and peninsular groups. These short divergence intervals were supported by the similar ranges of dates inferred in G-PHOCS for both models B and C (Fig. 5). Owing to the uncertainty between topologies, we treat the two species hypotheses for peninsular *Sauromalus* as equivocal to one another until future analyses are conducted.

Phylogeographical structure within major groups of *S. ater*

Comparative analyses across several taxonomic groups have shown that the Mohave/Sonoran Desert ecotone and lower Colorado river might limit or promote genetic divergence by facilitating geneflow between populations (Lamb *et al.*, 1991; Wood *et al.*, 2013; Dolby *et al.*, 2019). Within the continental group of *S. ater*, our concatenated phylogenetic and genomic clustering analyses strongly suggest there are geographically distinct genetic clusters that correspond roughly to the different desert regions in south-western North America. Our results are similar to early genetic studies by Lamb *et al.* (1991) and Petren & Case (1997, 2002) that identified significant mtDNA geographical variation within continental *S. ater*. Because of the wide-ranging and continuous geographical distribution of continental *S. ater* and the difficulty this poses for species delimitation, especially under the MSC (Chambers & Hillis, 2020), we did not attempt to validate any putative species within this group. The additional genetic clusters within continental *S. ater* require further investigation using demographic modelling and landscape genomic approaches to understand the diversification history of this lineage. Much denser sampling is also needed at potential contact zones, especially given evidence of admixture between south-eastern and north-western continental *S. ater*.

Numerous studies across a wide set of taxa occurring in Baja California (Riddle *et al.*, 2000; Crews & Hedin, 2006; Leaché & Mulcahy, 2007; Harrington *et al.*, 2017) have focused on evaluating the nature and timing of mid-peninsular vicariant events. These phylogeographical breaks correspond to the following regions: the mid-peninsula in the central Vizcaíno region, near the town of Loreto, and the Isthmus of La Paz (Riddle *et al.*, 2000). Our concatenated phylogeny did not clearly support the presence of any of the aforementioned breaks, and instead supported a gradual north-to-south cline within peninsular *S. ater*. This cline was similar to the structure presented by Petren & Case (2002), although their study had significantly less sampling. Petren & Case (2002) demonstrated that the genetic distance between the north and south peninsula was strongly correlated with geographical distance (i.e. isolation by distance), and there is a possibility that a similar pattern might be occurring within our data. Further sampling across the peninsular and additional landscape genomic and demographic analyses are required to determine the mechanisms causing such population structure and/or to evaluate whether this genomic structure exists.

Lastly, common garden experiments by Tracy (1999)demonstrated that populations of *S. ater* collected

from varying environmental conditions, habitat types and elevations exhibited heritable differences in body size and patterns of growth. Generally, populations of *S. ater* collected from higher elevations experienced lower frequencies of drought and longer growing seasons in comparison to low-elevation populations. This allowed them to allocate more resources to reaching larger adult body sizes, in comparison to the more quickly maturing but smaller low-elevation *S. ater*. Although we did not test for a correlation between genetic variation and adaptation to different environmental conditions, this might explain some of the phylogeographical structure observed in both continental and peninsular groups.

Historical biogeography

Murphy (1983) and Welsh (1988) hypothesized that *Sauromalus* evolved in the cape region of the Baja California peninsula, whereas Grismer (1994) suggested a continental (i.e. non-peninsular) origin for *Sauromalus*, followed by a relatively recent westward expansion into the peninsula. The deep phylogenetic divergence between the continental and peninsular groups of *S. ater* in the vicinity of the head of the Gulf of California supports neither of these previous biogeographical scenarios. Our phylogenetic and divergence dating analyses instead support the northern gulf vicariance scenario proposed by Grismer (1994) and Murphy (1983). Grismer (1994) suggested that lineages belonging to this biogeographical scenario exhibit a contemporary circum-gulf distribution and sister relationship between a peninsular and non-peninsular lineage at the vicinity of the head of the Gulf of California. Following this scenario, the ancestral *Sauromalus* is likely to have had a historical distribution around the head of the Gulf of California before the northernmost extent of the Gulf of California at San Gorgonio Pass in southern California. Grismer (1994) placed the date of northern gulf vicariance at 3 Mya, whereas more recent geomorphological studies have established an older date at ~6.3 Mya (Oskin & Stock, 2003; Bennett & Oskin, 2014). Our study placed the *Sauromalus* split at ~5.92 Mya in the *Cytb* time tree (Supporting Information, Fig. S4). The relaxed clock analyses of the concatenated ddRAD data and G-PHOCS (Fig. 5; Supporting Information, Figs S2, S3; Table S3) inferred mostly similar dates ranging from the early Pliocene to mid-Miocene. Thus, the basal split of the common ancestor of *Sauromalus* (i.e. the formation of a basal continental and peninsular clades) is likely to have occurred owing to the northwardly extension of the Gulf of California in the late Miocene, which then separated the basal *Sauromalus* clades onto opposite sides of the gulf. The regression of the gulf to its current distribution then allowed for the secondary contact of these sister lineages around the vicinity of northern Baja California.

The continental and peninsular clades are likely to come into contact with one another in northern Baja California, although the exact location and nature of this contact zone is currently unknown, and there are few records of chuckwallas from this region. Despite the paucity of samples in this region, there is no reason to believe there is an absence of chuckwallas, because the few known records fall within our sampling gap, and there is an abundance of rocky habitat. If there is a secondary contact, it is likely to occur in the north-eastern portion of the peninsula, between Cañon Guadalupe along the eastern escarpment of the Sierra Juarez and Arroyo Matomí (~200 km apart). Although not a significant result of this study, G-PHOCS detected minimal unidirectional gene flow (Supporting Information, Table S3) from the peninsular lineage to the continental lineage. Furthermore, admixture analysis detected minimal proportions of peninsular ancestry within continental samples from Cañon Guadalupe (SDF3961) and Sierra Cucapá (JAM11822), which represent the sampling locations closest to where this major contact zone could be. Other squamate (Leaché & Mulcahy, 2007; McGuire *et al.*, 2007; Wood *et al.*, 2008; Blair & Sanchez-Ramirez, 2016; Leavitt *et al.*, 2020), mammal (Riddle *et al.*, 2000) and spider (Crews & Hedin, 2006) species from south-western North America share a similar biogeographical scenario to continental and peninsular *S. ater*. Of the lizard species that diversified in northern Baja California, Leavitt *et al.* (2020) demonstrated that within western banded geckos (*Coleonyx variegatus*) two subspecies of this group, *Coleonyx variegatus variegatus* and *Coleonyx variegatus abbotti*, come into secondary contact and form a narrow hybridization zone in the Puertecitos region of north-eastern Baja California. This same study estimated that continental and peninsular *C. variegatus* diverged from each other ~6.28 Mya (5.28–7.32 Mya), which overlaps with dates estimated in our divergence dating analyses of the ddRAD and *Cytb* data.

Although not the focus of our study, the evolution of large body size in *Sauromalus* has remained a contentious biogeographical inquiry for several decades. Based on ecological models, Case (1982) hypothesized that large body size evolved in *S. varius* and *S. hispidus* as an adaptation to the island environment ('insular gigantism'). In both species, insular gigantism might then have evolved rapidly owing to intense selection pressures on the islands (Case, 1976, 1982). This hypothesis was supported by the placement of *S. varius* and *S. hispidus* as sister taxa relative to peninsular and continental *Sauromalus* in the mtDNA phylogenies of Petren & Case (1997, 2002). Similar to Shaw (1945) and based on phylogenetic

relationships in the study by Norell & de Queiroz (1991) inferred using fossil data, Grismer *et al.* (1995) presented an alternative hypothesis that the small size of mainland *Sauromalus* evolved through 'continental dwarfism' rather than 'insular gigantism'. Specifically, Grismer *et al.* (1995) suggested that because large iguanid body sizes were thought to be a characteristic of the ancestral *Sauromalus*, a continental dwarfism scenario would be more parsimonious than an insular gigantism scenario. Overall, because our results for the relationships between peninsular *S. ater*, *S. varius* and *S. hispidus* are equivocal, both body size evolution scenarios remain equally parsimonious.

Dates inferred in our divergence dating analyses among *S. varius* and *S. hispidus* (Fig. 5; Supporting Information, Figs S2, S3; Table S3) mostly overlap with those proposed for the ages of the formation of the volcanic Isla San Esteban (2.5–4.5 Mya; Calmus *et al.*, 2008) and the disconnection of the north-western shoreline of Isla Ángel de la Guarda (2–3.3 Mya; Aragón-Arreola & Martín-Barajas, 2007). The ancestor of *S. varius* would probably have needed to colonize Isla San Esteban through overwater dispersal, whereas the ancestor of *S. hispidus* could have existed in the vicinity of Isla Ángel de la Guarda before its disconnection from the peninsula. Alternatively, it could have reached Isla Ángel de la Guarda through overwater dispersal from the peninsula or from a neighbouring midriff island.

Taxonomic implications

Genomic species relationships within *S. ater* were mostly incongruent with any of the previous species and subspecies boundaries of *Sauromalus* based on morphology and failed to explain the geographical variation within *S. ater*. Both continental and peninsular *S. ater* exhibit similarly high levels of morphological variation compared with the single, widespread *S. ater*. Taken together, our results highlight the cryptic diversity within *S. ater* and strongly suggest that the continental and the peninsular groups of *S. ater* represent at least two separately evolving metapopulations (de Queiroz, 2007).

The taxonomy of *S. ater* has been complicated traditionally by the absence of the type locality of the holotype of *S. ater*, which was eventually restricted to southern Sonora based on the morphological analyses by Hollingsworth (1998) and further restricted to Guaymas, Sonora, Mexico by Montanucci (2008) based on historical ship logs from the expedition on which the type specimen was collected. Thus, the continental group would retain the name *S. ater* and the oldest available name for peninsular '*S. ater*' (exclusive of *S. klauberi* and *S. slevini*) would be *Sauromalus interbrachialis* (Dickerson, 1919). Both Schmidt (1922) and Shaw (1945) commented extensively on the proper type locality of *S. interbrachialis*, which they believed to be one of the southern islands off the coast of the Baja California peninsula, rather than the listed location of La Paz. Either way, both locations fall within the peninsular *S. ater* group, making it the oldest available name. Junior synonyms would include *S. australis* (Shaw, 1945) and *S. shawi* (Cliff, 1958).

Problematic to the recognition of *S. interbrachialis* are the embedded relationships of *S. klauberi* and *S. slevini*, within the peninsular clade. *Sauromalus klauberi* and *S. slevini* have been recognized as full species since the work of Murphy (1983). However, the validity of *S. slevini* and *S. klauberi* as distinct lineages is questionable because they do not form distinct groups in any genomic clustering analyses. We hypothesize that the insular morphology of *S. slevini* and *S. klauberi* might be caused by recent isolation from the peninsula and rapid phenotypic diversification associated with drift and/or local adaptation to island environments. However, because both *S. slevini* and *S. klauberi* have consistently represented a distinct morphological cluster in several studies (i.e. Hollingsworth, 1998; Montanucci, 2004), we advocate for further studies with improved sampling of individuals from *S. slevini*, *S. klauberi* and other southern Baja California islands. Such studies are necessary to investigate the genomic distinctiveness of *S. slevini* and *S. klauberi* before they can be synonymized with the peninsular species. There has been little dispute about the species status of *S. varius* and *S. hispidus*, and although testing their validity was not the focus of our study, the genomic results mostly supported their continued recognition.

Conclusion

Adopting a genomic approach, we argue that the widespread and morphologically variable *S. ater* represents at least two distinct continental and peninsular lineages that diversified around the vicinity of the head of the Gulf of California. Divergence dating estimates using concatenated and coalescent methods suggest that these two lineages speciated in the early Pliocene to mid-Miocene, during the formation of the northern Gulf of California. Although the exact location and nature of the contact zone are unknown, the two lineages are likely to come into secondary contact in northern Baja California, along the desert slopes of the eastern Sierras Juarez and San Pedro Mártir. Higher-level phylogenetic relationships within peninsular and insular lineages of *Sauromalus*, namely the sister relationship of *S. varius* and *S. hispidus*, are largely dependent on the inclusion of the outgroup taxon *I. iguana*. The insular species *S. klauberi* and *S. slevini* should continue to be recognized as full species, but their

species status requires further investigation with much-improved geographical sampling. The evolution of large body size in *S. varius* and *S. hispidus* is more complex than previous biogeographical studies have hypothesized and requires further examination. Our study contributes to our understanding of the herpetofauna of the Baja California peninsula and the evolution and systematics within *Sauromalus*. It also lays a foundation for future studies to explore potential secondary contact zones, additional species limits, trait evolution and island biogeography using chuckwallas as a model system.

ACKNOWLEDGEMENTS

We would like to thank Evan McCartney-Melstad and Peter Scott (UCLA La Kretz Center Conservation Genomics Workshop) for their assistance during the initial data-processing stage of this project; Sam Fellows, Sean Harrington, Melissa Stepek, Elizabeth García Aviña, Melba Alvarez Villegas, and Anny Peralta-García for help with collecting specimens; Tierney Bougie and Robert N. Fisher (USGS) for assistance with figures; Christopher Tracy for *Sauromalus* natural history advice; Dean Leavitt and Adam Leaché for analytical advice; Kieran Samuk for comments on the manuscript before submission; and two anonymous reviewers for thier comments and suggestions on the submmited manuscript. We are grateful to many individuals and institutions [Carol L. Spencer and Jim McGuire (Museum of Vertebrate Zoology), Kenneth Petren (University of Cincinnati, Ohio) and the late Ted J. Case (University of California, San Diego), Catherine Malone (Utah Valley University), Donna L. Dittmann (LSU Museum of Natural Science Collection of Genetic Resources) and Peter Holm (Organ Pipe National Monument)] who have provided us with specimens and tissue samples. A.S. was supported by the Society of Systematic Biologists Graduate Student Research Award, American Society of Ichthyologists and Herpetologists Gaige Fund Award, American Museum of Natural History Theodore Roosevelt Memorial Grant, and Anza-Borrego Foundation Howie Wier Memorial Conservation Grant.

AUTHOR CONTRIBUTIONS

Alexandra Sumarli (Conceptualization, Formal analysis, Investigation, Data Curation, Writing—Original Draft, Visualization, Project administration, Funding Acquisition); Bradford D. Hollingsworth (Conceptualization, Resources, Writing—Reviewing & Editing, Supervision); Jorge H. Valdez-Villavicencio (Resources, Writing—Reviewing & Editing); and Tod W. Reeder (Conceptualization, Resources, Writing—Reviewing & Editing, Supervision).

CONFLICT OF INTEREST

The authors do not declare any conflict of interests.

DATA AVAILABILITY

All input files used for analyses in this study are deposited in Figshare (https://doi.org/10.6084/m9.figshare.24069156). Raw sequence data are deposited in the NCBI Sequence Read Archive under BioProject (PRJNA1004988).

REFERENCES

Abts ML. 1987. Environment and variation in life history traits of the chuckwalla, *Sauromalus obesus*. *Ecological Monographs* **57**: 215–232. https://doi.org/10.2307/2937081

Alexander DH, **Novembre J**, **Lange K. 2009**. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**: 1655–1664. https://doi.org/10.1101/gr.094052.109

Aragón-Arreola M, **Martín-Barajas A. 2007**. Westward migration of extension in the northern Gulf of California, Mexico. *Geology* **35**: 571–574. https://doi.org/10.1130/g23360a.1

Arnold B, **Corbett-Detig RB**, **Hartl D**, **Bomblies K. 2013**. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology* **22**: 3179–3190. https://doi.org/10.1111/mec.12276

Barley AJ, **Datta-Roy A**, **Karanth KP**, **Brown RM. 2014**. Sun skink diversification across the Indian–Southeast Asian biogeographical interface. *Journal of Biogeography* **42**: 292–304. https://doi.org/10.1111/jbi.12397

Barley AJ, **Nieto-Montes de Oca A**, **Reeder TW**, **Manríquez-Morán NL**, **Arenas Monroy JC**, **Hernández-Gallegos O**, **Thomson RC. 2019**. Complex patterns of hybridization and introgression across evolutionary timescales in Mexican whiptail lizards (*Aspidoscelis*). *Molecular Phylogenetics and Evolution* **132**: 284–295. https://doi.org/10.1016/j.ympev.2018.12.016

Bennett SEK, **Oskin ME. 2014**. Oblique rifting ruptures continents: example from the Gulf of California shear zone. *Geology* **42**: 215–218. https://doi.org/10.1130/G34904.1

Berry KH. 1974. The ecology and social behavior of the chuckwalla, *Sauromalus obesus obesus* Baird. *University of California Publications in Zoology* **101**: 1–60.

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. https://doi.org/10.1016/j.ympev.2015.12.020

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**: 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bouckaert R, **Heled J**, **Kühnert D**, **Vaughan T**, **Wu C-H**, **Xie D**, **Suchard MA**, **Rambaut A**, **Drummond AJ. 2014**. BEAST 2: a software platform for Bayesian evolutionary

analysis. *PLoS Computational Biology* **10**: e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bryant D, **Bouckaert R**, **Felsenstein J**, **Rosenberg NA**, **RoyChoudhury A. 2012**. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29**: 1917–1932. https://doi.org/10.1093/molbev/mss086

Calmus T, **Pallares C**, **Maury RC**, **Bellon Ĥ**, **Pérez-Segura Én**, **Aguillón-Robles A**, **Carreño A-L**, **Bourgois J**, **Cotten J**, **Benoit M. 2008**. Petrologic diversity of Plio-Quaternary post-subduction volcanism in northwestern Mexico: an example from Isla San Esteban, Gulf of California. *Bulletin de La Société Géologique de France* **179**: 465–481. https://doi.org/10.2113/gssgfbull.179.5.465

Carstens BC, **Pelletier TA**, **Reid NM**, **Satler JD. 2013**. How to fail at species delimitation. *Molecular Ecology* **22**: 4369–4383. https://doi.org/10.1111/mec.12413

Case TJ. 1976. Body size differences between populations of the chuckwalla, *Sauromalus obesus*. *Ecology* **57**: 313–323. https://doi.org/10.2307/1934819

Case TJ. 1982. Ecology and evolution of the insular gigantic chuckwallas, *Sauromalus hispidus* and *Sauromalus varius*. In: Burghardt GM, Rand AS, eds. *Iguanas of the world: their behavior, ecology, and conservation*. Park Ridge: Noyes, 184–212.

Chambers EA, **Hillis DM. 2020**. The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Systematic Biology* **69**: 184–193. https://doi.org/10.1093/sysbio/syz042

Chifman J, **Kubatko L. 2014**. Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**: 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Cliff FS. 1958. A new species of *Sauromalus* from Mexico. *Copeia* **4**: 259–261. https://doi.org/10.2307/1439955

Crews SC, **Hedin M**. **2006**. Studies of morphological and molecular phylogenetic divergence in spiders (Araneae: Homalonychus) from the American southwest, including divergence along the Baja California Peninsula. *Molecular phylogenetics and evolution*. **38**: 470–487. https://doi.org/10.1016/j.ympev.2005.11.010

Crotti M, **Barratt CD**, **Loader SP**, **Gower DJ**, **Streicher JW. 2019**. Causes and analytical impacts of missing data in RADseq phylogenetics: insights from an African frog (*Arixalus*). *Zoologica Scripta* **48**: 157–167. https://doi.org/10.1111/zsc.12335

Danecek P, **Auton A**, **Abecasis G**, **Albers CA**, **Banks E**, **DePristo MA**, **Handsaker RE**, **Lunter G**, **Marth GT**, **Sherry ST**, **McVean G**, **Durbin R**; **1000 Genomes Project Analysis Group**. **2011**. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Davy DM, **Méndez de la Cruz FR**, **Lathrop A**, **Murphy RW. 2011**. Seri Indian traditional knowledge and molecular biology agree: no express train for island-hopping spiny-tailed iguanas in the Sea of Cortés. *Journal of Biogeography* **38**: 272–284. https://doi.org/10.1111/j.1365-2699.2010.02422.x

Dickerson MC. 1919. Diagnosis of twenty-three new species and a new genus of lizards from Lower California. *Bulletin of the American Museum of Natural History* **41**: 461–477.

Dolby GA, **Bennett SEK**, **Lira-Noriega A**, **Wilder BT**, **Munguía-Vega A. 2015**. Assessing the geological and climatic forcing of biodiversity and evolution surrounding the Gulf of California. *Journal of the Southwest* **57**: 391–455.

Dolby GA, **Dorsey RJ**, **Graham MR. 2019**. A legacy of geo-climatic complexity and genetic divergence along the lower Colorado River: insights from the geological record and 33 desert-adapted animals. *Journal of Biogeography* **46**: 2479–2505. https://doi.org/10.1111/jbi.13685

Eaton DAR, **Overcast I**. **2020**. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* **36**: 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797. https://doi.org/10.1093/nar/gkh340

Edwards SV, **Xi Z**, **Janke A**, **Faircloth BC**, **McCormack JE**, **Glenn TC**, **Zhong B**, **Wu S**, **Lemmon EM**, **Lemmon AR**, **Leaché AD**, **Liu L**, **Davis CC. 2016**. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution* **94**: 447–462. https://doi.org/10.1016/j.ympev.2015.10.027

Flouri T, **Jiao X**, **Rannala B**, **Yang Z. 2018**. Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution* **35**: 2585–2593. https://doi.org/10.1093/molbev/msy147

Fujita MK, **Leaché AD**, **Burbrink FT**, **McGuire JA**, **Moritz C. 2012**. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* **27**: 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Gottscho AD, **Wood DA**, **Vandergast AG**, **Lemos-Espinal J**, **Gatesy J**, **Reeder TW. 2017**. Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado desert: delimiting species in the presence of gene flow. *Molecular Phylogenetics and Evolution* **106**: 103–117. https://doi.org/10.1016/j.ympev.2016.09.008

Grismer LL. 1994. The origin and evolution of the peninsular herpetofauna of Baja California, México. *Herpetological Natural History* **5**: 51–106.

Grismer LL. 2002. *Amphibians and reptiles of Baja California, including its Pacific islands and the islands in the Sea of Cortés*. Berkeley: University of California Press.

Grismer LL, Beaman KR, Lawler HE. 1995. *Sauromalus hispidus*. *Catalogue of American Amphibians and Reptiles* **615**: 1–4.

Gronau I, **Hubisz MJ**, **Gulko B**, **Danko CG**, **Siepel A. 2011**. Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics* **43**: 1031–1034. https://doi.org/10.1038/ng.937

Grummer JA, **Bryson RW**, **Reeder TW. 2014**. Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology* **63**: 119–133. https://doi.org/10.1093/sysbio/syt069

Harrington SM, **Hollingsworth BD**, **Higham TE**, **Reeder TW. 2017**. Pleistocene climatic fluctuations drive isolation and secondary contact in the red diamond rattlesnake (*Crotalus ruber*) in Baja California. *Journal of Biogeography* **45**: 64–75. https://doi.org/10.1111/jbi.13114

Hoang DT, **Chernomor O**, **von Haeseler A**, **Minh BQ**, **Vinh LS. 2018**. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* **35**: 518–522. https://doi.org/10.1093/molbev/msx281

Hollingsworth BD. 1998. The systematics of chuckwallas (*Sauromalus*) with a phylogenetic analysis of other iguanid lizards. *Herpetological Monographs* **12**: 38–191. https://doi.org/10.2307/1467020

Huang H, **Knowles LL. 2016**. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology* **65**: 357–365. https://doi.org/10.1093/sysbio/syu046

Hudson DH, **Bryant D. 2006**. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254–267.

Kalyaanamoorthy S, **Minh BQ**, **Wong TKF**, **von Haeseler A**, **Jermiin LS. 2017**. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**: 587–589. https://doi.org/10.1038/nmeth.4285

Klauber LM. 1947. Classification and ranges of the gopher snakes of the genus *Pituophis* in the western United States. *Bulletin of the Zoological Society of San Diego* **22**: 7–81. https://doi.org/10.1093/molbev/msj030

Lamb T, **Jones TR**, **Avise JC. 1991**. Phylogeographic histories of representative herpetofauna of the southwestern U.S.: mitochondrial DNA variation in the desert iguana (*Dipsosaurus dorsalis*) and the chuckwalla (*Sauromalus obesus*). *Journal of Evolutionary Biology* **5**: 465–480. https://doi.org/10.1046/j.1420-9101.1992.5030465.x

Leaché AD, **Banbury BL**, **Felsenstein J**, **de Oca AN**, **Stamatakis A. 2015**. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology* **64**: 1032–1047. https://doi.org/10.1093/sysbio/syv053

Leaché AD, **Fujita MK**, **Minin VN**, **Bouckaert RR. 2014**. Species delimitation using genome-wide SNP data. *Systematic Biology* **63**: 534–542. https://doi.org/10.1093/sysbio/syu018

Leaché AD, **Mulcahy DG. 2007**. Phylogeny, divergence times and species limits of spiny lizards (*Sceloporus magister* species group) in western North American deserts and Baja California. *Molecular Ecology* **16**: 5216–5233. https://doi.org/10.1111/j.1365-294X.2007.03556.x

Leaché AD, **Oaks JR. 2017**. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* **48**: 69–84. https://doi.org/10.1146/annurev-ecolsys-110316-022645

Leaché AD, **Zhu T**, **Rannala B**, **Yang Z. 2019**. The spectre of too many species. *Systematic Biology* **68**: 168–181. https://doi.org/10.1093/sysbio/syy051

Leavitt DH, **Hollingsworth BD**, **Fisher RN**, **Reeder TW. 2020**. Introgression obscures lineage boundaries and phylogeographic history in the western banded gecko, *Coleonyx variegatus* (Squamata: Eublepharidae). *Zoological Journal of the Linnean Society* **190**: 181–226. https://doi.org/10.1093/zoolinnean/zlz143

Leavitt DH, **Marion AB**, **Hollingsworth BD**, **Reeder TW. 2017**. Multilocus phylogeny of alligator lizards (*Elgaria*, Anguidae): testing mtDNA introgression as the source of discordant molecular phylogenetic hypotheses. *Molecular Phylogenetics and Evolution* **110**: 104–121. https://doi.org/10.1016/j.ympev.2017.02.010

Lovich RE, **Grismer LL**, **Danemann G. 2009**. Conservation status of the herpetofauna of Baja California, México and associated islands in the Sea of Cortez and Pacific Ocean. *Herpetological Conservation and Biology* **4**: 358–378.

Madeira F, **Park YM**, **Lee J**, **Buso N**, **Gur T**, **Madhusoodanan N**, **Basutkar P**, **Tivey ARN**, **Potter SC**, **Finn RD**, **Lopez R. 2019**. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47**: W636–W641. https://doi.org/10.1093/nar/gkz268

Malone CL, **Reynoso VH**, **Buckley L. 2017**. Never judge an iguana by its spines: systematics of the Yucatan spiny tailed iguana, *Ctenosaura defensor* (Cope, 1866). *Molecular Phylogenetics and Evolution* **115**: 27–39. https://doi.org/10.1016/j.ympev.2017.07.010

McGuire JA, **Linkem CW**, **Koo MS**, **Hutchison DW**, **Lappin AK**, **Orange DI**, **Lemos-Espinal J**, **Riddle BR**, **Jaeger JR. 2007**. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. *Evolution* **62**: 2879–2897. https://doi.org/10.1111/j.1558-5646.2007.00239.x

Meik JM, **Schaack S**, **Flores-Villela O**, **Streicher JW. 2018**. Integrative taxonomy at the nexus of population divergence and speciation in insular speckled rattlesnakes. *Journal of Natural History* **52**: 989–1016. https://doi.org/10.1080/00222933.2018.1429689

Minh BQ, **Nguyen MAT**, **von Haeseler A. 2013**. Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution* **30**: 1188–1195. https://doi.org/10.1093/molbev/mst024

Montanucci RR. 1987. A phylogenetic study of the horned lizard, genus *Phrynosoma*, based on skeletal and external morphology. *Contributions in Science, Natural History Museum of Los Angeles County* **390**: 1–36.

Montanucci RR. 2004. A note on the identity of Chuckwallas inhabiting Isla Danzante, Baja California Sur. *Herpetological Review* **35**: 223–224.

Montanucci RR. 2008. Historical evidence for the type locality of *Sauromalus ater* Duméril, 1856. *Herpetological Review* **39**: 326–328.

Mulcahy DG, **Macey JR. 2009**. Vicariance and dispersal form a ring distribution in nightsnakes around the Gulf of California. *Molecular Phylogenetics and Evolution* **53**: 537–546. https://doi.org/10.1016/j.ympev.2009.05.037

Murphy RW. 1983. Paleobiogeography and genetic differentiation of the Baja California herpetofauna. *Occasional Papers of the California Academy of Sciences* **137**: 1–48.

Nguyen LT, **Schmidt HA**, **von Haeseler A**, **Minh BQ. 2015**. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**: 268–274. https://doi.org/10.1093/molbev/msu300

Norell MA. 1989. Late Cenozoic lizards of the Anza Borrego Desert, California. *Natural History Museum of Los Angeles*

County, Contributions in Science **414**: 1–31. https://doi.org/10.5962/p.208135

Norell MA, **de Queiroz K. 1991**. The earliest iguanine lizard (Reptilia: Squamata) and its bearing on iguanine phylogeny. *American Museum Novitates* **2997**: 1–16.

Oskin M, **Stock J. 2003**. Marine incursion synchronous with plate-boundary localization in the Gulf of California. *Geology* **31**: 23–26. https://doi.org/10.1130/0091-7613(2003)031<0023:miswpb>2.0.co;2

Perry BW, **Card DC**, **McGlothlin JW**, **Pasquesi GIM**, **Adams RH**, **Schield DR**, **Hales NR**, **Corbin AB**, **Demuth JP**, **Hoffmann FG**, **Vandewege MW**, **Schott RK**, **Bhattacharyya N**, **Chang BSW**, **Casewell NR**, **Whiteley G**, **Reyes-Velasco J**, **Mackessy SP**, **Gamble T**, **Storey KB**, **Biggar KK**, **Passow CN**, **Kuo C-H**, **McGaugh SE**, **Bronikowski AM**, **de Koning APJ**, **Edwards SV**, **Pfrender ME**, **Minx P**, **Brodie ED III**, **Brodie ED Jr**, **Warren WC**, **Castoe TA. 2018**. Molecular adaptations for sensing and securing prey and insight into amniote genome diversity from the garter snake genome. *Genome Biology and Evolution* **10**: 2110–2129. https://doi.org/10.1093/gbe/evy157

Petren K, **Case TJ. 1997**. A phylogenetic analysis of body size evolution and biogeography in chuckwallas (*Sauromalus*) and other iguanines. *Evolution* **51**: 206–219. https://doi.org/10.1111/j.1558-5646.1997.tb02402.x

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

Petren K, **Case TJ. 2002**. An updated mtDNA phylogeny for *Sauromalus* and implications for the evolution of gigantism. In: Case TJ., Cody ML, Ezcurra E, eds. *A new island biogeography of the Sea of Cortés*. New York: Oxford University Press, 574–579. https://doi.org/10.1371/journal.pone.0037135

Pyron RA, **Burbrink FT**, **Wiens JJ. 2013**. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* **13**: 93. https://doi.org/10.1186/1471-2148-13-93

de Queiroz A, **Lawson R. 2008**. A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake *Thamnophis validus*. *Biological Journal of the Linnean Society* **95**: 409–424. https://doi.org/10.1111/j.1095-8312.2008.01049.x

de Queiroz KD. 2007. Species concepts and species delimitation. *Systematic Biology* **56**: 879–886. https://doi.org/10.1080/10635150701701083

Riddle BR, **Hafner DJ**, **Alexander LF**, **Jaeger JR. 2000**. Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 14438–14443. https://doi.org/10.1073/pnas.250413397

Robinson M. 1974. Chromosomes of the insular species of chuckwalla lizards (genus *Sauromalus*) in the Gulf of California, Mexico. *Herpetologica* **30**: 162–167.

Savage JM. 1960. Evolution of a peninsular herpetofauna. *Systematic Zoology* **9**: 184–212. https://doi.org/10.2307/2411967

Schmidt KP. 1922. The amphibians and reptiles of Lower California and the neighboring islands. *Bulletin of the American Museum of Natural History* **46**: 607–707.

Schmidt-Lebuhn AN, **Aitken NC**, **Chuah A. 2017**. Species trees from consensus single nucleotide polymorphism (SNP) data: testing phylogenetic approaches with simulated and empirical data. *Molecular Phylogenetics and Evolution* **116**: 192–201. https://doi.org/10.1016/j.ympev.2017.07.018

Shaw CE. 1945. The chuckwallas, genus *Sauromalus*. *Transactions of the San Diego Society of Natural History* **10**: 269–306.

Soulé ME, **Sloan AJ. 1966**. Biogeography and distribution of the reptiles and amphibians on islands in the Gulf of California, Mexico. *Transactions of the San Diego Society of Natural History* **14**: 137–156.

Sukumaran J, **Knowles LL. 2017**. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America* **114**: 1607–1612. https://doi.org/10.1073/pnas.1607921114

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

Sylber CK. 1985. Eggs and hatchlings of the yellow giant chuckwalla and the black giant chuckwalla in captivity. *Herpetological Review* **16**: 18–21.

Tracy. **1999**. Differences in Body Size among Chuckwalla (*Sauromalus obesus*) Populations. *Ecology* **80**: 259–271. https://doi.org/10.2307/176995

Tripp EA, **Tsai YHE**, **Zhuang Y**, **Dexter KG. 2017**. RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. *Ecology and Evolution* **7**: 7920–7936. https://doi.org/10.1002/ece3.3274

Upton DE, **Murphy RW. 1997**. Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. *Molecular Phylogenetics and Evolution* **8**: 104–113. https://doi.org/10.1006/mpev.1996.0392

Valdivia-Carrillo T, **García-De León FJ**, **Blázquez MC**, **Gutiérrez-Flores C**, **González Zamorano P. 2017**. Phylogeography and ecological niche modeling of the Desert Iguana (*Dipsosaurus dorsalis*, Baird & Girard 1852) in the Baja California Peninsula. *Journal of Heredity* **108**: 640–649. https://doi.org/10.1093/jhered/esx064

Weins JJ. 1993. Phylogenetic systematics of the tree lizards (genus *Urosaurus*). *Herpetologica* **49**: 399–420.

Welsh HH Jr. 1988. An ecogeographic analysis of the herpetofauna of the Sierra San Pedro Mátir Region, Baja California with a contribution to the biogeography of the Baja California herpetofauna. *Proceedings of the California Academy of Sciences* **46**: 1–72.

Wiens JJ, **Hollingsworth BD. 2000**. War of the iguanas: conflicting molecular and morphological phylogenies and long-branch attraction in iguanid lizards. *Systematic Biology* **49**: 143–159. https://doi.org/10.1080/10635150050207447

Wood DA, **Fisher RN**, **Reeder TW. 2008**. Novel patterns of historical isolation, dispersal and secondary contact across Baja California in the Rosy Boa (*Lichanura trivirgata*).

Molecular Phylogenetics and Evolution **46**: 484–502. https://doi.org/10.1016/j.ympev.2007.11.014

Wood DA, **Vandergast AG**, **Barr KR**, **Inman RD**, **Esque TC**, **Nussear KE**, **Fisher RN. 2013**. Comparative phylogeography reveals deep lineages and regional evolutionary hotspots in the Mojave and Sonoran Deserts. *Diversity and Distributions* **19**: 722–737. https://doi.org/10.1111/ddi.12022

Xi Z, **Liu L**, **Davis CC. 2016**. The impact of missing data on species tree estimation. *Molecular Biology and Evolution* **33**: 838–860. https://doi.org/10.1093/molbev/msv266

Zheng X, **Levine D**, **Shen J**, **Gogarten SM**, **Laurie C**, **Weir BS. 2012**. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**: 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article on the publisher's website.

Table S1. Individuals used for genomic DNA sequencing separated by sample identity, taxon and locality information.
Table S2. Individuals obtained from GenBank for dated *Cytb* phylogeny separated by accession number, sample identity, taxon and locality information.
Table S3. Results of the three G-PHOCS models tested in this study.
Figure S1. Quartet species tree based on double-digest restriction site-associated DNA sequencing data showing all putative *Sauromalus* species. The *Sauromalus* tree is rooted with *Iguana iguana*. Numbers indicate bootstrap support.
Figure S2. Bayesian time tree of the reduced set of individuals of *Sauromalus* and *Iguana iguana* using a relaxed clock based on concatenated double-digest restriction site-associated DNA sequencing data. Numbers at nodes indicate posterior probabilities, and shaded bars on nodes indicate 95% highest posterior density confidence intervals. Major groups and currently recognized taxa are labelled.
Figure S3. Bayesian time tree of reduced set of individuals of only *Sauromalus* using a relaxed clock based on concatenated double-digest restriction site-associated DNA sequencing data. Numbers at nodes indicate posterior probabilities, and shaded bars on nodes indicate 95% highest posterior density confidence intervals. Major groups and currently recognized taxa are labelled.
Figure S4. Bayesian time tree of individuals of *Sauromalus* based on *Cytb* data
obtained from GenBank (Supporting Information, Table S2). Numbers at nodes indicate posterior probabilities, and shaded bars on nodes indicate 95% highest posterior density confidence intervals. Major groups and currently recognized taxa are labelled.