Molecular Phylogenetics and Evolution 127 (2018) 669–681

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Cryptic genetic diversity, population structure, and gene flow in the Mojave rattlesnake (*Crotalus scutulatus*)

Drew R. Schield[a], Richard H. Adams[a], Daren C. Card[a], Andrew B. Corbin[a], Tereza Jezkova[b], Nicole R. Hales[a], Jesse M. Meik[c], Blair W. Perry[a], Carol L. Spencer[d], Lydia L. Smith[d], Gustavo Campillo García[e], Nassima M. Bouzid[d,f], Jason L. Strickland[g], Christopher L. Parkinson[g,h], Miguel Borja[i], Gamaliel Castañeda-Gaytán[i], Robert W. Bryson Jr.[f], Oscar A. Flores-Villela[e], Stephen P. Mackessy[j], Todd A. Castoe[a,*]

[a] Department of Biology, 501 S. Nedderman Drive, University of Texas at Arlington, Arlington, TX 76019, USA
[b] Department of Biology, 501 E. High Street, Miami University, Oxford, OH 45056, USA
[c] Department of Biological Sciences, Tarleton State University, 1333 W. Washington Street, Stephenville, TX 76402, USA
[d] Museum of Vertebrate Zoology, 3101 Valley Life Sciences Building, University of California, Berkeley, CA 94720, USA
[e] Museo de Zoología, Department of Evolutionary Biology, Facultad de Ciencias, Universidad Nacional Autónoma de México, External Circuit of Ciudad Universitaria, México City 04510, Mexico
[f] Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Box 351800, Seattle, WA 98195, USA
[g] Department of Biology, Biological Sciences Building, 4110 Libra Drive, University of Central Florida, Orlando, FL 32816, USA
[h] Department of Biological Sciences and Department of Forestry and Environmental Conservation, Clemson University, 190 Collings Street, Clemson, SC 29634, USA
[i] Facultad de Ciencias Biológicas, Universidad Juárez del Estado de Durango, Gómez Palacio, Durango, Mexico
[j] School of Biological Sciences, 501 20th Street, University of Northern Colorado, Greeley, CO 80639, USA

ARTICLE INFO

Keywords:
Gene flow
Population genomics
Population structure
RADseq
Secondary contact
Viperidae

ABSTRACT

The Mojave rattlesnake (*Crotalus scutulatus*) inhabits deserts and arid grasslands of the western United States and Mexico. Despite considerable interest in its highly toxic venom and the recognition of two subspecies, no molecular studies have characterized range-wide genetic diversity and population structure or tested species limits within *C. scutulatus*. We used mitochondrial DNA and thousands of nuclear loci from double-digest restriction site associated DNA sequencing to infer population genetic structure throughout the range of *C. scutulatus*, and to evaluate divergence times and gene flow between populations. We find strong support for several divergent mitochondrial and nuclear clades of *C. scutulatus*, including splits coincident with two major phylogeographic barriers: the Continental Divide and the elevational increase associated with the Central Mexican Plateau. We apply Bayesian clustering, phylogenetic inference, and coalescent-based species delimitation to our nuclear genetic data to test hypotheses of population structure. We also performed demographic analyses to test hypotheses relating to population divergence and gene flow. Collectively, our results support the existence of four distinct lineages within *C. scutulatus*, and genetically defined populations do not correspond with currently recognized subspecies ranges. Finally, we use approximate Bayesian computation to test hypotheses of divergence among multiple rattlesnake species groups distributed across the Continental Divide, and find evidence for co-divergence at this boundary during the mid-Pleistocene.

1. Introduction

The drivers of differentiation between populations and the subsequent divergence of once panmictic lineages are of considerable interest and debate (Bolnick and Fitzpatrick, 2007; Coyne and Orr, 2004; Feder et al., 2012). Widespread organisms, particularly those with low vagility, are particularly well-suited for studies of speciation, as their populations are more likely to be subjected to disruptions (e.g., geological, biological, anthropogenic) that lead to reductions in gene flow and population divergence. Thus, examining the population genetic structure of widely distributed lineages that occupy diverse habitats and climatic conditions is useful for improving our understanding of environmental features and historical processes that drive diversification. Organisms inhabiting the North American warm desert regions of

* Corresponding author at: Department of Biology, University of Texas at Arlington, 501 S. Nedderman Drive, 337 Life Science, Arlington, TX 76010-0498, USA.
E-mail address: todd.castoe@uta.edu (T.A. Castoe).

https://doi.org/10.1016/j.ympev.2018.06.013
Received 10 July 2017; Received in revised form 30 May 2018; Accepted 6 June 2018
Available online 15 June 2018
1055-7903/

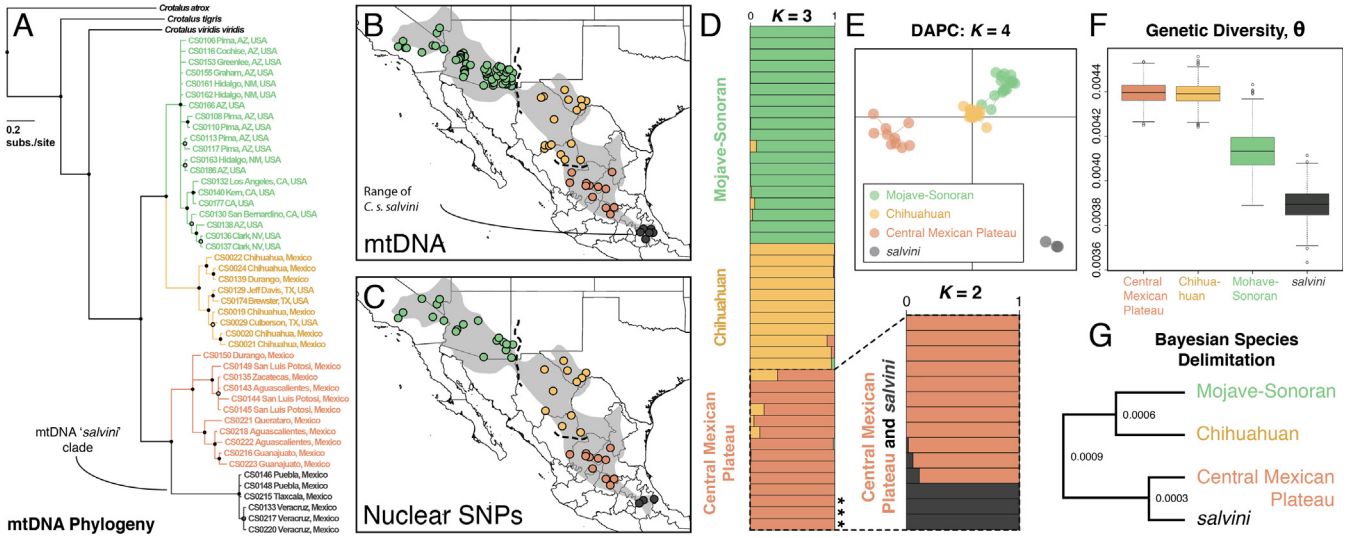

Fig. 1. Mitochondrial and nuclear evidence for phylogenetic and population structure across the range of *C. scutulatus*. A. Bayesian mitochondrial gene (ND4) tree depicting four distinct clades, with sample identifiers and general locality descriptions provided. B. Map of sampling localities used in mtDNA analyses; dots are colored according to mitochondrial clade (the current range of *C. s. salvini* is indicated). Dashed lines represent the Continental Divide (northern) and elevational increase of the Mexican Plateau (southern) biogeographic barriers. C. Sampling localities used in nuclear analyses; dots are colored according to major nuclear clades. The range of *C. scutulatus* shaded in gray on each map was redrawn from spatial data available from the IUCN. D. Bayesian population clustering models from STRUCTURE and phylogenetic tree (based on concatenated RAD loci). Horizontal bars on depict the posterior probability of assignment to one or more population clusters, colored to correspond with mtDNA/nuclear trees. Samples from the currently recognized range of *C. s. salvini* are marked with asterisks. The inset to the right shows the results of the fine scale analysis of Central Mexican Plateau and *salvini* populations. E. Results of population clustering analysis in DAPC based on a clustering scheme of $K = 4$. F. Posterior distributions of θ estimated using ThetaMater. Whiskers on the boxplots depict the bounds of the 95% HPD, boundaries of the boxes show the interquartile range, and the black line shows the median posterior estimate per population. G. The best-supported topology from Bayesian species delimitation analysis in bpp. Values at nodes represent relative branch lengths.

the southwestern United States and north-central Mexico have become model systems in this regard, and multiple studies have demonstrated repeated patterns of diversification associated with historical and current habitat and climate variation in this region (Bryson et al., 2011a, 2011b; Cox et al., 2012; Jaeger et al., 2005; Riddle et al., 2000).

Here, we studied the population genetic structure of the Mojave rattlesnake (*Crotalus scutulatus*). This species is widely distributed throughout the Mojave, Sonoran, and Chihuahuan Deserts and associated grasslands of the western United States and Mexico (Klauber, 1956), and is thus also a useful model for understanding patterns of diversification across the North American warm desert region. These rattlesnakes primarily inhabit dry scrubland habitat in flat terrain, but also occur in montane woodlands near the southern end of their distribution (Campbell and Lamar, 2004), and occupy elevations from sea level to over 2100 m. *Crotalus scutulatus* is considered among the most dangerous North American pitvipers due to its aggressive disposition, high venom yield, and highly neurotoxic venom (Cate and Bieber, 1978; Hardy, 1983; Mackessy, 2008; Sanchez et al., 2005). For these reasons, *C. scutulatus* venom has been the subject of numerous studies (e.g., (Durban et al., 2013; Glenn and Straight, 1978; Glenn et al., 1983; Massey et al., 2012; Smith and Mackessy, 2016). These studies have largely sought to characterize populations with and without expression of neurotoxic venom compounds. Despite efforts to understand the toxicity and geographic variation of *C. scutulatus* venom, remarkably little is known regarding the overall genetic structure and variation throughout the large range of this species.

Crotalus scutulatus has been included in several broad phylogenetic analyses of pitvipers, and has been well supported as the sister lineage to the Western Rattlesnake (*Crotalus viridis* + *C. oreganus*) species complex (Castoe and Parkinson, 2006; Murphy et al., 2002; Reyes-Velasco et al., 2013), from which it diverged during the Pliocene (Reyes-Velasco et al., 2013). Two phenotypically distinct subspecies of *C. scutulatus* are currently recognized: the Mojave Rattlesnake (*C. s. scutulatus*) and the Huamantlan Rattlesnake (*C. s. salvini*) (Campbell and Lamar, 2004). The majority of the range of *C. scutulatus* is occupied by the northern subspecies *C. s. scutulatus,* which ranges from southern California to central Mexico. The southern subspecies *C. s. salvini* has a comparatively smaller range in the south-central Mexican states of Hidalgo, Tlaxcala, Estado de México, Puebla, and Veracruz. Although Klauber (1956) studied intraspecific variation in *C. scutulatus* morphology, no studies have assessed genetic relationships among populations and subspecies. Considerable cryptic diversity within the species is plausible given the large geographic range and distinctive ecoregions inhabited by *C. scutulatus*. Given the large range, medical relevance, and absence of any molecular studies of intraspecific variation of *C. scutulatus*, a range-wide molecular appraisal of the species would also serve as a valuable foundation for understanding regional biogeography as well as venom variation in this lineage.

Here we evaluate patterns of genetic diversity and structure of *C. scutulatus* throughout its range and provide the first molecular-based inferences of *C. scutulatus* population genetic structure using a combined dataset of a single mitochondrial gene and thousands of nuclear loci. We test the hypothesis that *C. s. scutulatus* and *C. s. salvini* are distinct evolutionary lineages and explore potential cryptic diversity within *C. scutulatus* that is not recognized by current subspecies taxonomy. We also estimate population structure and gene flow between populations and characterize biogeographic and demographic patterns within lineages of *C. scutulatus*. More broadly, we test the hypothesis that the Continental Divide represents a partial barrier to gene flow between subpopulations of *C. s. scutulatus*, as has been observed in the closely related Western Diamondback Rattlesnake (*C. atrox*; (Castoe et al., 2007; Schield et al., 2015) and other snake taxa (Myers et al., 2016). *Crotalus scutulatus* also traverses a breadth of topographies and ecoregions in Mexico, including the elevated Central Mexican Plateau, and we also tested for divergence and gene flow across this known biogeographic barrier. Overall our results highlight substantial and previously unappreciated genetic structure and lineage diversity within *C. scutulatus* that is valuable for interpreting data on venom variation, snakebite treatment, and ultimately will help guide future taxonomic revision of this group.

2. Materials and methods

2.1. Taxon sampling and DNA extraction

We obtained tissue samples from 141 *Crotalus scutulatus* specimens, including representatives of both recognized subspecies (Fig. 1; Supplementary Table 1). We sampled specimens from diverse localities to maximize our ability to estimate population structure and genetic diversity range-wide. Skin or liver tissue was dissected and snap frozen in liquid nitrogen, or stored in DNA lysis buffer or in ethanol. Non-lethal caudal punctures were used to obtain blood tissue from some specimens, and blood was snap-frozen in liquid nitrogen. We isolated genomic DNA from blood and whole tissue using phenol-chloroform-isoamyl alcohol and NaCl-isopropanol precipitation extractions (Miller et al., 1988).

2.2. Mitochondrial and nuclear DNA sequence generation

We used PCR to amplify a fragment of the mitochondrial NADH4 (ND4) gene and downstream tRNAs from all samples, using the primers ND4 and Leu (Arevalo et al., 1994). PCR products were purified using Serapure beads and were quantified using gel electrophoresis and a Qubit Fluorometer (Life Technologies, Grand Island, NY, USA). Purified PCR products were sequenced in both directions using amplification primers and BigDye on an ABI 3730 capillary sequencer (Life Technologies), and ND4 sequences for outgroup taxa were retrieved from our previous studies (Castoe et al., 2007; Schield et al., 2015) or from GenBank (see Supplementary Table 1 for accession numbers).

We chose a subset of *C. scutulatus* sequenced for mitochondrial markers to generate double-digest RADseq libraries (n = 48), using the Peterson et al. (2012) protocol with minor adjustments detailed in (Schield et al., 2015). In brief, we began with 0.5 μg genomic DNA per sample and performed digestions overnight at 37 °C using rare-cutting (*SbfI*) and common-cutting (*Sau3AI*) restriction enzymes. Digested samples were purified with Ampure (Invitrogen) beads; purified products were eluted in 40 μL of TE and quantified using a Qubit Fluorometer. Based on these quantifications, we organized samples by concentration, divided samples into groups of six to eight, and standardized input DNA for ligation reactions based on the lowest concentration among within-group samples (these were variable). We then ligated double-stranded indexed DNA adapters (including 8 bp unique molecular identifier regions; UMIs), and pooled samples into their respective groups, with each sample having a specific barcoded adapter. Pooled samples were purified and size-selected using a 1.5% agarose cassette on a Blue Pippin Prep (Sage Science) for fragments within a range of 575–655 bp. Based on *in silico* digestion and size selection using the Burmese Python genome (Castoe et al., 2013), we estimated this protocol would target ~20,000 loci. Size selected samples were amplified using indexed primers to provide a second index specific to each pooled sample. Indexed pools were mixed in equimolar ratios and were sequenced together using 100 bp single-end reads on an Illumina HiSeq 2500.

2.3. Estimates of mitochondrial gene phylogeny, diversity, and demography

Raw mitochondrial ND4 sequence chromatograms were edited and consensus sequences were generated using Geneious v6.1.6 (Biomatters Ltd., Aukland, NZ). Sequences were aligned using MUSCLE v3.8 (Edgar, 2004) and trimmed manually to remove spurious and potentially erroneous low quality base calls. The final ND4 alignment consisted of 808 bases for all individuals, including outgroup taxa, with no indels. We estimated mitochondrial relationships within *C. scutulatus* using MrBayes v3.2.1 (Huelsenbeck and Ronquist, 2001), after inferring best-fit models of evolution for partitioned 1st, 2nd, and 3rd codon positions using the Bayesian Information Criterion and the Greedy algorithm implemented in PartitionFinder v1.1.1 (Lanfear et al., 2012). The best-fit partitioning scheme was an independent partition for each of the three codon positions (HKY + Γ for 1st and 3rd codon positions, F81 + Γ +I for 2nd codon positions); this partitioned model was used in all MrBayes runs. Two initial trial runs were performed using our total dataset, including 141 *C. scutulatus* individuals and four outgroup taxa to identify unique haplotypes. We then condensed the alignment to single representatives of each haplotype and performed four independent runs of MrBayes on this condensed dataset, each consisting of 10^7 generations with samples taken every 500 generations and four MCMC chains (one cold and three heated). We discarded the first 25% of MCMC samples in each run as burnin, as the potential scale reduction factor (PSRF) values, parameter estimates, and marginal likelihood estimates indicated that runs had converged by this period, which we assessed using Tracer v1.6 (Drummond and Rambaut, 2007). Effective sample sizes exceeded 2000 for all parameters indicating that the posterior probability distribution had been effectively sampled in each run. We generated a 50% majority rule tree using combined tree estimates from post-burnin samples.

We estimated mitochondrial genetic diversity for populations using the average within-group pairwise distance, nucleotide diversity (π), Watterson's estimator ($\hat{\theta}$), and haplotype diversity calculated in DNAsp v5 (Librado and Rozas, 2009). Here, we refer to mitochondrial clades and 'populations' interchangeably; spatial organization of mitochondrial clades was consistent with populations in distinct geographic regions (see Section 3). The populations included in these analyses mirrored the mitochondrial gene tree (see Fig. 1) and included four populations: (i) a northern population consisting of samples from Nevada, California, Arizona, and New Mexico ('Mojave-Sonoran' hereafter), (ii) a 'Chihuahuan' population including samples from Texas, Chihuahua, Coahuila, and northern Durango, (iii) a population occupying higher elevation regions of the central Mexican Plateau ('Central Plateau' hereafter), and (iv) a population from Puebla, Tlaxcala, and Veracruz, which represents the recognized range of *C. s. salvini*. For simplicity, we refer to this population as '*salvini*'. We also tested for patterns of allelic diversity in each population consistent with population expansion using Fu's *Fs* (Fu, 1997). We obtained p-values for *Fs* by performing 1000 coalescent simulations and computing the probability of observing values equal to or less than each simulated value.

2.4. Nuclear analyses of population structure, demography, and phylogeny

We processed raw Illumina sequencing reads first using the Stacks v1.35 clone_filter module (Catchen et al., 2013) to remove PCR clones using the UMI regions included in our sequence library adapters. We then trimmed the 8 bp UMIs from reads using the Fastx-Toolkit trimmer (Hannon, 2014). Trimmed reads were demultiplexed into individual samples using the Stacks process_RADtags function, which also trimmed the 6 bp barcode region of each read. Demultiplexed and trimmed read data are available from the NCBI SRA (accession SRP150237). We performed *de novo* assembly and variant calling using the Stacks ustacks, cstacks, and sstacks modules, and used the downstream rxstacks filter module to retain only biologically plausible loci. We then used the populations module to calculate heterozygosity and private alleles and generate sequence alignments by parsing the dataset to the individual level. We also performed pairwise analyses of allelic differentiation (F_{ST}) between populations to evaluate population structure. We filtered data to retain only loci that had at least 5x read depth per individual and for which data were present in at least 30 samples (Stacks options 'm' = 5 and 'p' = 30).

We used STRUCTURE (Pritchard et al., 2000) to estimate broad scale population structure and admixture within our nuclear SNP dataset. Initial runs to determine a reduced range of likely models to explain the data were performed using a *K* range of 1–10 clusters, and we performed 3 replicate runs per model of *K* with a burnin of 10,000 and 10,000 sampled MCMC generations. The results of preliminary runs

were processed using Structure Harvester (Earl and Vonholdt, 2012), from which we determined the best-fit model to be $K = 3$ using the Evanno method (Evanno et al., 2005). We then performed longer triplicate runs across a smaller range of $K = 2$–5, with a 10,000 generation burnin and 100,000 sampled generations, and assessed model likelihoods in Structure Harvester. We also performed a secondary STRUCTURE analysis focused only on Central Mexican Plateau and *salvini* samples ($K = 2$) to discern if nuclear structure exists between these populations that is not detected when all populations are analyzed simultaneously. To complement STRUCTURE analyses and to investigate evidence of genetic structure without model assumptions, we performed Discriminant Analysis of Principal Components (DAPC; (Jombart et al., 2010)). A benefit of this approach is that it can be sensitive to fine scale structure between populations, even when they share substantial gene flow (Jombart et al., 2010). For DAPC analyses, we first performed *K*-means clustering of the dataset specifying a possible 30 genetic clusters and estimated likelihoods over 1,000,000 iterations. We then evaluated model likelihoods using Bayesian Information Criterion (BIC), as recommended by Zhao (2006) and (Lee et al., 2009) to determine the best-supported number of clusters.

To test the mtDNA-based hypothesis of four major clades within *C. scutulatus* (see Section 3), we conducted Bayesian species delimitation using bpp v.3.4 (Yang and Rannala, 2010) with 922 randomly sampled loci from our RADseq dataset (to make analyses computationally tractable). We set the prior distributions for the tree height parameter (τ) and effective population size parameters (θ) based on our mtDNA phylogeny and the average pairwise genetic distance (π) measured across our entire RADseq dataset. We set the shape and scale parameters of the gamma prior distribution on θ to 1.2 and 1.0, respectively, which reflects an expected θ of 0.12 (average $\pi = 0.12$). For the prior on τ, we set the shape and scale parameters to 1.6 and 10000, respectively, to reflect the tree height for these four clades that was estimated from our mtDNA phylogeny (distance to ancestral node = 0.00016). We also set the starting topology to match the mtDNA topology ((Mojave-Sonoran, Chihuahuan), (Central Mexican Plateau, *salvini*)) and conducted unguided species delimitation using prior algorithm0. We ran the MCMC chain for 110,000 iterations, sampling every 5th iteration, and discarded the first 10,000 iterations as burnin.

We estimated effective population sizes from our RADseq data by first estimating the population genetic diversity parameter θ (equal to $4N_e\mu$) for each population using ThetaMater (Adams et al., 2017). ThetaMater uses an infinite-sites likelihood function to simulate the posterior probability of θ based on the probability of observing k segregating sites in a sample size of n sequences obtained from a locus of size l. For each of the four populations identified from mtDNA analyses (see above and Section 3), we used the Stacks populations module to output sequences of RAD loci that were present in at least two thirds of population samples, and used ThetaMater's 'Read.AllelesFile' function to convert RAD loci into ThetaMater input. We ran the 'ThetaMater.M1' MCMC simulation model to estimate the posterior θ distribution for each population for a total of 1×10^6 generations and discarded 10% of generations as burn in per population. We then derived an estimate of effective population size (N_e) by dividing the median of the posterior distribution of θ per population by the generation time of 3 years (a general estimate for rattlesnakes; (Castoe et al., 2007)), and assuming a previously estimated rate for squamate 4-fold degenerate sites (2.4×10^{-9}) as the substitution rate (Green et al., 2014).

We also estimated relationships among individuals using Bayesian MCMC inference on a concatenated nuclear dataset in BEAST v2.1.3 (Bouckaert et al., 2014). While useful as a first-pass analysis of relationships, we note that this approach must be interpreted with caution because gene flow, incomplete lineage sorting, and other processes may substantially mislead inferences from such concatenated analyses (Kubatko and Degnan, 2007; Mendes and Hahn, 2017; Roch and Steel, 2015). Analyses were run using a GTR + Γsubstitution model, and we ran two independent BEAST runs for a total of 50 x 10^6 generations each, and discarded the first 25% (12.5 $\times$ 10^6 generations) as burn-in based on likelihood stationarity and parameter effective sample sizes (ESS) analyzed in Tracer v1.6 (Drummond and Rambaut, 2007). We combined the two runs into a single posterior distribution consisting of 75 x 10^6 MCMC samples and generated a maximum clade credibility tree using TreeAnnotater (Bouckaert et al., 2014) based on mean heights, which is shown in Supplementary Fig. 1.

We tested for evidence of gene flow between lineages of *C. scutulatus* using Patterson's D statistics in the ABBA-BABA framework (Durand et al., 2011), which quantify shared derived alleles that are expected to arrive via introgression rather than incomplete lineage sorting. We performed two analyses that tested for introgression across two biogeographic barriers: the Continental Divide and the elevational increase associated with the Central Mexican Plateau. For both analyses, we used the four-population 'CalcD' algorithm implemented in the R package evobiR (https://cran.r-project.org/web/packages/evobiR/index.html; see Streicher et al. (2014) for additional details). The 'Continental Divide' analysis included Central Mexican Plateau *C. scutulatus* as the outgroup and the Chihuahuan group as the first ingroup. We then split the Mojave-Sonoran group into two subclades consisting of samples from New Mexico and eastern Arizona ('Mojave-Sonoran B') and samples from western Arizona and California ('Mojave-Sonoran A'; Fig. 2). We generated consensus sequences for each population using Stacks, specifying a minimum read depth of five per locus per individual ('−m' = 5), requiring each population to contain data for a given SNP locus ('−p' = 4), and requiring each SNP to be represented by at least 75% of individuals within a population ('−r' = 0.75). These parameters resulted in 6,058 SNPs for the 'Continental Divide' analysis. The 'Central Mexican Plateau' analysis required an outgroup outside of *C. scutulatus* to test for gene flow between the Central Mexican Plateau and Chihuahuan populations, so we generated RADseq data from four individual *C. viridis* (Supplementary Table 1), the sister lineage to *C. scutulatus*, following the same library preparation and sequencing protocols. The Stacks parameters detailed above yielded 5077 SNPs for the 'Central Mexican Plateau' analysis. Schematic representations and

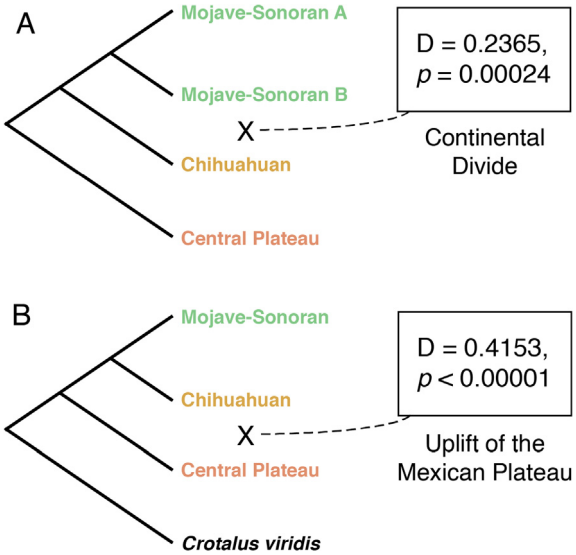

Fig. 2. Schematic design and results of ABBA-BABA tests using Patterson's D statistics for introgression across biogeographic barriers. (A) Tree topology used to test for introgression across the Continental Divide and D-statistic information. (B) Tree topology used to test for introgression across the uplift of the Mexican Plateau, including *C. viridis* as outgroup, and information about the resulting D-statistic. Colors used to label the *C. scutulatus* tree tips in these schematics correspond with the nuclear tree colors in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

details of each analysis are provided in Fig. 2.

We used *δaδi* to further estimate the demographic histories of *C. scutulatus* lineages, including modes and timing of divergence and migration rates between populations. We tested eight competing two-population divergence models (including a model of no divergence between populations; Supplementary Table 1) for parallel pairwise analyses of Mohave-Sonoran + Chihuahuan, Chihuahuan + Central Mexican Plateau, and Central Mexican Plateau + *salvini* populations. These models were also used previously to estimate demographic parameters for populations of *Crotalus atrox* (Schield et al., 2017), and the *δaδi* pipeline and associated scripts used here are described in detail in Portik et al. (2017). We generated input files using the Stacks output described above with python scripts available from https://github.com/dportik/Stacks_pipeline. We then generated the folded 2D allele frequency spectrum (AFS) for each of the three population pairs listed above, downsampling a minimum of 10 alleles from each population to minimize missing data among loci, with the exception of *salvini*, from which we could only sample a maximum of 6 alleles. We then fit each of the eight alternative demographic models to each 2D AFS. We used the Nelder-Mead method to optimize 20 sets of randomly perturbed parameters for a maximum of 50 iterations, followed by two rounds of optimization using the highest scoring parameter estimates per model from the previous round. We then simulated the 2D AFS for each parameter set and performed extrapolation using a grid size of [40,50,60]. We then estimated log-likelihoods of model fit using the multinomial approach and assessed model fit using the Akaike Information Criterion (AIC) based on log likelihood of the highest scoring replicate.

2.5. Inferences of historical biogeography and population expansion

We estimated the distribution of *C. scutulatus* at the height the Last Glacial Maximum (LGM) during the Pleistocene. We were specifically interested to determine if patterns of genetic divergence and structure observed in our datasets were consistent with predictions of historical geographic isolation during Pleistocene glacial conditions (represented by LGM conditions). We obtained geographic coordinates from museum records for 740 *C. scutulatus* specimens from Vertnet (www.vertnet.org) and combined these with coordinates from our genetic dataset. We then used ecological niche modeling (ENM) to estimate the present-day range of *C. scutulatus*, as implemented in MAXENT v3.3.3k (Phillips et al., 2006), which extracts environmental data associated with occurrence records and predicts the suitability of climatic conditions within a given range using prediction algorithms (Elith et al., 2006).

Because occurrence records were biased towards heavier sampling in the United States than in Mexico, we subsampled occurrence records to account for potential biases in ENM analysis, setting a minimum distance between samples of 0.1°, 0.2°, 0.3°, and 0.4°; we then determined if subsampling schemes quantitatively altered ENM results. The environmental data were represented by temperature and precipitation variables from the WorldClim dataset v1.4 with a 2.5-min resolution (Hijmans et al., 2005). Following the methodology of Jezkova et al. (2011), we removed 8 highly correlated variables (i.e., correlation coefficient > 0.9), resulting in the selection of 11 climatic variables (Supplementary Table 5), and we confined models to the southwestern region of North America. We used default MAXENT parameters and used cross-validation as a replicated run type. We ran 20 replicates for each model, determined an average model using logistic probability classes of climatic niche suitability, and visualized this model in ArcGIS v9.2 (ESRI). We used the receiver-operating characteristic (ROC) to determine an area under the curve (AUC) value to evaluate model performance, which range from 0.5 (i.e., random prediction) to 1 (perfect prediction; (Raes and ter Steege, 2007). We then projected present-day models onto reconstructions of the LGM, with the assumption that the climatic niche of *C. scutulatus* has not changed between the LGM and present day (Elith et al., 2010). We used three ocean-atmosphere simulation models for environmental layers representing the LGM: CCSM4, MIROC-ESM, and MPI-ESM-P (www.worldclim.org). To validate the use of default MAXENT parameters, we also ran analyses by first performing model testing and selection, which we assessed using the R package ENMeval (Muscarella et al., 2014). The best-supported model was LQHP 3.5, and the results of predictions based on this model matched those predicted using default MAXENT parameters.

Population genetic theory predicts that with range expansion, populations at or near to the expanding range-front will harbor lower genetic diversity (i.e., heterozygosity) than the ancestral population (Mayr, 1942; Slatkin and Excoffier, 2012), and we have observed this pattern studying RADseq loci in several empirical systems (Jezkova et al., 2015; Schield et al., 2015; Streicher et al., 2016). We tested the hypothesis that populations of *C. scutulatus* have experienced range expansion out of an ancestral region using linear models implemented in R (R Core Team, 2017). Specifically, we evaluated whether estimates of individual heterozygosity and private alleles correlated with latitude and longitude.

2.6. Divergence dating and tests of co-divergence

We analyzed an expanded sample of *Crotalus* lineages using BEAST 2 v2.4 (Bouckaert et al., 2014) to estimate divergence times for major splits between *C. scutulatus* lineages, and to test for evidence of co-divergence (i.e., correlated divergence events among taxa) of two or more distinct lineages across a *C. scutulatus* + *C. atrox* phylogeny, particularly between populations adjacent to the Continental Divide. The expanded dataset used for these analyses included all *C. scutulatus* ND4 sequences (condensed to unique haplotypes), as well as ND4 sequences of western diamondback rattlesnakes (*C. atrox*), and several outgroup taxa (*C. ruber*, *C. horridus*, *C. molossus*, *Agkistrodon contortrix*; Supplementary Table 1). Sequences were aligned using MUSCLE and bases were trimmed at each end of the alignment to reduce missing data. The final alignment consisted of 807 bases. We estimated substitution models and partition assignments using PartitionFinder v1.1.1 (Lanfear et al., 2012), and partitioned the final alignment independently for each codon position. We applied an HKY model of sequence evolution to first and second codon position partitions and a TN93 model for the thirds codon position. To calibrate the tree we specified four clade constraints based on previous divergence estimates for pitvipers. The ancestor of *Agkistrodon*, the most distant outgroup used in this analysis, was constrained at 6 MYA, with a specified mean of 0.01 and standard deviation (SD) of 0.42. The ancestor of *Sistrurus*, the more closely related outgroup to *Crotalus*, was constrained with an offset of 8 MYA, a mean of 0.01, and SD of 0.76. These priors were specified using lognormal prior distributions and constraints according to Reyes-Velasco et al. (2013). We constrained the ancestral node for *C. atrox* + *C. ruber* with an offset of 3.2 MYA following Castoe et al. (2007), and used a normal distribution with mean of 0 and SD of 1. Finally, we constrained the ancestral node for all rattlesnakes (i.e., *Crotalus* + *Sistrurus*), setting a normal distribution with an offset of 11.2 MYA, mean of 0, and SD of 3, following Reyes-Velasco et al. (2013). We ran two BEAST 2 runs for 5×10^9 generations each and evaluated effective sample sizes, stationarity, and run convergence using Tracer. We discarded the first 10% of generations as burnin, combined estimates from both runs, and summarized parameter estimates on a maximum clade credibility tree using LogCombiner and TreeAnnotator v2.1.3 (Bouckaert et al., 2014).

We tested for evidence of simultaneous divergence of multiple lineages using the hierarchical approximate Bayesian computation (hABC) algorithm used in MTML-msBayes (Huang et al., 2011). Here, we specifically tested the hypothesis that divergence events of *C. scutulatus* and *C. atrox* populations across the Continental Divide were coincident. We prepared two inputs of lineage pairs: (1) Mojave-Sonoran + Chihuahuan *C. scutulatus* and western + eastern *C. atrox*.

Using these inputs, we simulated 1,000,000 randomly drawn hyper-parameters and summary statistics and estimated the posterior density for numbers of possible divergence times, Ψ (i.e., 1–2), specifying a proportion of acceptable draws from the prior ('-t' option) to 0.005 (i.e., 500 draws from simulated priors).

3. Results

3.1. Mitochondrial gene phylogeny and genetic diversity

The 50% majority rule consensus phylogeny (Fig. 1A) of 46 unique *C. scutulatus* haplotypes revealed strong support (i.e., > 0.95 posterior probability) for *C. scutulatus* as monophyletic with respect to outgroup taxa, and for four distinct clades within *C. scutulatus*. The first major split within *C. scutulatus* was inferred as the divergence between two northern (Mojave-Sonoran and Chihuahuan) and two central/southern Mexico clades (Central Mexican Plateau and *salvini*). This split corresponds to the geographic break occurring in central Mexico (Fig. 1B and C), located along the northern edge of the Central Mexican Plateau where elevation increases by ~1000 m; samples from northern and southern Durango fall into northern and southern clades, respectively. We find evidence for two major subclades within each of these larger northern and southern clades (Fig. 1A). The major split within the northern clade segregates samples east and west of the Continental Divide (analogous to patterns observed in *C. atrox*; (Castoe et al., 2007; Schield et al., 2015)). The split within the southern clade segregates samples from central and southern Mexican localities from samples in the known geographic range of *C. s. salvini* (Fig. 1B) (Campbell and Lamar, 2004).

Mitochondrial genetic diversity was greatest in the Central Mexican Plateau and Chihuahuan populations (Table 1), with lower estimates in the Mojave-Sonoran population, and very low estimates in *salvini*. This pattern was consistent across different measures of genetic diversity and our estimates of population size, with the exception of haplotype diversity. Here, both the Mojave-Sonoran and Chihuahuan populations had greater haplotype diversity than the Central Mexican Plateau, yet this difference may also represent an artifact of ascertainment bias in haplotype sampling. We therefore consider the consistent patterns from within-population pairwise distances (π), and $\hat{\theta}$ to provide a more accurate depiction of the relative mitochondrial diversity across populations. Levels of expected versus observed mtDNA polymorphism resulted in negative *Fs* (consistent with population expansion) for the Chihuahuan and Mojave-Sonoran populations (Table 1), yet coalescent simulations suggest that only the northern population had significant evidence of expansion at ($p < 0.02$; (Fu, 1997)).

3.2. Nuclear estimates of phylogeny, population structure, and gene flow

The filtering scheme used for the total nuclear RAD loci from all samples resulted in 6337 SNPs from 2799 loci. Because we filtered to retain loci only if they were present in at least 30 of 44 individuals, the resulting data matrix had a relatively low proportion of missing data per individual (mean = 26.25%). In some cases, the number of loci or SNPs analyzed differed depending on the populations compared in various downstream analyses (see below). Bayesian clustering analysis on nuclear SNPs (using a single, randomly sampled SNP per locus) from the entire sample dataset in STRUCTURE provides consistent evidence for a best-fit model (ΔK) of three population clusters (Fig. 1D, Supplementary Fig. 1). Estimates of population assignments and admixture proportions from STRUCTURE runs were highly consistent across iterations. Under the $K = 3$ model, we find support for distinct Mojave-Sonoran and Chihuahuan genetic clusters, with Central Mexican Plateau and *salvini* samples assigned with high posterior probability to a single Central Mexican Plateau cluster. This model also provides implicit evidence of gene flow between populations based on shared posterior assignments of some individuals to multiple clusters – mixed-assignment individuals originate from localities adjacent to the neighboring population cluster, matching expectations of secondary contact between two distinct genetic lineages. At this scale, increasing values of K did not further segregate samples into unique genetic clusters, but instead identified small proportions of shared alleles among nearly all samples, potentially from shared ancestral variation. When we performed a more fine-scale analysis including only samples belonging to the Central Mexican Plateau and *salvini* mitochondrial clades, a $K = 2$ model clearly demarcated *salvini* as having distinct genetic structure from the Central Mexican Plateau population. This *salvini* cluster was also clearly delineated by DAPC analyses (Fig. 1E), which are independent of the model assumptions of STRUCTURE and tree-based analyses, despite a BIC best-fit model of three distinct genetics clusters (similar to STRUCTURE analysis).

Estimation of the population genetic diversity parameter, θ, using ThetaMater matched patterns of mitochondrial gene diversity in the four populations (Fig. 1F), with the Central Mexican Plateau and Chihuahuan populations having higher values of θ, *salvini* having the lowest θ, and an intermediate θ estimated for the Mohave-Sonoran population. After scaling the median posterior θ estimates for each population, we estimated effective population sizes (N_e) of 143,000 Mohave-Sonoran individuals, 152,000 Chihuahuan individuals, 153,000 Central Mexican Plateau individuals, and 135,000 *salvini* individuals.

Estimates from Bayesian tree analysis of concatenated SNP loci support three distinct clades corresponding to major populations and regions throughout the range of *C. scutulatus* that coincide with inferences from population clustering analyses (Supplementary Fig. 1). While the three RADseq *C. s. salvini* samples were monophyletic, their subclade is nested within the clade comprising populations from the Central Mexican Plateau and the recognized range of *salvini* (i.e., south of central Durango, Mexico; Supplementary Fig. 1), which contrasts the unique and genetically distant mitochondrial clade of *salvini* endemic to the subspecies' proposed range. In our nuclear phylogenetic inferences, the Central Mexican Plateau + *salvini* clade had strong posterior support, and was inferred as the sister lineage to the Mojave-Sonoran plus Chihuahuan clades. Further, the nuclear topology of Mojave-Sonoran and Chihuahuan clades was congruent with the mtDNA tree topology, splitting populations east and west of the Continental Divide (Supplementary Fig. 1). It is notable, however, that these results from concatenated nuclear SNP analysis should be interpreted with caution because population-level data do not adhere to assumptions of such a phylogenetic approach (Kubatko and Degnan, 2007), and incomplete lineages sorting, as well as gene flow, can substantially impact results (Mendes and Hahn, 2017). We therefore consider these results to represent a hypothesis that we further evaluated with other analyses (e.g., bpp species tree estimation). In contrast to our results from concatenated nuclear SNP phylogenetic analysis, Bayesian species delimitation in bpp did indicate 100% posterior probability for a four species model, comprising Mojave-Sonoran, Chihuahuan, Central Mexican Plateau, and *salvini* populations. Here, the best-supported topology

Table 1
Measures of nucleotide diversity and tests for evidence of population expansion within *C. scutulatus* mtDNA clades. The asterisk represents a significant value of Fu's Fs.

Population	Pairwise Distance	Haplotype Diversity	Pi	Theta	Fu's Fs	p-value
Mohave-Sonoran	0.0687	0.963	0.005	3.221	−7.967	< 0.00001*
Chihuahuan	0.085	0.987	0.0083	5.889	−2.278	0.089
Central Mexican Plateau	0.1085	0.905	0.013	9.818	2.225	0.26
Salvini	0.0212	0.6	0.0008	0.6	0.795	0.8

matches the mtDNA topology (Fig. 1G). Given the more appropriate modeling of population-level processes (e.g., gene flow and incomplete lineage sorting) by bpp, and the correspondence between the bpp and mtDNA-based estimates, we conclude that the most compelling evidence points to *salvini* as a distinct clade (as in the bpp and mtDNA analyses).

We find moderate to high levels of nuclear genome-wide allelic differentiation between the major nuclear clades, and patterns of differentiation are predicted by the mtDNA and nuclear phylogenies, with one exception. The highest F_{ST} is found between the Mojave-Sonoran population and *salvini* (mean $F_{ST} = 0.197$), followed by the comparison between the Mojave-Sonoran population and the Central Mexican Plateau population ($F_{ST} = 0.158$), and a slightly lower F_{ST} (0.113) was observed between the Mojave-Sonoran and Chihuahuan populations. Interestingly, we observe lower differentiation between the Chihuahuan and Central Mexican Plateau populations than between the Mojave-Sonoran and Chihuahuan populations ($F_{ST} = 0.098$), which we did not expect given the nuclear tree topology (Fig. 1G), but may reflect greater gene flow between populations adjacent to the phylogeographic break associated with the elevational increase of the Central Mexican Plateau and the possible filter barrier formed by the Río Nazas and Río Mezquital basins and the western reaches of the Sierra Madre Oriental (i.e., Arteaga anticline; (Hafner et al., 2008; Morafka, 1977a; Neiswenter and Riddle, 2010)). Alternatively, this inference may also be an artifact of the geographic distance between sampled localities on respective sides of the Continental Divide, resulting in a stronger signal of isolation-by-distance. Finally, we find that allelic differentiation between *salvini* samples and the Central Mexican Plateau is very low ($F_{ST} = 0.026$).

Tests of Patterson's D statistics across both major phylogeographic breaks within the range of *C. scutulatus* (i.e., the Continental Divide and the uplift of the Central Mexican Plateau) revealed significant evidence of introgression, suggesting that these boundaries are permeable to gene flow between adjacent populations (Fig. 2). The Continental Divide analysis resulted in a significantly positive value of D (0.2365, $p = 0.00024$), with 149 of 6058 SNPs fitting the 'ABBA' pattern, and 92 'BABA' sites. Of the 5077 SNPs used in the Central Mexican Plateau analysis, there are 380 'ABBA' sites and 157 'BABA' sites, also resulting in a significantly positive D value (0.4153, $p < 0.00001$).

Results from analyses of the 2D allele frequency spectrum using *δaδi* supported best-fit population divergence models of: (i) secondary contact with asymmetric gene flow between the Mohave-Sonoran and Chihuahuan populations (Fig. 3A), (ii) secondary contact with symmetric gene flow between the Chihuahuan and Central Mexican Plateau populations (Fig. 3B), and (iii) symmetrical ancestral migration followed by more recent divergence in isolation between the Central Mexican Plateau and *salvini* populations (Fig. 3C). Migration parameter estimates between the Mohave-Sonoran and Chihuahuan populations, and between the Chihuahuan and Central Mexican Plateau populations, corroborate evidence of introgression between adjacent populations (Supplementary Tables 2 and 3). Further, these estimates indicate a stronger overall signal of gene flow between the Chihuahuan and Central Mexican Plateau populations across the Mexican Plateau barrier than at the Continental Divide (Fig. 3B, Supplementary Tables 2 and 3), and further suggest that gene flow from the Chihuahuan population into the Mohave-Sonoran population is greater than in the reverse direction (Fig. 3A). To obtain estimates of population divergence times for each of the secondary contact models, we first calculated a scaled average population size parameter estimate (*Nref*), equal to the average of the unscaled population size estimates (*nu*) for each population divided by the effective sequence length per analysis, which is equal to the number of segregating sites analyzed in *δaδi* after downsampling of alleles multiplied by the total length of RAD loci analyzed, then divided by the total available segregating sites prior to downsampling. This value was then multiplied by 4 times the mutation rate (2.4×10^{-9}; see Methods for justification), and multiplied by the generation time of 3 years; these values were then multiplied by the sum of the *T1* and *T2* estimates. We repeated this process for the Central Mexican Plateau and *salvini* ancestral migration model, using only the *T1* estimate in the final calculation. This resulted in estimated nuclear divergence times of 1.459 MYA between Mohave-Sonoran and Chihuahuan populations, 4.138 MYA between Chihuahuan and Central Mexican Plateau populations, and 1.817 MYA between Central Mexican Plateau and *salvini* populations.

3.3. LGM modeling and historical biogeography

Ecological niche models derived from present-day environmental variables predicted suitable habitat that is largely overlapping with the known distribution of *C. scutulatus* (Fig. 4A); AUC values for models under the four subsampling schemes ranged from 0.925 to 0.945, and we obtained an average AUC value of 0.934 for our combined model. The minimum training presence-absence threshold corresponded to a logistic probability of 0.05. We found that our subsampling scheme did not appreciably affect the results, and predictions were consistent across all 12 different reconstructions (Supplementary Fig. 2). The average LGM model (Fig. 4B) suggests that *C. scutulatus* has persisted within the Mojave and Sonoran Deserts and the southern Chihuahuan Desert into central Mexico during glacial periods. Within the Mojave-Sonoran region, populations would have been restricted such that their range in all directions was less expansive relative to the current northern distribution of the species. Likewise, the inferred LGM distribution within Mexico was also restricted in size and confined to the southern extent of the current predicted distribution, leaving a considerable gap in the LGM distribution between predicted Mojave-Sonoran and Chihuahuan + Central Mexican Plateau refugia (Fig. 4B). This inference is consistent with discontinuous suitable habitat during this and perhaps more ancient glacial periods. Fragmented LGM habitat in the northern extent of the ancestral *C. scutulatus* distribution could support one of two alternative hypotheses: (1) existing populations inhabiting this region were restricted in space and isolated from one another; partial or complete barriers to gene flow contributed to substantial divergence, or (2) the species occupied one of these refugia during the LGM and underwent subsequent range expansion into the region of the other inferred refugium when suitable habitat became connected. Estimates of population structure and divergence dating (see below) are consistent with the former (i.e., the split between these lineages predates the LGM), and evidence of geographic isolation between populations east and west of the Continental Divide during glacial periods provides historical and geographic context for their divergence. Additionally, we find evidence of discontinuity in predicted suitable habitat in the extreme southern extent of the Mexican LGM distribution (Fig. 4B). The currently recognized distribution of *C. s. salvini* is restricted to this range, and potential isolation from adjacent southern populations may have facilitated spatial structuring of genetic diversity in this region.

Linear model tests of relationships between geography and nuclear genetic diversity revealed a pattern of spatial sorting consistent with greater diversity in the Chihuahuan and Central Mexican Plateau populations, and low levels of diversity in localities of the Mojave-Sonoran population consistent with range expansion (Fig. 4C–F). We found significant negative correlations between heterozygosity and both latitude ($N = 44$, $R^2 = -0.368$, $p = 0.0139$; Fig. 4C) and longitude ($N = 44$, $R^2 = -0.38$, $p = 0.0109$; Fig. 4D). We also found significant negative relationships between the number of private alleles and latitude ($N = 28$, $R^2 = -0.34$, $p = 0.0237$; Fig. 4E), as well as longitude ($N = 28$, $R^2 = -0.386$, $p = 0.0097$; Fig. 4F). These results imply an overall greater degree of genetic variation in the Chihuahuan and Central Mexican Plateau regions, and lower estimates with greater latitude and longitude are consistent with range expansion out of an ancestral range in central-southern Mexico. These estimates are also broadly consistent with patterns of mtDNA genetic diversity and

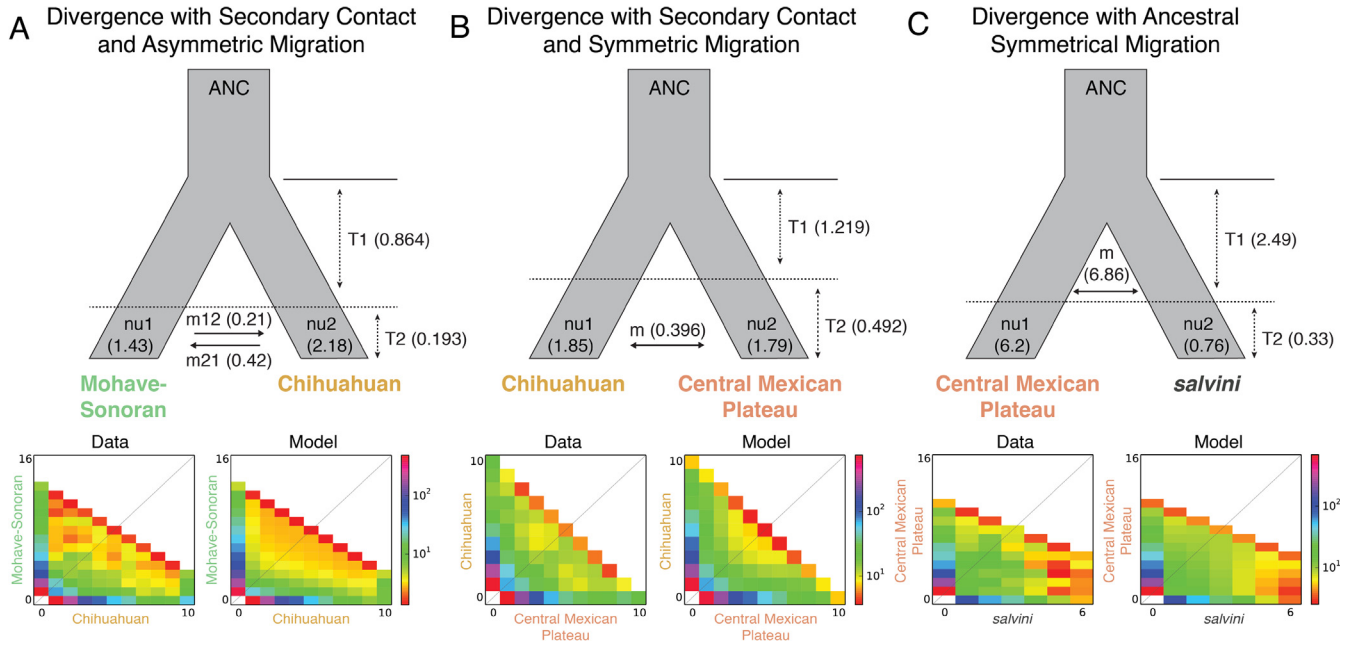

Fig. 3. Inferred demographic histories of *C. scutulatus* populations from analyses of the 2D allele frequency spectrum in *δaδi*. For each analysis, the best-fit model of population divergence is depicted in grey, with unscaled population size (nu), migration rate (m), and divergence time (T) parameters labeled. Below the best-fit models, the 2D allele frequency spectrum observed from the data is shown in comparison to the simulated model. The scale bar to the right of each data-model comparison shows the spectrum of allele frequencies.

nuclear θ estimates for each population. With regard to the three samples from the currently recognized range of *C. s. salvini* (and corresponding with the mtDNA '*salvini*' clade), low levels of heterozygosity, but moderate numbers of private alleles, may be result of a population bottleneck and/or founder's effects during southern range expansion.

3.4. Divergence time estimates and evidence for lineage co-divergence

Parameter estimates across replicate BEAST runs were similar, with low variation in median divergence time estimates and sufficient ESS values (6013–42,834). Median divergence time estimates for focal clades range from 173 KYA for the southern *salvini* mitochondrial clade to 14.13 MYA for the split between *Crotalus* and *Agkistrodon* (Fig. 5A). Median estimates and 95% highest posterior density distributions (HPD) for other divergence times among rattlesnake lineages correspond with estimates from Reyes-Velasco et al. (2013), with the exception of the common ancestor of *C. atrox* and the *C. scutulatus* group (median = 8.55 MYA), which is similar instead to estimates from Anderson and Greenbaum (2012). The median estimates and HPD for the splits between Mojave-Sonoran and Chihuahuan *C. scutulatus* (1.45 MYA; 0.78–2.5 MYA), and eastern *C. atrox* + western *C. atrox* (1.21 MYA; 0.76–1.75 MYA) overlap substantially, suggesting co-incident divergence events in these groups over the Continental Divide (Fig. 5A). The median divergence time estimate for the split between the Central Mexican Plateau and *salvini* mitochondrial clades is more ancient, and while the HPD overlaps with the distributions from Mojave-Sonoran/Chihuahuan *C. scutulatus* and eastern/western *C. atrox*, this divergence event more closely corresponds with the relative age of the split between *C. ruber* and *C. atrox* (2.09 MYA; 1.14–3.49 MYA). The 95% HPDs for each major split within *C. scutulatus* include the values of the nuclear estimates detailed above from analysis in *δaδi*. Explicit inference of co-occurring divergence events within our dataset focused on divergence between populations adjacent to the Continental Divide (Mojave-Sonoran and Chihuahuan *C. scutulatus* and eastern and western *C. atrox*). The results from msBayes analysis indicated strong posterior support for synchronous divergence within these taxon pairs – the estimated dispersion index (Ω) heavily sampled 0 (mean = 0.0046, 95% HPD = 0.0–0.05; Fig. 5B). The simple rejection and categorical rejection methods for inferring the number of possible divergence times (Ψ) identified the greatest posterior probability support for $\Psi = 1$ (simple rejection $pp = 0.821$, rejection with multinomial logistic regression $pp = 0.815$).

4. Discussion

4.1. Evolutionary history and phylogeography of Crotalus scutulatus

Our analysis of population genetic variation across the range of *C. scutulatus* provides the first molecular evidence of substantial population genetic structure within this group of highly venomous snakes (Fig. 1). Our divergence time estimates suggest that *C. scutulatus* lineages diverged from a common ancestor during the Pliocene (roughly 3.4 MYA), and have subsequently diversified over an expansive geographic range. Our phylogenetic and population genetic inferences from mitochondrial and nuclear data generally support four distinct lineages of *C. scutulatus* associated with relatively ancient divergences (Figs. 1, 3 and 5). Major lineages correspond with populations residing (from north to south) in the Mojave and Sonoran deserts in the United States, a large region of Chihuahuan Desert in Texas and northern Mexico, higher elevation regions of the Central Mexican Plateau (Morafka, 1977a), and the extreme southern extent of the species range in the Mexican states of México, Puebla, and Veracruz (the recognized range of *C. s. salvini*; Fig. 1).

We find consistent evidence from both mtDNA and nuclear SNPs for two major phylogeographic breaks within *C. scutulatus*: one break at the Continental Divide and a second associated with higher elevations of the Central Mexican Plateau. The 'Cochise Filter Barrier' region surrounding the Continental Divide is a biogeographic zone of restricted gene flow between numerous taxa (Morafka, 1977b), including multiple snake species (Myers et al., 2016), and patterns of *C. scutulatus* divergence further implicate this region as an important vicariant barrier among the North American desert taxa. Restriction of populations to suitable habitat on either side of this region during the

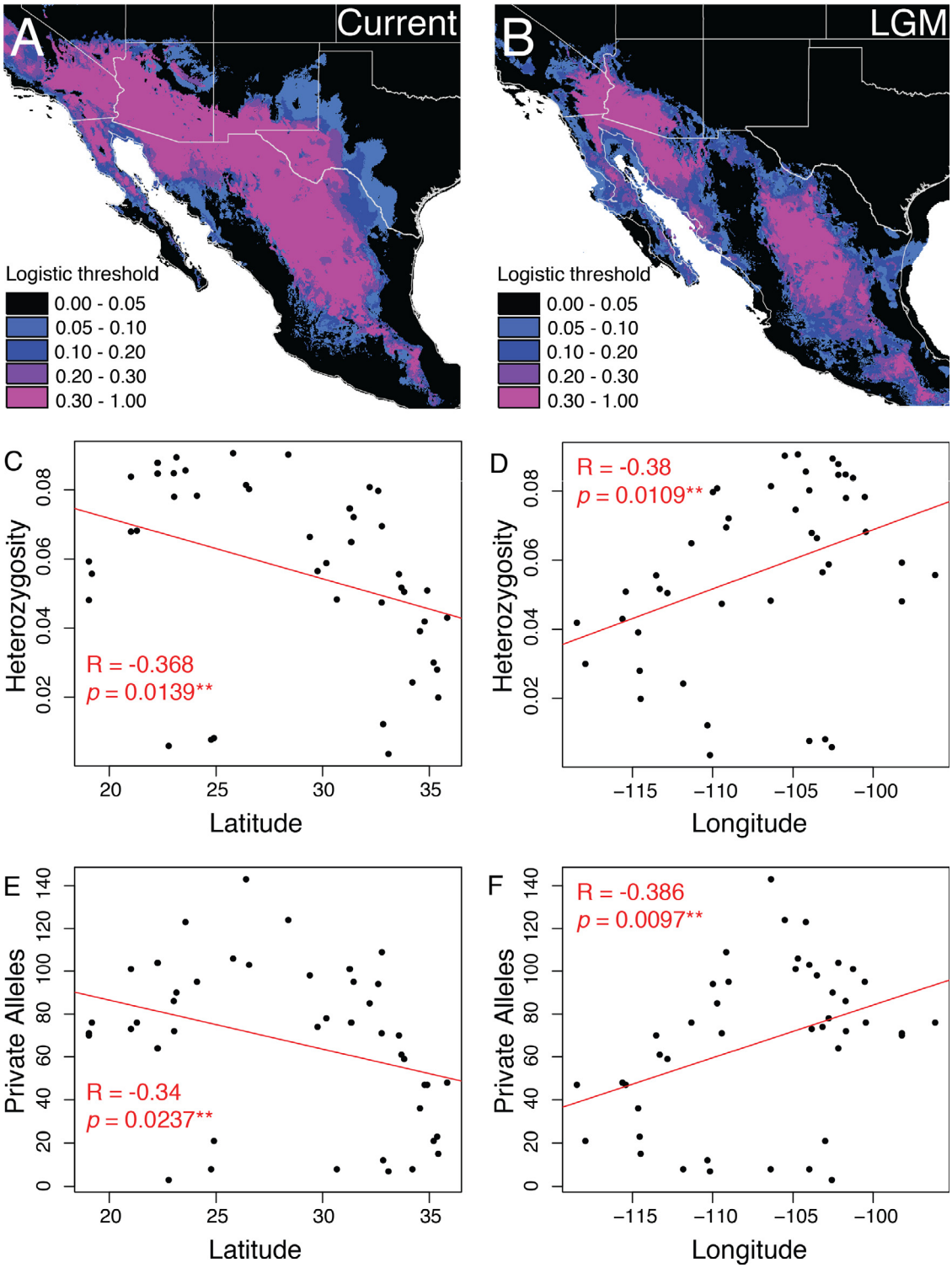

Fig. 4. Inferred current and historical suitable habitat based on ecological niche modeling of *C. scutulatus* localities and relationships between nuclear genetic diversity and geography. (A and B) Current and Last Glacial Maximum projections, respectively, of suitable habitat for *C. scutulatus* estimated using niche modeling. Warmer colors depict regions with a high logistic threshold probability of suitable habitat, while dark colors represent areas with low probability of *C. scutulatus* presence. C-D. Scatterplots of estimated heterozygosity versus latitude and longitude, respectively. (E and F) Scatterplots of numbers of private alleles versus latitude and longitude, respectively. Black dots represent individuals and trendlines, R-values, and *p* values are shown in red for each linear model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Pleistocene likely promoted divergence between lineages adjacent to the Continental Divide (Fig. 4) until more recent secondary contact after range expansion from both the Mojave-Sonoran and Chihuahuan regions, with asymmetric gene flow between populations across the Continental Divide (Fig. 3; Supplementary Table 2).

The second major phylogeographic break in *C. scutulatus* identified by our results appears to be associated generally with the elevational increase of the Central Mexican Plateau, which represents a known biogeographic barrier found along the transition between lowland Chihuahuan Desert habitat and higher elevation semi-arid habitat

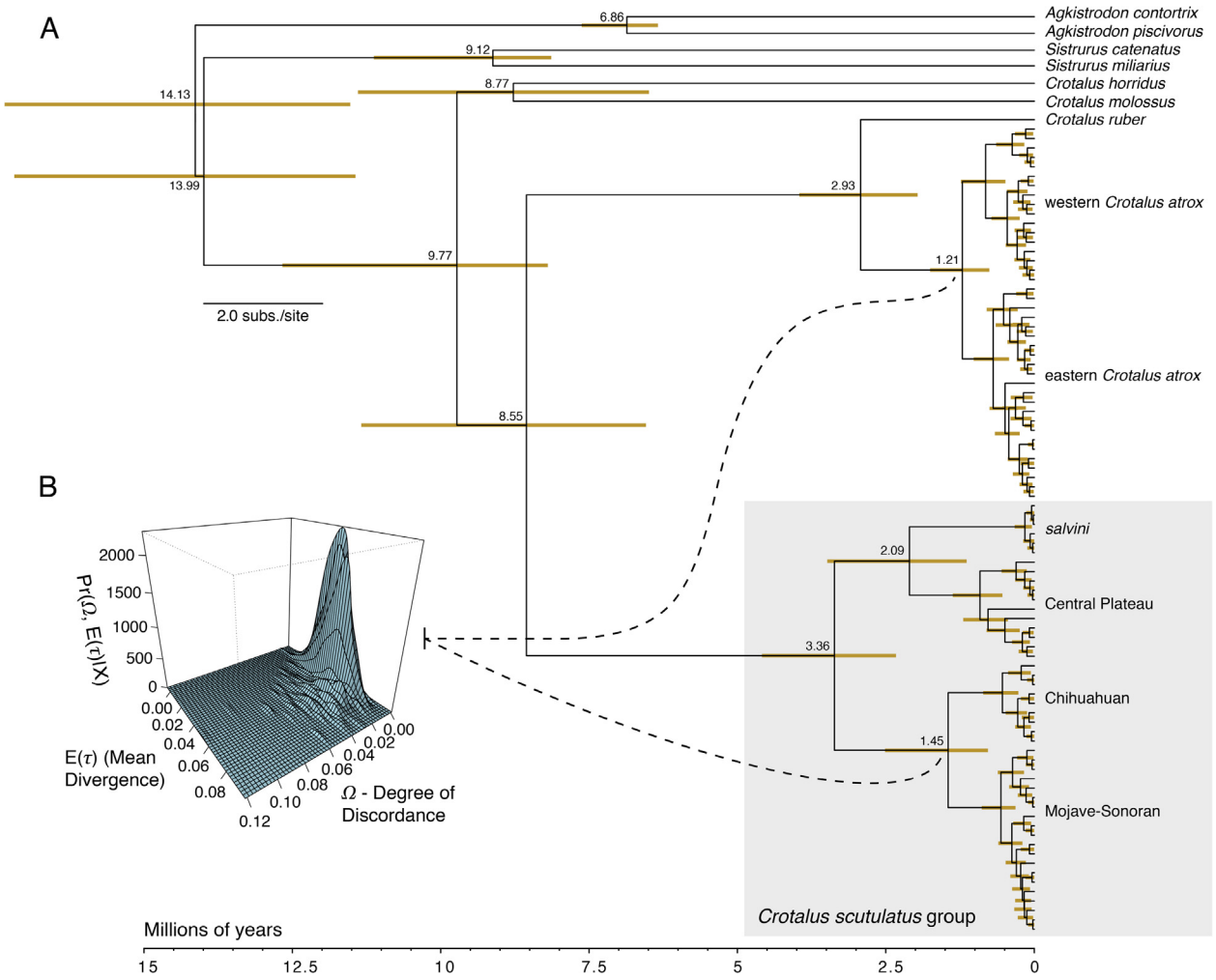

Fig. 5. Divergence time estimates and evidence for synchronous divergence of rattlesnake populations across the Continental Divide at the 'Cochise Filter Barrier'. (A) BEAST time tree based on analysis of mtDNA, showing *C. scutulatus*, *C. atrox*, and outgroup relationships and divergence time estimates, with the *C. scutulatus* group shaded in grey. Median estimates of node ages for major divergence events are labeled on the tree, and yellow bars at each node represent the 95% HPD. A scale of time in millions of years is shown at the bottom. (B) Results of tests of co-divergence in msBayes, depicting the combined distributions of the degree of discordance (Ω), mean relative divergence time (t), and probability distribution of mean divergence time and discordance given alignments of Mojave-Sonoran and Chihuahuan *C. scutulatus* and eastern and western *C. atrox*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Bryson et al., 2011a; Marshall and Liebherr, 2000; Morafka, 1977a). Our results suggest that Pleistocene climatic cycles may not have contributed to *C. scutulatus* diversification across this barrier, as we infer suitable habitat during glacial periods to have broad overlap with both the Chihuahuan and Central Mexican Plateau regions (Fig. 4B). Thus, we expect divergence across the uplift of the Central Mexican Plateau and the potential filter barrier formed by multiple river basins and the Arteaga anticline to have been driven by ecological constraints associated with the transition from arid Chihuahuan desert habitat to the semi-arid Central Mexican matorral, rather than changes in suitable habitat driven by climatic fluctuations. This prediction is also supported by the divergence time estimate between the ancestral lineages of extant *C. scutulatus* predating Pleistocene glaciation, and our inference that the Chihuahuan and Central Mexican Plateau populations have experienced secondary contact with gene flow (Fig. 2B, 3B; Supplementary Table 2).

We observed a diversification pattern in *C. scutulatus* across the Continental Divide that was similar to our previous findings for *C. atrox* (Castoe et al., 2007; Schield et al., 2017; Schield et al., 2015). Based on our new analysis of co-divergence, we found evidence for the synchronous divergence of eastern and western *C. atrox* lineages and Mojave-Sonoran and Chihuahuan *C. scutulatus* (Fig. 5A and B) at this boundary, suggesting that common historical processes during the Pleistocene drove simultaneous divergence of these taxa across this boundary. In a recent study, Myers et al. (2016) investigated diversification across the Continental Divide in twelve species and found broad evidence for asynchronous diversification sufficient to reject the hypothesis of a single vicariant event among all taxa studied. With respect to *C. scutulatus* and *C. atrox,* however, they observed overlapping posterior divergence time distributions, consistent with our findings. Our results indicating synchronous Pleistocene diversification for multiple rattlesnake lineages, suggest that taxa with similar ecologies, life histories, and climatic affinities (like *C. atrox* and *C. scutulatus*) may also show co-diverge across this barrier.

While there is substantial evidence from both mtDNA and nuclear loci for four distinct lineages within *C. scutulatus*, some inferences do not demonstrate substantial genetic structure between the Central Mexican Plateau and *salvini* populations, or disagree with regard to the relative placement of *C. s. salvini* on the phylogeny. Both our mitochondrial phylogenetic and nuclear-based species delimitation inferences strongly support four distinct clades of *C. scutulatus* (Fig. 1A and G), and DAPC clustering and more fine-scale analysis in

STRUCTURE clearly show distinct structure in *salvini* (Fig. 1D–E). However, the concatenated RADseq phylogeny and population clustering analyses support only three distinct clades in which *C. s. salvini* is nested within the Central Mexican Plateau lineage (Supplementary Fig. 1). Furthermore, genome-wide F_{ST} is very low between *salvini* and Central Mexican Plateau populations, suggesting pervasive gene flow, very recent divergence, or both. Inferences of the demographic history of the Central Mexican Plateau and *salvini* populations suggest that these populations have diverged in isolation relatively recently, following a prolonged period of pervasive gene flow. We expect that this particular demographic scenario could have resulted in genetic patterns that make nuclear population structure within *salvini* difficult to discern by broad scale analyses in STRUCTURE, and would also tend to mislead phylogenetic inferences based on concatenated SNPs. We therefore expect the four-lineage model supported by mtDNA and species-tree analysis of nuclear loci to be the most likely hypothesis.

4.2. Species hypotheses and evidence for gene flow between divergent lineages

Two subspecies of *C. scutulatus* are currently recognized: the Mojave Rattlesnake (*C. s. scutulatus*) and the Huamantlan Rattlesnake (*C. s. salvini*; (Campbell and Lamar, 2004). Under the current classification, the Mojave Rattlesnake occupies the vast majority of the northern portion of the species' distribution, while the Huamantlan Rattlesnake is restricted to a smaller, high-elevation (i.e., > 1800 m; (Campbell and Lamar, 2004) region at the southern end of the species' distribution. Importantly, our results indicate that populations currently recognized as *C. s. scutulatus* are not monophyletic. Further, our results highlight cryptic, taxonomically unrecognized lineage diversity within *C. scutulatus*, as phylogenetic and population genetic analyses of both mtDNA and nuclear loci consistently support four distinct lineages. Thus, the current taxonomy does not capture the diversity within *C. scutulatus*, nor the important biogeographic barriers and historical processes that have driven diversification within the group.

The divergence time estimate between the clades containing *salvini* and Central Mexican Plateau populations and the common ancestor of Mojave-Sonoran and Chihuahuan *C. scutulatus* (3.4 MYA; Fig. 5A) is more ancient than several recognized rattlesnake species pairs (e.g., *C. ruber* and *C. atrox*, 2.99 MYA (Reyes-Velasco et al., 2013); *C. viridis* and *C. oreganus*, 3.1 MYA (Anderson and Greenbaum, 2012). Additionally, the divergence between the Mojave-Sonoran and Chihuahuan *C. scutulatus* (1.45 MYA; Fig. 5A) is also more ancient than some recognized rattlesnake species (e.g., *C. durissus* and *C. simus*, ~1 MYA (Blair and Sanchez-Ramirez, 2016). Based on this precedent, the four *C. scutulatus* lineages represent excellent 'species hypotheses' to be tested further.

It is becoming clear from molecular studies of rattlesnakes that considerable divergence (i.e., millions of years) does not necessarily prevent gene flow and introgression in situations of secondary contact (Castoe et al., 2007; Meik et al., 2015; Murphy and Ben Crabtree, 1988; Schield et al., 2015). For example, divergent populations of the Western Diamondback Rattlesnake (*C. atrox*) east and west of the Continental Divide introgress over a broad geographic region in the northern Chihuahuan Desert (Schield et al., 2017). Divergent lineages of *C. scutulatus* also exhibit this pattern, with evidence of individuals with admixed assignment from Bayesian population clustering analysis in STRUCTURE (Fig. 1D) and a significantly positive value of Patterson's D (Fig. 2A). These findings also highlight the apparent permeability of major historic barriers to gene flow. For example, although the Continental Divide has been repeatedly demonstrated as an important historical driver of diversification in many arid-adapted taxa, migrants from divergent rattlesnake lineages are able to introgress across this barrier. Evidence for gene flow across the elevational increase at the Central Mexican Plateau between Chihuahuan and Central Mexican Plateau populations is also notable, given that these lineages diverged from an even more ancient ancestor than did the lineages that meet at the Continental Divide. Patterns of gene flow in *C. scutulatus* therefore add to a growing body of evidence that distinct lineages (even species) may experience substantial levels of admixture or hybridization in secondary contact (Murphy and Ben Crabtree, 1988; Schield et al., 2017; Zancolli et al., 2016), underscoring rattlesnakes as an important model system for understanding incomplete reproductive isolation and the evolution (or lack thereof) of postzygotic isolation mechanisms in secondary contact.

5. Conclusion

In this study, we investigated range-wide patterns of diversity, population structure, and gene flow in the Mojave Rattlesnake. Through analyses of both mtDNA and nuclear genomic data we find consistent evidence for previously undocumented lineage diversity within this group. These lineages represent intriguing species hypotheses to be further tested using complementary data to the genomic data presented here (e.g., morphological and venom variation data). Our findings also underscore the importance of using multiple types of genetic markers to explore range-wide diversity in such wide-ranging lineages, which in our case led to the identification of multiple apparently distinct lineages and of important physiographic features promoting divergence. Beyond biogeographic and taxonomic implications, previously undocumented diversity within *Crotalus scutulatus* has immediate medical relevance due to the toxicity and diversity of its venom composition. Evidence of population genetic structure (and gene flow) presented here may thus provide a useful framework for detailed comparative and evolutionary analyses of venom variation in general, and specifically for understanding the apparently complex gain and loss of neurotoxic function in this broadly distributed species complex (Glenn and Straight, 1978, 1989; Glenn et al., 1983).

Acknowledgments

We thank Ann Trápaga, Daniel Antonio Rangel, Edmundo Pérez-Ramos, J. Jesús Sigala Rodríguez, Raul Hernández Árciga, Carlos A. Hernández Jiménez, Luis Felipe Vázquez Vega, Luis Fernando Dávila Galavíz, Erick Alejandro García Trejo, Gustavo E. Quintero Díaz, Valente Bernal, Citlalli Edith Esparza Estrada, Jesús Lenin Lara Galván, Rogelio Rosales García, Juan José Ayala Rodríguez, Alejandra Guadalupe López Campos, Cruz Estefanía López Campos, Antonio Martínez Méndez, Gustavo Contreras Cobos, Víctor Hugo Reynoso Rosales, and Agustín Álvarez Trillo (and his students from the herpetário del Zoológico de Zacango, Estado de México) for assistance in the field collecting and/or providing access to live animals and specimens. We thank Jonathan Campbell, Carl Franklin, and Corey Roelke (UTA), Robert Murphy (Royal Ontario Museum), Hans Werner Herrmann, Leonardo Badillo, Jens Vindum (California Academy of Sciences), Adrián Neito Montes de Oca (UNAM), Juan Carlos López Vidal (Instituto Politécnico Nacional) for providing tissue samples, and Elda Sanchez, Mark Hockmuller, and Juan Salinas for assistance with tissue samples from the Texas A&M Kingsville Natural Toxins Research Center (NTRC). Scientific collecting permits were issued by the California Department of Fish and Wildlife (SC-12985), the New Mexico Department of Game and Fish (3563, 3576), the State of Arizona Game and Fish Department (SP628489, SCP660656, SP673390, SP673626, SP715023), Nevada Department of Wildlife (S 33773), Texas Parks and Wildlife (SPR-0390-029) and Secretaria de Medio Ambiente y Recursos Naturales of the Estados Unidos Mexicanos (SGPA/DGVS/03562/15). Support for this work was provided by faculty startup funds from the University of Texas at Arlington to TAC, NSF Grant DEB-1655571 to TAC, SPM, and JMM, UC Mexus CN-11-548 to CLS, NSF grants DUE-1161228 and DEB 1638879 to CLS, NSF DDIG Grant DEB-1501886 awarded to DRS and TAC, NSF DDIG Grant DEB-1501747 to DCC and TAC, a Phi Sigma Beta Phi Chapter research grant to DRS, and a Theodore Roosevelt Memorial Fund research grant, Prairie Biotic

Research Inc. Grant, Sigma Xi Grants-in-aid-of-research, SnakeDays Research Grant, and Southwestern Association of Naturalists Howard McCarley research grant to JLS. Collecting permits in Mexico were issued by SEMARNAT to OFV, permit number FAUT-0015.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.06.013.

References

Adams, R.H., Schield, D.R., Card, D.C., Corbin, A., Castoe, T.A., 2017. ThetaMater: Bayesian estimation of population size parameter h from genomic data. Bioinformatics 34, 1072–1073.

Anderson, C.G., Greenbaum, E., 2012. Phylogeography of northern populations of the Black-Tailed Rattlesnake (*Crotalus Molossus* Baird and Girard, 1853), with the revalidation of *C. ornatus* Hallowell, 1854. Herpetological Monogr. 26, 19–57.

Arevalo, E., Davis, S.K., Sites, J.W., 1994. Mitochondrial-DNA sequence divergence and phylogenetic relationships among 8 chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Systematic Biol. 43, 387–418.

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

Bolnick, D.I., Fitzpatrick, B.M., 2007. Sympatric speciation: models and empirical evidence. Annu. Rev. Ecol., Evol., System. 38, 459–487.

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.

Bryson, R.W., García-Vázquez, U.O., Riddle, B.R., 2011a. Phylogeography of Middle American gophersnakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. J. Biogeogr. 38, 1570–1584.

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

Campbell, J.A., Lamar, W.W., 2004. The Venomous Reptiles of the Western Hemisphere. Cornell University Press, Ithaca, NY.

Castoe, T.A., de Koning, A.P.J., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P., Degner, J.F., Daza, J.M., Gu, W.J., Reyes-Velasco, J., Shaney, K.J., Castoe, J.M., Fox, S.E., Poole, A.W., Polanco, D., Dobry, J., Vandewege, M.W., Li, Q., Schott, R.K., Kapusta, A., Minx, P., Feschotte, C., Uetz, P., Ray, D.A., Hoffmann, F.G., Bogden, R., Smith, E.N., Chang, B.S.W., Vonk, F.J., Casewell, N.R., Henkel, C.V., Richardson, M.K., Mackessy, S.P., Bronikowsi, A.M., Yandell, M., Warren, W.C., Secor, S.M., Pollock, D.D., 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc. Natl. Acad. Sci., USA 110, 20645–20650.

Castoe, T.A., Parkinson, C.L., 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol. Phylogenet. Evol. 39, 91–110.

Castoe, T.A., Spencer, C.L., Parkinson, C.L., 2007. Phylogeographic structure and historical demography of the western diamondback rattlesnake (*Crotalus atrox*): a perspective on North American desert biogeography. Mol. Phylogenet. Evol. 42, 193–212.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140.

Cate, R.L., Bieber, A.L., 1978. Purification and characterization of Mojave (*Crotalus scutulatus scutulatus*) Toxin and its subunits. Arch. Biochem. Biophys. 189, 397–408.

Cox, C.L., Davis Rabosky, A.R., Reyes-Velasco, J., Ponce-Campos, P., Smith, E.N., Flores-Villela, O., Campbell, J.A., 2012. Molecular systematics of the genus *Sonora* (Squamata: Colubridae) in central and western Mexico. System. Biodiversity 10, 93–108.

Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates, Sunderland, MA.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2252.

Durban, J., Perez, A., Sanz, L., Gomez, A., Bonilla, F., Rodriguez, S., Chacon, D., Sasa, M., Angulo, Y., Gutierrez, J.M., Calvete, J.J., 2013. Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake, Crotalus simus simus. BMC Genomics 14, 234.

Earl, D.A., Vonholdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genetics Resour. 4, 359–361.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods Ecol. Evol. 1, 330–342.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.

Feder, J.L., Egan, S.P., Nosil, P., 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28, 342–350.

Fu, Y.-X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Glenn, J.L., Straight, R., 1978. Mojave rattlesnake (*Crotalus scutulatus scutulatus*) venom: variation in toxicity with geographical origin. Toxicon 16, 81–84.

Glenn, J.L., Straight, R.C., 1989. Intergradation of two different venom populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. Toxicon 27, 411–418.

Glenn, J.L., Straight, R.C., Wolfe, M.C., Hardy, D.L., 1983. Geographical variation in *Crotalus scutulatus scutulatus* (Mojave rattlesnake) venom properties. Toxicon 21, 119–130.

Green, R.E., Braun, E.L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., Vandewege, M.W., John, J.A.S., Capella-Gutiérrez, S., Castoe, T.A., 2014. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science 346, 1254449.

Hafner, D.J., Hafner, M.S., Hasty, G.L., Spradling, T.A., Demastes, J.W., 2008. Evolutionary relationships of pocket gophers (*Cratogeomys castanops* species group) of the Mexican Altiplano. J. Mammal. 89, 190–208.

Hannon, G.J., 2014. FASTX-Toolkit.

Hardy, D.L., 1983. Envenomation by the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Southern Arizona, USA. Toxicon 21, 111–118.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Huang, W., Takebayashi, N., Qi, Y., Hickerson, M.J., 2011. MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. BMC Bioinform. 12, 1.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Jaeger, J.R., Riddle, B.R., Bradford, D.F., 2005. Cryptic Neogene vicariance and Quaternary dispersal of the red-spotted toad (*Bufo punctatus*): insights on the evolution of North American warm desert biotas. Mol. Ecol. 14, 3033–3048.

Jezkova, T., Olah-Hemmings, V., Riddle, B.R., 2011. Niche shifting in response to warming climate after the last glacial maximum: inference from genetic data and niche assessments in the chisel-toothed kangaroo rat (*Dipodomys microps*). Global Change Biol. 17, 3486–3502.

Jezkova, T., Riddle, B.R., Card, D.C., Schield, D.R., Eckstut, M.E., Castoe, T.A., 2015. Genetic consequences of postglacial range expansion in two codistributed rodents (genus *Dipodomys*) depend on ecology and genetic locus. Mol. Ecol. 24, 83–97.

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94.

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

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. System. Biol. 56, 17–24.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.

Lee, C., Abdool, A., Huang, C.-H., 2009. PCA-based population structure inference with generic clustering algorithms. BMC Bioinform. 10, S73.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Mackessy, S.P., 2008. Venom composition in rattlesnakes: trends and biological significance. In: Hayes, W.K., Beaman, K.R., Cardwell, M.D., Bush, S.P. (Eds.), The Biology of Rattlesnakes. Loma Linda University Press, Loma Linda, CA.

Marshall, C.J., Liebherr, J.K., 2000. Cladistic biogeography of the Mexican transition zone. J. Biogeogr. 27, 203–216.

Massey, D.J., Calvete, J.J., Sanchez, E.E., Sanz, L., Richards, K., Curtis, R., Boesen, K., 2012. Venom variability and envenoming severity outcomes of the *Crotalus scutulatus scutulatus* (Mojave rattlesnake) from Southern Arizona. J. Proteom. 75, 2576–2587.

Mayr, E., 1942. Systematics and the Origin of Species, From the Viewpoint of a Zoologist. Harvard University Press.

Meik, J.M., Streicher, J.W., Lawing, A.M., Flores-Villela, O., Fujita, M.K., 2015. Limitations of climatic data for inferring species boundaries: insights from speckled rattlesnakes. PloS one 10, e0131435.

Mendes, F.K., Hahn, M.W., 2017. Why concatenation fails near the anomaly zone. System. Biol. 67, 158–169.

Miller, S.A., Dykes, D.D., Polesky, H., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucl. Acids Res. 16, 1215.

Morafka, D.J., 1977a. A biogeographical analysis of the Chihuahuan desert through its herpetofauna. W. Junk BV Publishers, The Hague, Dr.

Morafka, D.J., 1977b. Is there a Chihuahuan Desert? A quantitative evaluation through a herpetofaunal perspective. In: Transactions of the Symposium on the Biological Resources of the Chihuahuan Desert Region, United States and Mexico. National Park Service, Washington, DC, pp. 437–454.

Murphy, R.W., Ben Crabtree, C., 1988. Genetic identification of a natural hybrid rattlesnake: *Crotalus scutulatus scutulatus* × *C. viridis viridis*. Herpetologica 119–123.

Murphy, R.W., Fu, J., Lathrop, A., Feltham, J.V., Kovac, V., 2002. Phylogeny of the rattlesnakes (*Crotalus* and *Sistrurus*) inferred from sequences of five mitochondrial DNA genes. In: Schuett, G.W., Hoggren, M., Douglas, M.E., Greene, H.W. (Eds.), Biology of the Vipers. Eagle Mountain Publishing.

Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M., Anderson, R.P., 2014. ENMeval: An R package for conducting spatially independent

evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol. Evol. 5, 1198–1205.
Myers, E.A., Hickerson, M.J., Burbrink, F.T., 2016. Asynchronous diversification of snakes in the North American warm deserts. J. Biogeogr. 44, 461–474.
Neiswenter, S.A., Riddle, B.R., 2010. Diversification of the *Perognathus flavus* species group in emerging arid grasslands of western North America. J. Mammal. 91, 348–362.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one 7, e37135.
Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259.
Portik, D.M., Leaché, A.D., Rivera, D., Barej, M.F., Burger, M., Hirschfeld, M., Rödel, M.O., Blackburn, D.C., Fujita, M.K., 2017. Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Mol. Ecol. 26, 5245–5263.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
R Core Team, 2017. R: A language and environment for statistical computing.
Raes, N., ter Steege, H., 2007. A null-model for significance testing of presence-only species distribution models. Ecography 30, 727–736.
Reyes-Velasco, J., Meik, J.M., Smith, E.N., Castoe, T.A., 2013. Phylogenetic relationships of the enigmatic longtailed rattlesnakes (*Crotalus ericsmithi, C. lannomi,* and *C. stejnegeri*). Mol. Phylogenet. Evol. 69, 524–534.
Riddle, B.R., Hafner, D.J., Alexander, L.F., 2000. Phylogeography and systematics of the *Peromyscus eremicus* species group and the historical biogeography of North American warm regional deserts. Mol. Phylogenet. Evol. 17, 145–160.
Roch, S., Steel, M., 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theoret. Populat. Biol. 100, 56–62.
Sanchez, E.E., Galan, J.A., Powell, R.L., Reyes, S.R., Soto, J.G., Russell, W.K., Russell, D.H., Perez, J.C., 2005. Disintegrin, hemorrhagic, and proteolytic activities of Mohave rattlesnake, *Crotalus scutulatus scutulatus* venoms lacking Mojave toxin. Comp. Biochem. Physiol. C: Toxinol. Pharmacol. 141, 124–132.
Schield, D.R., Adams, R.H., Card, D.C., Perry, B.W., Pasquesi, G.M., Jezkova, T., Portik, D.M., Andrew, A.L., Spencer, C.L., Sanchez, E.E., 2017. Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. Ecol. Evol. 7, 3951–3966.
Schield, D.R., Card, D.C., Adams, R.H., Jezkova, T., Reyes-Velasco, J., Proctor, F.N., Spencer, C.L., Herrmann, H.W., Mackessy, S.P., Castoe, T.A., 2015. Incipient speciation with biased gene flow between two lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*). Mol. Phylogenet. Evol. 83, 213–223.
Slatkin, M., Excoffier, L., 2012. Serial founder effects during range expansion: a spatial analog of genetic drift. Genetics 191, 171–181.
Smith, C.F., Mackessy, S.P., 2016. The effects of hybridization on divergent venom phenotypes: characterization of venom from *Crotalus scutulatus scutulatus* x *Crotalus oreganus helleri* hybrids. Toxicon 120, 110–123.
Streicher, J.W., Devitt, T.J., Goldberg, C.S., Malone, J.H., Blackmon, H., Fujita, M.K., 2014. Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. Mol. Ecol. 2014, 3273–3291.
Streicher, J.W., McEntee, J.P., Drzich, L.C., Card, D.C., Schield, D.R., Smart, U., Parkinson, C.L., Jezkova, T., Smith, E.N., Castoe, T.A., 2016. Genetic surfing, not allopatric divergence, explains spatial sorting of mitochondrial haplotypes in venomous coralsnakes. Evolution 70, 1435–1449.
Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci., USA 107, 9264–9269.
Zancolli, G., Baker, T.G., Barlow, A., Bradley, R.K., Calvete, J.J., Carter, K.C., de Jager, K., Owens, J.B., Price, J.F., Sanz, L., 2016. Is hybridization a source of adaptive venom variation in rattlesnakes? A test, using a *Crotalus scutulatus* × *viridis* hybrid zone in southwestern New Mexico. Toxins 8, 188.
Zhao, N.L.a.H., 2006. A non-parametric approach to population structure inference using multilocus genotypes. Human Genom. 2, 353.