Molecular Phylogenetics and Evolution 83 (2015) 213–223

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Incipient speciation with biased gene flow between two lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*)

Drew R. Schield [a], Daren C. Card [a], Richard H. Adams [a], Tereza Jezkova [b], Jacobo Reyes-Velasco [a], F. Nicole Proctor [a], Carol L. Spencer [c], Hans-Werner Herrmann [d], Stephen P. Mackessy [e], Todd A. Castoe [a,*]

[a] Department of Biology & Amphibian and Reptile Diversity Research Center, 501 S. Nedderman Drive, University of Texas at Arlington, Arlington, TX 76019, USA
[b] School of Life Sciences, University of Nevada, Las Vegas, 4505 Maryland Parkway, Las Vegas, NV 89154, USA
[c] Museum of Vertebrate Zoology, 3101 Valley Life Sciences Building, University of California, Berkeley, CA 94720, USA
[d] School of Natural Resources and the Environment, 1041 E Lowell Street, University of Arizona, Tuscon, AZ 85721, USA
[e] School of Biological Sciences, 501 20th Street, University of Northern Colorado, Greeley, CO 80639, USA

ARTICLE INFO

Article history:
Received 11 July 2014
Revised 3 December 2014
Accepted 9 December 2014
Available online 19 December 2014

Keywords:
Population genomics
Population structure
Introgression
Speciation continuum
Viperidae
RADseq

ABSTRACT

We used mitochondrial DNA sequence data from 151 individuals to estimate population genetic structure across the range of the Western Diamondback Rattlesnake (*Crotalus atrox*), a widely distributed North American pitviper. We also tested hypotheses of population structure using double-digest restriction site associated DNA (ddRADseq) data, incorporating thousands of nuclear genome-wide SNPs from 42 individuals. We found strong mitochondrial support for a deep divergence between eastern and western *C. atrox* populations, and subsequent intermixing of these populations in the Inter-Pecos region of the United States and Mexico. Our nuclear RADseq data also identify these two distinct lineages of *C. atrox*, and provide evidence for nuclear admixture of eastern and western alleles across a broad geographic region. We identified contrasting patterns of mitochondrial and nuclear genetic variation across this genetic fusion zone that indicate partially restricted patterns of gene flow, which may be due to either pre- or post-zygotic isolating mechanisms. The failure of these two lineages to maintain complete genetic isolation, and evidence for partially-restricted gene flow, imply that these lineages were in the early stages of speciation prior to secondary contact.

1. Introduction

Speciation proceeds with the origin of barriers to gene flow, which permits the maintenance of genetic and phenotypic divergence (Coyne and Orr, 2004; Nosil and Feder, 2012). Isolation mechanisms may vary in strength, leading to a continuum of speciation based on the degree to which such mechanisms promote reproductive isolation between lineages (Coyne and Orr, 2004; Nosil and Feder, 2012). Lineages in the early stages of this speciation continuum are particularly valuable as model systems for understanding the mechanisms that drive speciation. They are valuable for understanding primary mechanisms of isolation because they lack confounding secondary isolation mechanisms that evolve later (Orr, 1995; Good et al., 2008). Accordingly, studying these model systems may also provide novel insight into why speciation might not occur, leading to the secondary fusion of divergent lineages (Taylor et al., 2006; Wiens et al., 2006; Webb et al., 2011). Comparing patterns of mitochondrial and nuclear genetic variation can be useful for identifying such patterns of partial genetic isolation, due to their different modes of inheritance. Here we conduct a detailed analysis of nuclear and mitochondrial gene flow between two divergent lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*), and assess evidence for where these divergent lineages may exist on the speciation continuum.

The Western Diamondback Rattlesnake is a large venomous rattlesnake native to the United States and Mexico. It inhabits among the broadest distributions of all rattlesnake species, ranging across much of the southwestern United States and northern Mexico (Campbell and Lamar, 2004). This species is viewed as a habitat and diet generalist, living in a diversity of lowland habitats and feeding on a large variety of prey. There are no described subspecies within *C. atrox*, although there is evidence of morphological variation across their range (Klauber, 1956; Spencer, 2008). Given its large range, local abundance, large size, and toxic venom, *C.*

* Corresponding author at: Department of Biology, University of Texas at Arlington, 501 S. Nedderman Drive, 337 Life Science, Arlington, TX 76010-0498, USA. Fax: +1 817 272 2855.
E-mail address: todd.castoe@uta.edu (T.A. Castoe).

http://dx.doi.org/10.1016/j.ympev.2014.12.006
1055-7903/

atrox is among the most medically relevant species in North America, with the highest number of human fatalities due to envenomation of any snake (Campbell and Lamar, 2004).

In a previous study, Castoe et al. (2007) found evidence for two distinct *C. atrox* mitochondrial lineages that diverged approximately 1.36 MYA, corresponding roughly to populations east and west of the Continental Divide of the United States and Mexico. Although sampling was limited, the authors found some evidence that these two mitochondrial lineages overlap in distribution near the Continental Divide, suggesting possible gene flow between these lineages. Although mitochondrially-encoded loci have been used extensively to study snake phylogeography and population genetics (e.g., Ashton and de Queiroz, 2001; Douglas et al., 2002; Castoe et al., 2007; Burbrink and Castoe, 2009; Meik et al., 2012), they may obscure some patterns of introgression due to their matrilineal inheritance, lack of recombination, and rapid coalescence (Avise and Vrijenhoek, 1987; Palumbi and Baker, 1994; Avise, 2000). The combination of mitochondrial data with data from nuclear single nucleotide polymorphisms (SNPs), however, provides a more powerful means of testing population genetic hypotheses and of comparing matrilineal versus whole genome patterns of genetic variation.

Here we combined extensive sampling of individuals for a mitochondrial gene with sampling of thousands of nuclear SNPs from restriction site associated DNA sequencing (RADseq) to investigate patterns of historical and current divergence, and gene flow across the range of *C. atrox*. Our aim in this study was to use both mitochondrial and nuclear data to characterize patterns of divergence and secondary contact in *C. atrox*, and to detect signatures of progression towards speciation of lineages within *C. atrox* prior to recent introgression. To address this aim, we tested the following hypotheses: (1) Mitochondrial and nuclear data provide evidence of two divergent *C. atrox* lineages, (2) mitochondrial and nuclear data show evidence of recent gene flow between these lineages, (3) both eastern and western populations show similar patterns of genetic diversity and historical demography based on mitochondrial and nuclear data, and (4) gene flow is unrestricted between eastern and western lineages, and there is no evidence of reproductive isolation.

2. Materials and methods

2.1. Taxon sampling and DNA extraction

We obtained tissues from a total of 151 *Crotalus atrox* from throughout their range (Fig. 1B; Table S1), including the 48 samples in Castoe et al. (2007). Tissues included samples of blood, liver, and skin preserved by snap-freezing or via lysis buffer, RNALater, shed skins or rattles. Genomic DNA was isolated in one of four ways: using a Qiagen DNeasy extraction kit (samples from Castoe et al., 2007; Qiagen, Inc., Valencia, CA, USA), Zymo Research Genomic DNA Tissue MiniPrep kit (most solid tissues; Zymo Research Corporation, Irvine, CA, USA), Thermo Scientific GeneJet whole blood DNA extraction kit (blood; Thermo Fisher Scientific, Inc., Waltham, MA, USA), or phenol-choroform-isoamyl alcohol (some shed skins).

2.2. Mitochondrial locus amplification and sequencing

We used PCR to amplify a fragment of the mitochondrially-encoded NADH dehydrogenase subunit 4 (ND4), plus the downstream Serine, Histidine, and Leucine tRNAs, using the primers ND4 and Leu (Arevalo et al., 1994). PCR products were purified using AgenCourt AMPure XP beads (Beckman Coulter, Inc., Irving, TX, USA). Purified PCR products were quantified and sequenced in both directions with the use of amplification primers, using BigDye on an ABI 3730 capillary sequencer (Life Technologies, Grand Island, NY, USA) at the UTA Genomics Core Facility.

2.3. ddRADseq library generation and sequencing

A subset of the DNA samples used for ND4 PCR were also used to generate ddRAD libraries ($n = 42$; Fig. 3B, Supplementary Online Table 1), largely following the protocol of Peterson et al. (2012) except for notable exceptions outlined below. Samples were chosen based on locality in order to include representative specimens of the putative eastern, western, and hybrid zone populations. Genomic DNA was digested using a combination of rare (*SbfI*; 8 bp recognition site) and common (*Sau3AI*; 4 bp recognition site) cutting restriction enzymes. Double-stranded indexed DNA adapters were ligated to the ends of digested fragments that also contained unique molecular identifiers (UMIs; eight consecutive N's prior to the ligation site). Following adapter ligation, samples were pooled in sets of eight, size selected for a range of 590–640 bp using the Blue Pippin Prep (Sage Science, Beverly, MA, USA), and PCR amplified using primers to complete attachment of flow-cell binding sequences and addition of a second index specific to each sub-pool. Sub-pools were pooled again based on molarity calculations from analysis on a Bioanalyzer (Agilent, Santa Clara, CA, USA) using a DNA 7500 chip, and sequenced using 100 bp paired-end reads on an Illumina HiSeq 2500.

2.4. mtDNA sequence analysis

Raw mitochondrial gene sequence chromatograms were edited using Geneious v6.1.6 (Biomatters Ltd., Auckland, NZ), and we aligned edited sequences using MUSCLE (Edgar, 2004) with minimal manual adjustments to improve alignment of tRNA gene regions and to trim the 5′ and 3′ ends of all sequences to reduce columns with high levels of missing data. The final alignment contained 813 aligned bases and included no indels.

We estimated phylogenetic relationships among unique *C. atrox* haplotypes and outgroup species using Bayesian phylogenetic inference in MrBayes v3.2.1. (Huelsenbeck and Ronquist, 2001). For outgroups, we used single representatives *C. molossus*, *C. tigris*, and *C. ruber*, which were chosen based on previous estimates of relationships among rattlesnakes (Reyes-Velasco et al., 2013) We used the Bayesian Information Criterion (BIC) implemented in PartitionFinder v1.1.1 (Lanfear et al., 2012) to select best-fit models, which were HKY for 1st and 2nd codon positions, as well as for tRNAs, and TN + Γ + invariant sites for 3rd codon positions. This partitioned model was used for analyses in MrBayes, which consisted of four runs, each run for 10^7 generations with four chains (one cold and three heated), sampled every 500 generations. We accounted for among partition rate variation using the "prset ratepr = variable" command. Potential scale reduction factor value estimates (PSRF) indicated that individual runs had converged by 10^5 generations, and thus we discarded the first 10^5 samples as burn-in. We also confirmed that independent runs had converged based on overlap in likelihood and parameter estimates among runs, as well as effective sample size (ESS) and PSRF values, which we evaluated in Tracer v1.5 (Drummond and Rambaut, 2007). We generated a 50% majority rule consensus phylogram using combined estimates from post burn-in samples from independent runs. We also constructed a median-joining haplotype network to visualize relationships among unique haplotypes using Network v4.5.1.6 (Bandelt et al., 1999), weighting transitions 2:1 over transversions (recommended in the Network manual to distinguish between haplotype connections that would be equally parsimonious with a 1:1 ratio) and using the maximum parsimony option to reduce excess links among haplotypes from the resulting network.

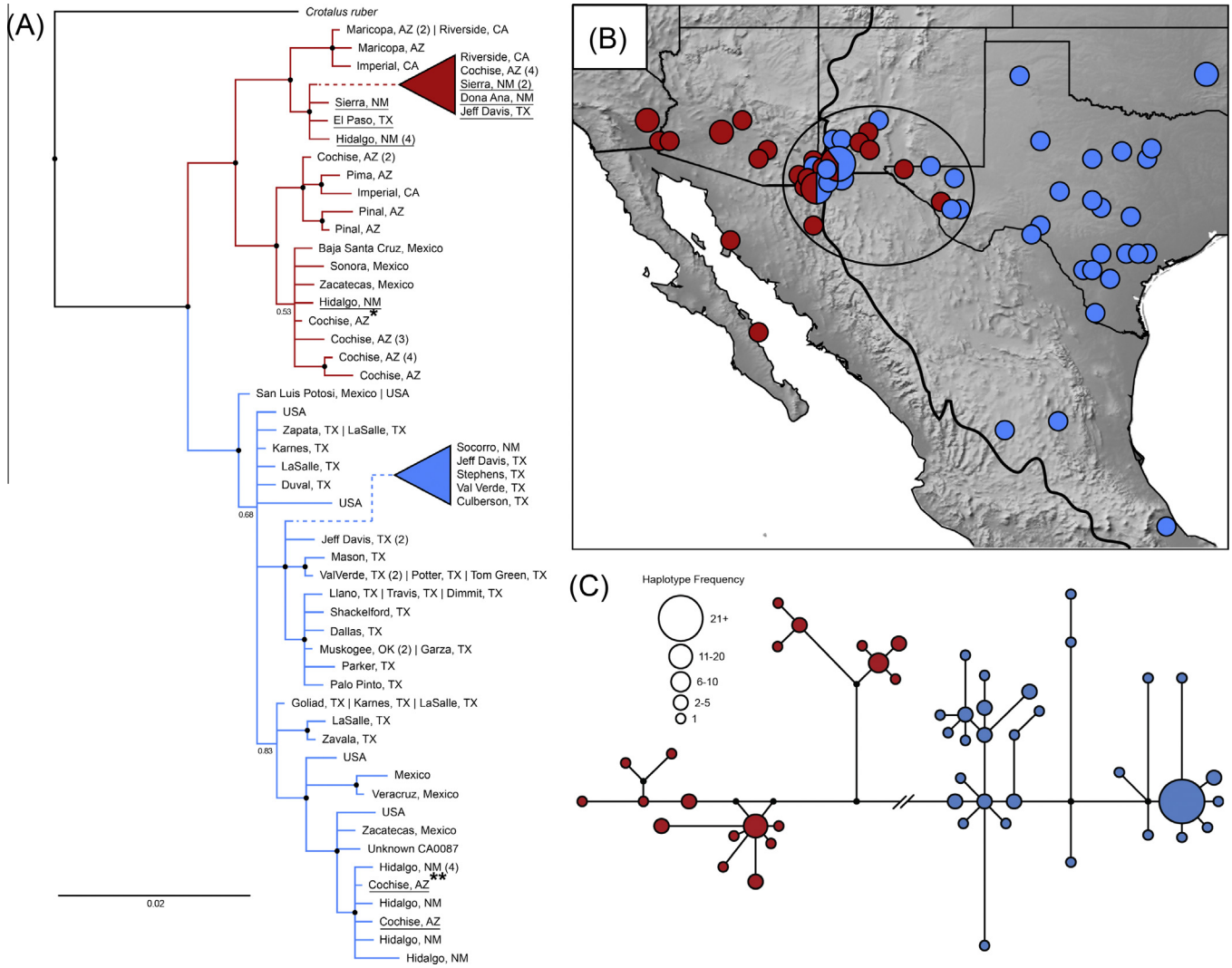

Fig. 1. Results of analyses of the ND4 mitochondrial gene. A. Bayesian phylogenetic tree estimate of relationships among condensed *C. atrox* haplotypes. The western clade is represented by red branches and the eastern clade by blue branches. Bipartition posterior probabilities greater than 0.90 are represented by black circles. Haplotypes marked with * and ** include 16 and 49 individuals, respectively. Samples falling within the western or eastern clade counter to geographic expectation are underlined. B. Map of western (red circles) and eastern (blue circles) clades, and the Continental Divide (dark line). The size of each circle corresponds to sampling frequency at that locality. For localities with pie charts, the relative contribution of each color to the circle reflects the frequency of eastern or western clade individuals. The black ellipse highlights the putative fusion region between eastern and western clades. C. Median-joining network of mtDNA haplotypes. The two mitochondrial clades are represented by blue (eastern) and red (western) circles. Circle sizes correspond to the frequency of each haplotype in our sampling. Branch lengths between circles are proportional to the number of nucleotide differences between adjacent haplotypes. Black circles represent a haplotype that is absent in our sampling. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We estimated haplotype diversity in each population using the Nei and Tajima equation (Nei and Tajima, 1981) implemented in custom Python scripts.

We analyzed changes in effective population size through time in the two primary mitochondrial lineages of *C. atrox* using the Bayesian Skyline Plot (BSP) coalescent model (Drummond et al., 2005) implemented in BEAST v1.7.5 (Drummond and Rambaut, 2007). We partitioned the dataset by gene (ND4 and tRNA) and ND4 was further partitioned by codon position (1st and 2nd position combined, 3rd position) and used the substitution model HKY, along with a strict molecular clock (recommended for intraspecific inferences in the BEAST manual) and a coalescent Bayesian skyline tree prior. We applied a 0.7% per lineage per million years mutation rate, which has been used in other studies using the ND4 gene in snakes, including *C. atrox* (Wuster et al., 2002; Castoe et al., 2007; Lane and Shine, 2011b). We conducted two independent runs of 4×10^7 generations with a burn-in of 25%. Additionally, we used IMa2 (Hey and Nielsen, 2007) to obtain estimates of the marginal posterior probability density of the parameters of the isolation-migration model (Hey and Nielsen, 2004), including t_0 (time since population splitting), q_0 (effective western population size), q_1 (effective eastern population size), q_A (effective ancestral population size), $m_1 > m_0$ (migration rate from west to east), and $m_1 < m_0$ (migration rate from east to west). We ran IMa2 using our total mtDNA dataset, with population sample sizes assigned based on the locality of each individual relative to the Continental Divide. Each of four independent runs had a 2×10^6 generation burn-in period (length was determined in initial trial runs) followed by a 10^7 generation post-burn-in sampling period, which we determined based on chain mixing and ESS values exceeding 1000 for parameters in all runs, and convergence was assessed by congruent results from independent runs. We rescaled parameter estimates using generation time and mutation rate estimates for *C. atrox* from Castoe et al. (2007).

We assessed landscape-level patterns of genetic differentiation for the eastern and western mitochondrial clades separately by interpolating pairwise mitochondrial genetic distances (mismatch distances) among sampling localities. Mismatch distances were calculated in Alleles in Space v1.0 (Miller, 2005) and assigned to geographic coordinates representing midpoints between sampling localities using the Delaunay triangulation-based connectivity network (Miller et al., 2006). To account for the correlation between genetic and geographical distances, we used residual values derived from the linear regression of mismatch distances derived from the linear regression of mismatch versus geographical distance (Manni et al., 2004). Any data points with residuals outside of the 95% confidence interval of the linear regression were removed from analysis. We imported the mismatch distances and associated coordinates into ArcGIS v9.2 (ESRI, Redlands, CA, USA) and interpolated them across 2.5-minute grids using the inverse distance weighted procedure (Watson and Philip, 1985) in the ArcGIS Spatial Analyst extension. In order to restrict the interpolation analysis within an area that is actually occupied by the species, we approximated the range of *C. atrox* via ecological niche modeling implemented in MAXENT v3.2.1 (Phillips et al., 2006) using default parameters, with 50 replicates per population, and average model probabilities converted to presence-absence maps in ArcGIS. This methodology extracts environmental data (obtained from the Worldclim dataset; (Hijmans et al., 2005) corresponding to occurrence records which were represented by geographic coordinates obtained from HerpNet (www.herpnet.org) and evaluates habitat suitability across the landscape using program-specific algorithms (Elith et al., 2006). The resulting maps were used as an approximation of the range of *C. atrox* for geographically masking the surface derived from the interpolation analyses.

To estimate how cooler climatic cycles during the Pleistocene might have restricted the range of *C. atrox*, we estimated the geographic distribution of the eastern and the western population during the last glacial maximum (LGM). We projected the present-day models for eastern and western populations of *C. atrox* (discussed above) onto climatic reconstructions of the LGM under the assumption that the climatic niche of each population remained conserved between the LGM and present time (Elith et al., 2010; Jezkova et al., 2011). For environmental layers representing the climatic conditions of the LGM, we used ocean-atmosphere simulations available through the Paleoclimatic Modelling Intercomparison Project (Braconnot et al., 2007). We used Community Climate System Model v. 3 (CCSM) that has been previously downscaled to the spatial resolution of 2.5 minutes and converted to bioclimatic variables (Waltari et al., 2007). We constructed LGM models in MAXENT using default parameters and ran models for 20 replicates per model, and obtained an average model using logistic probability classes of climatic niche suitability. Maps representing suitable climatic niche for eastern and western populations were determined using a logistic probability threshold that balances omission, predicted area, and a threshold value (Liu et al., 2005).

2.5. Analysis of RADseq data

Raw ddRADseq Illumina sequencing reads were processed using the Stacks pipeline (Catchen et al., 2013). PCR clones were removed using the Stacks clone filter program using the in-line UMI regions of our adapter design, which were subsequently trimmed away using the FASTX-Toolkit trimmer (Hannon, 2014). Trimmed reads were processed using the process radtags function in Stacks, which parses reads by barcodes, confirms the presence of restriction digest cut sites, and discards reads that lacked these features or those with poor quality scores. We used the *de novo* main Stacks pipeline, including Ustacks, Cstacks, and Sstacks, to summarize SNP information for downstream analyses.

We used the populations program in Stacks to obtain various population genetic parameters estimates. We used two alternative population assignments for analyses: (1) all individuals as a single population (these outputs were used for downstream Structure analyses); and (2) a two-population model, with individuals partitioned into western and eastern populations. For the two-population model, we only used individuals that fit the following two criteria: (1) nuclear allele assignments of >90% to either the western or eastern population cluster in Structure analyses, and (2) mitochondrial haplotypes that matched with the majority of nuclear genome assignments (e.g., >90% western nuclear alleles and containing a western mitochondrial haplotype); this population assignment model was used for genetic diversity comparisons between western and eastern populations. We used thresholds for missing data (50%) as well as minimum read depth per stack (5) in populations for all analyses. We used these thresholds to maximize the number of loci available for analyses after determining that other threshold settings (e.g., 75% missing data, 10X stack depth) did not qualitatively alter our estimates (Supplementary Online Table 2).

We inferred population structure and admixture using Structure (Pritchard et al., 2000), based on 4494 SNPs recovered by our Stacks analyses. We first estimated the allele frequency distribution parameter (λ), using a trial run with K set to 1. We used this estimation ($\lambda = 3.13$) in short clustering runs (10^4 burn-in, 10^4 data collection) under a mixed ancestry model (Hubisz et al., 2009) for $K = 2$–9 with no putative population origins specified for individuals (single population model). From this, we determined that a K range of 2–5 population clusters was most likely based on model likelihood estimates. We then performed longer runs (10^7 burn-in, 10^8 data collection) for $K = 2$–5 (each with 3 iterations). For each of these runs, the optimal K clusters was determined using the ΔK method described by Evanno et al. (2005) implemented in StructureHarvester (Earl and Vonholdt, 2012). The results of Structure runs were visualized using Distruct (Rosenberg, 2004).

To quantify genomic introgression in *C. atrox*, we conducted Bayesian estimation of genomic clines using the program bgc (Gompert and Buerkle, 2011). This program estimates the probability that an individual with a hybrid index H has inherited a gene copy at a given locus from one of two parental populations using two locus-specific cline parameters (α, the genomic cline center parameter, and β, the genomic cline rate parameter; estimated using MCMC). Under this model, if α and β both equal zero, an individual's hybrid index will be equal to the probability of ancestry (ϕ) from a given parental population (Gompert and Buerkle, 2011; Gompert et al., 2012b). We were specifically interested in genome wide estimates of both cline parameters to understand the degree of variation in locus specific introgression across the genome of our admixed population. We used data from all individuals included in our RADseq dataset, and partitioned samples into western parental, eastern parental, and admixed populations. We used allele output files for each individual from Stacks and custom Python scripts to generate biallelic input files for each population, which included information for a total of 19,123 loci (though some loci suffered from missing data for particular individuals). We ran bgc using the genotype uncertainty model (recommended for next-generation sequencing data; see bgc manual) on our parental and admixed population dataset using four chains for 50,000 generations each, discarding the first 5000 generations as burnin, and using default settings assuming free recombination between all loci (maximum distance between loci set to 0.5). We then combined the output from the four chains after inspecting the MCMC output for convergence onto a stationary distribution.

3. Results

3.1. Mitochondrial gene variation

Within *C. atrox*, we identified 52 unique ND4 haplotypes in our sampling. The four independent runs of Bayesian phylogenetic analyses using unique haplotypes converged on nearly identical estimates of the likelihood score, had PSRF values very near 1.0 throughout non burn-in generations, and had ESS values above 400 for all parameters. The 50% majority rule consensus phylogram from these runs is presented in Fig. 1A. Within *C. atrox*, we found strong support (>0.95 posterior probability) for a deep split between two haplotype lineages that correspond approximately to populations east and west of the Continental Divide of the US and Mexico (Fig. 1B). Both of these major clades, however, contain subclades that include haplotypes from individuals on either side of the Continental Divide. The two major mitochondrial clades are also apparent in the median-joining haplotype network (Fig. 1C). The eastern clade contains a greater number of haplotypes (31 of 52 haplotypes) compared to the western clade. Despite this, the western population has a higher degree of haplotype diversity (0.89) than the east (0.75), which appears to be due to many eastern samples possessing one of a small number of high-frequency haplotypes.

Our map of eastern and western haplotypes highlights a putative zone of population fusion, which we term the Inter-Pecos region, implicating the Continental Divide as a barrier to gene flow for eastern haplotypes, as very few individuals belonging to this clade were observed west of the divide (Fig. 1B). Likewise, we do not observe any western haplotypes east of the Pecos River in Texas, suggesting that this region in general represents a barrier to gene flow for the western clade. The results of our BSP analysis of historical demography in eastern and western clades indicate that both lineages have experienced recent population expansion (i.e., within the last 100 KYA). These results also suggest that the eastern population has undergone a greater increase in effective population size relative to the west (approximately 1.7 million and 3 million in the west and east, respectively; Supplementary Online Fig. 1).

Our independent IMa2 runs resulted in nearly identical estimates of marginal posterior probability densities for each parameter, and ESS values were greater than 2000 for all parameters; we present the results from the independent run with the highest ESS values in Supplementary Online Fig. 2. Based on our IMa2 analyses, we found evidence of substantial population expansion in both the eastern and western population (Supplementary Online Fig. 2A), with a greater degree of population expansion in the eastern population; these findings are all consistent with our inference from BSP analysis (Supplementary Online Fig. 1). We were unable to obtain a robust estimate of the ancestral effective population size from IMa2 analyses, however, as parameter estimation appears to have suffered from insufficient data resulting an essentially flat posterior distribution for this parameter. The posterior estimates for migration rate parameters support very low migration rates ($\sim 7 \times 10^{-8}$ migrants per year) for both populations (Supplementary Online Fig. 2B), though estimates for these parameters may have also suffered from insufficient data for accurate estimation of migration rates as indicated by posterior densities that do not reach low levels near the upper or lower limits of the prior (even after multiple runs with increasingly broad priors). We found support for a TMRCA for the eastern and western populations of approximately 675,000 years ago (95% posterior interval values ranging from 412,933 to 1,027,939; Supplementary Online Fig. 2C), which falls well within the time span of Pleistocene glaciation cycles.

When residual pairwise genetic distances are interpolated on geographic distance and are visualized across the range of both clades, we find evidence for highly contrasting levels of genetic diversity. Although there are varying degrees of diversity in each lineage, there is an overall greater degree of genetic diversity (inferred via mismatch distances) in the east (range = −0.012 to 0.029; Fig. 2B) relative to the west (range = −0.012 to 0.012; Fig. 2A), particularly in the central and western regions of Texas. Both lineages harbor relatively low diversity within areas of the Inter-Pecos fusion zone, consistent with range expansion in each lineage towards the Continental Divide. Furthermore, we find that a large portion of the high-diversity core observed in the eastern lineage falls within the eastern portion of the Inter-Pecos region, while the western lineage shows no evidence of high diversity anywhere within the Inter-Pecos.

The results of our LGM climatic niche modeling highlight major differences in estimated Pleistocene refugia for the eastern and western populations, and provide context for our observed patterns of contrasting diversity in western versus eastern populations (Fig. 2C). At all thresholds, we find that the eastern population inhabited a large region of the Chihuahuan Desert and adjacent Gulf Coastal Plains. This region is comprised of multiple large segments that are largely in contact; thus, the eastern refugium exhibits little fragmentation, whereas the western refugium occupies a much smaller and more linear range. Our models also demonstrate that the eastern population occupied part of the Inter-Pecos region during the LGM, while the western population was completely absent. Additionally, these models suggest that the eastern and western populations of *C. atrox* were most likely not in contact during Pleistocene glacial cycles, and were instead isolated by an extensive region of high-elevation.

3.2. Nuclear SNP variation

RADseq filtering thresholds of 50% missing data per locus and 5X read depth per locus provided the most numerous RAD loci (4,519) for analyses among threshold combinations used under the two-population model explained above, and genetic diversity estimates did not vary qualitatively with this filtering scheme relative to others (for results from Stacks populations analyses under different filtering thresholds, see Supplementary Online Table 2). We find substantial population differentiation among western and eastern populations based on these data ($F_{ST} = 0.15$). This estimate is consistent with moderate-to-high levels of differentiation (Lewontin, 1972; Lewontin and Krakauer, 1973), and emphasizes that, prior to secondary contact, eastern and western lineages were well-differentiated incipient species.

Our nuclear SNP dataset provided consistent evidence of higher genetic diversity in the eastern population relative to the west. We find a much higher number of private alleles in the eastern population than in the western population (453 and 253, respectively; Fig. 3A). Observed heterozygosity is significantly higher in the eastern population than in the western population ($p < 0.0001$; Fig. 3B). We find the same pattern in the estimates of nucleotide diversity (π), with significantly higher diversity in the east ($p < 0.0001$; Fig. 3C). Altogether, the pattern of higher nuclear genetic diversity in the eastern population is consistent with similar patterns observed in our mitochondrial dataset.

In our final Structure analyses, we estimated $K = 4$ as the optimal model of population clustering, and populations assignments are shown in Fig. 3D. Additionally, we visualized the results of the two-cluster ($K = 2$) model for comparison, because a two-population model was most consistent with the general findings from our mitochondrial data. We find that both two and four population models recover similar population structure, such that there are identifiable western and eastern clusters. Both models also indicate a similar fusion zone between lineages, with the majority of individuals within the zone showing evidence of substantial

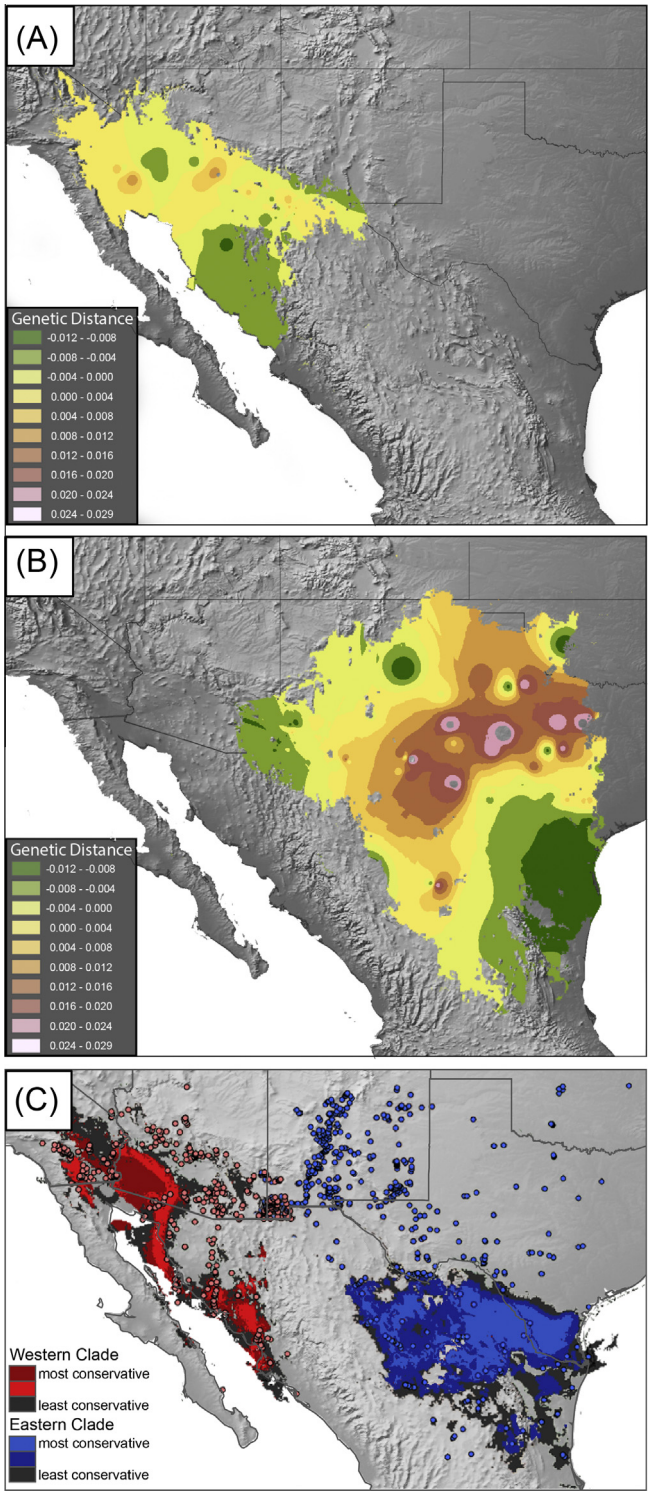

Fig. 2. Landscape genetic diversity estimates and Last Glacial Maximum (LGM) ecological niche models. A-B. Residual pairwise genetic distances from the calculated linear regression interpolated across landscape for western (A) and eastern (B) mitochondrial clades of *C. atrox*. Interpolations are restricted by ecological niche models used to determine suitable habitat for each population and by the distribution of sampled localities. C. LGM models for western (red) and eastern (blue) populations at three stringency thresholds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

admixture. These samples predominately fall within the Inter-Pecos region as defined by our mitochondrial dataset, although nuclear SNPs highlight the Continental Divide as the location of a very steep cline of genotype intergradation between eastern and western population clusters (also see Section 3.3. below).

When we ordered samples by longitude (as shown in Fig. 3D), we found a visible relationship between the longitudinal gradient from west to east and relative levels of population cluster assignment, suggesting a cline of genetic introgression of western and eastern populations. We find that as the complexity of our model increases from $K = 2$ to $K = 4$, so does evidence of increased structure in the eastern population (Fig. 3D). Additionally, the more complex model recovered a genetic cluster that is endemic to the Inter-Pecos region and northern Chihuahuan Desert (Fig. 3E), consistent with the hypothesis that the eastern population has been long-established in this area (e.g. prior to the LGM). To further test this, we examined the relative frequencies of private alleles (from nuclear SNP data) for the entire eastern population used for diversity estimates versus a subset that falls within this more restricted geographic zone of high mitochondrial diversity. Because these compared sets of individuals differ in the number of loci available after filtering and in the number of individuals, we standardized our comparison by the number of loci and numbers of individuals to obtain a relative frequency of private alleles. We find that the entire eastern population sampling (for nuclear data) has a frequency of approximately 0.17 private alleles per individual per locus, while the eastern sampling limited to the inferred endemic diversity region has a frequency of 0.20, corresponding to a 17% increase in private alleles in this region. Thus the mitochondrial and nuclear data broadly agree in identifying a region of high endemic diversity in the eastern population centered in Texas (e.g., Fig. 2B).

We found evidence for variable introgression across the genome for the admixed population of *C. atrox* in our bgc analyses (Fig. 4). In particular, we found median estimates of the genomic cline center parameter α to vary moderately among the loci sampled (minimum = −0.0127, maximum = 0.0147) and the 95% confidence interval for α encompassed zero for all loci, indicating a lack of excess ancestry from either parental population. The genomic cline rate parameter β was slightly less variable (minimum = −0.0108, maximum = 0.0116) and also had a 95% confidence interval that encompassed zero for all loci. Collectively, the variation in these parameters is less extreme than that observed in other recently studied systems (e.g., Gompert et al., 2012a; Lindtke et al., 2012), and would indicate that, though genome wide variation in introgression is occurring in *C. atrox*, the degree to which parental population ancestry is observed appears to be driven by drift over selection (which would be inferred in the case of large proportions of excess ancestry from one parental population over the other).

3.3. Relationship between mitochondrial and nuclear genomes

In addition to the formal tests of genomic introgression in our nuclear dataset, we examined the longitudinal relationship of our mitochondrial and nuclear genetic datasets to identify visible clines or shifts in genetic composition across their distribution. We combined samples from both datasets into two-degree longitudinal bins and calculated eastern versus western mitochondrial haplotype frequencies (per longitudinal bin), as well as average western or eastern nuclear genome proportion (inferred via population assignments of individuals under the two population model in Structure). We then plotted these values to compare nuclear and mitochondrial patterns (Fig. 5). We find a relatively steep gradient of nuclear composition shift between 110 and 108 degrees W, corresponding to the Continental Divide (and approximately the Arizona-New Mexico border), supporting this region as a barrier to western nuclear alleles moving eastward and vice versa. In contrast, the region of mitochondrial overlap is considerably broader and the shift between higher frequency western or eastern haplo-

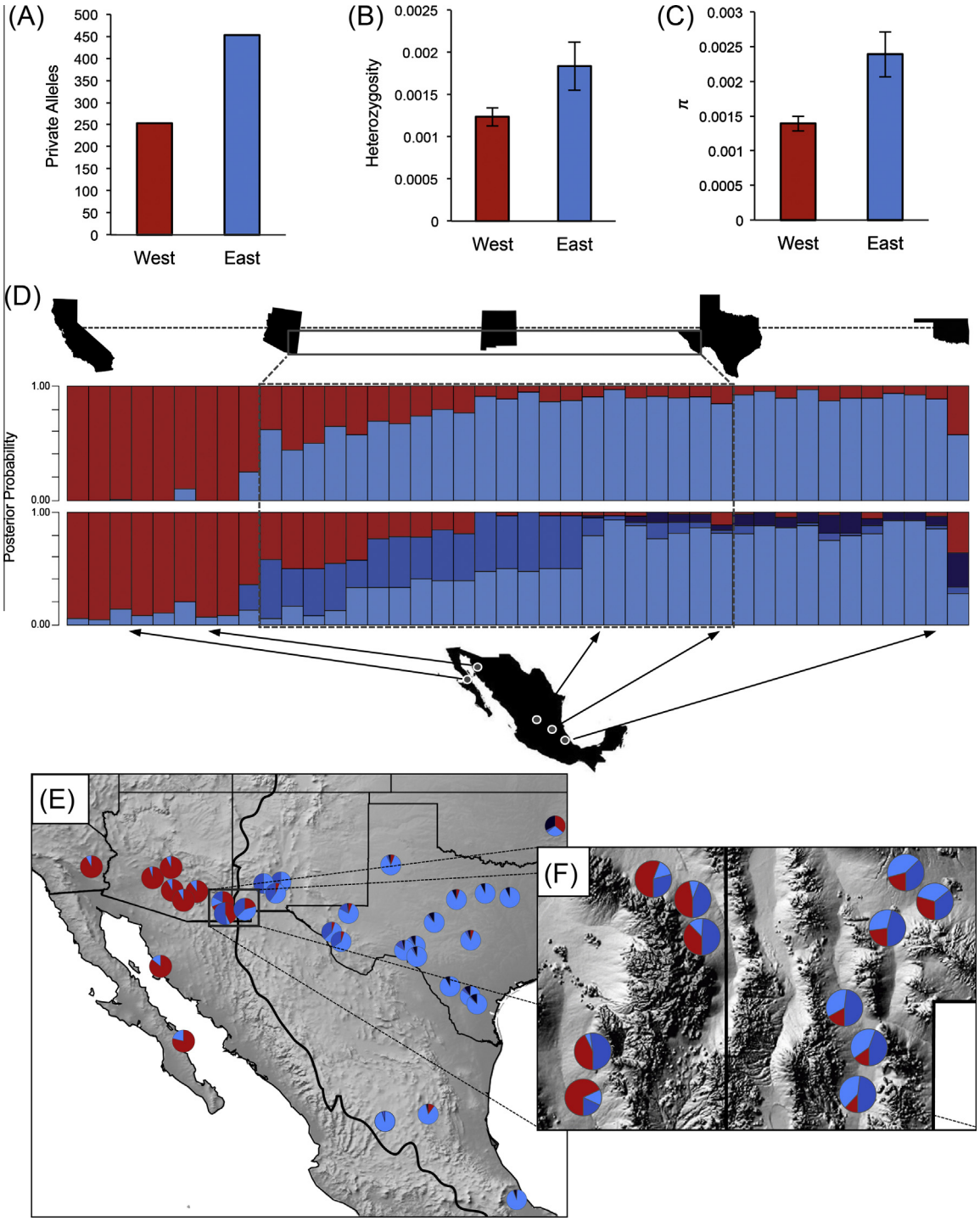

Fig. 3. Patterns of genetic structure and diversity based on nuclear SNP data. A., B., and C. Estimates of private alleles, heterozygosity, and nucleotide diversity (respectively) for western and eastern populations inferred from analyses in *Stacks*. Error bars represent the variance in estimates of heterozygosity and nucleotide diversity. D. Structure plots for $K = 2$ (top) and $K = 4$ (bottom) population cluster models. Individuals are represented by individual bars of the posterior probability of their assignment to various population clusters. Individuals are ordered longitudinally from west to east. The putative region of lineage fusion (introgression) is highlighted in grey, and the dashed grey box includes samples that fall within this region. E. Map of samples used in nuclear SNP analyses. Individuals are represented by pie charts with colors corresponding to the structure plot for the $K = 4$ model, and the Continental Divide is represented by the bold black line. The grey box highlights a region of dense sampling at the continental divide, which is shown in more detail in the inset to the right (F). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

types is less steep than the shift observed in the nuclear data (Fig. 5. There is also an immediate decrease to zero in observed western mitochondrial haplotypes east of the Inter-Pecos region (east of the Pecos River), and the same pattern is observed for eastern haplotypes west of the Continental Divide. While the greatest degree of shifting nuclear genome composition is observed within the discrete region mentioned above, there is evidence of small amounts of nuclear introgression throughout the entire longitudinal gradient. Thus, while mitochondrial introgression is confined to within the Inter-Pecos region, nuclear gene flow appears to

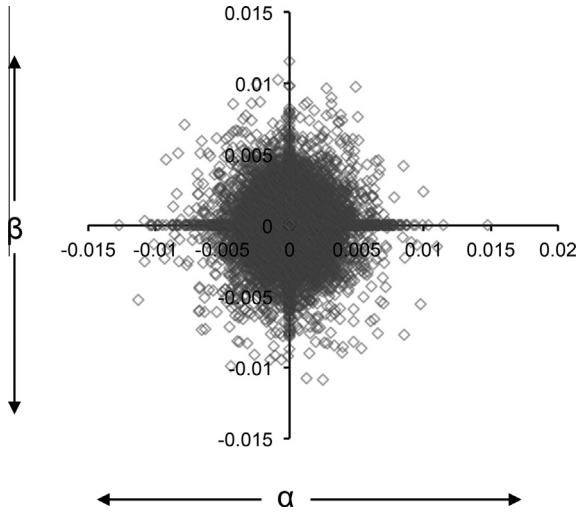

Fig. 4. Variable levels of introgression in the admixed *C. atrox* population as determined by the genomic cline center parameter α and the genomic cline rate parameter β estimated in bgc for each nuclear locus. Each individual locus is represented by a grey box.

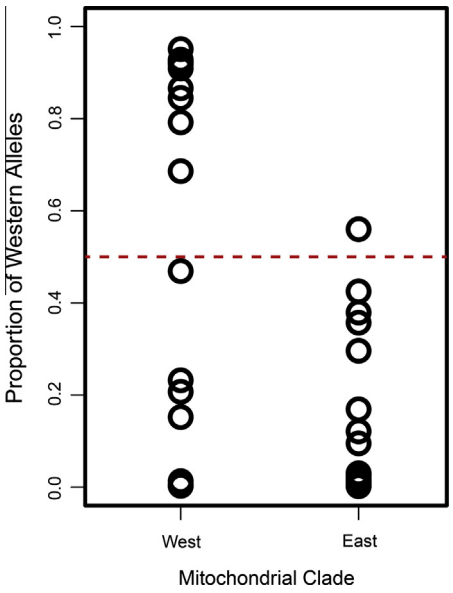

Fig. 6. Deficiency of individuals with eastern mitochondrial haplotypes and moderate to high levels of western nuclear alleles. Comparison of mitochondrial haplotype clade membership with relative proportion of western nuclear alleles inferred from *Structure* analysis based on a two population model. The red line indicates a 50% threshold of western alleles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

penetrate far beyond the western and eastern geographic barriers of this region.

To test the hypothesis of unbiased gene flow between eastern and western populations of *C. atrox*, we compared mitochondrial and nuclear estimates of ancestral genetic origins of individuals for which we had both datasets (Fig. 6). If gene flow is completely unrestricted between populations we would expect to see mitochondrial haplotypes from one lineage existing with a range of different nuclear backgrounds (i.e., individuals belonging to the western mitochondrial clade may have a range of proportions of alleles from western and eastern nuclear clusters). We might also expect to see this range weighted such that nuclear genetic composition would tend to agree with the mitochondrial clade an individual belongs to, including, for example, western mitochondrial clade individuals having a higher proportion of western nuclear cluster alleles.

We find that there is a lack of high proportion western nuclear allele individuals with eastern mitochondrial haplotypes (Fig. 6). We do see evidence of the reverse, however, with several low proportion western nuclear allele individuals belonging to the western mitochondrial clade. This pattern is particularly interesting given the close geographic proximity of many of our samples, especially those within the western Inter-Pecos region (Fig. 3F). Eastern mitochondria only occur with roughly 50% or less western nuclear genome proportion, and most often possess a much lower frequency of western alleles (<20%). Even at the steep cline of gene flow between eastern and western populations at the Continental Divide (Fig. 3F), it is notable that eastern nuclear alleles seem to penetrate this barrier while eastern mitochondrial haplotypes do not.

4. Discussion

4.1. Evidence of two distinct yet introgressing lineages within C. atrox

We find evidence from both mitochondrial and nuclear datasets that *Crotalus atrox* comprises two well-differentiated lineages. Indeed, phylogenetic and network analyses of our mitochondrial

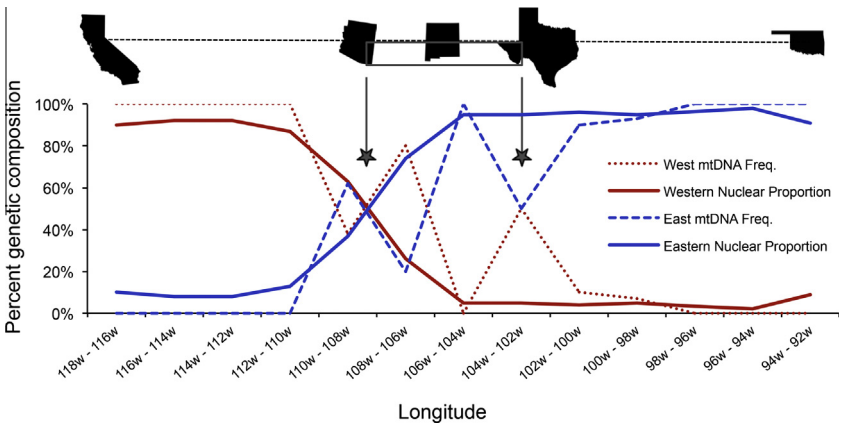

Fig. 5. Clines of mitochondrial and nuclear genetic composition across a longitudinal gradient. Estimates of genetic composition are organized in two-longitudinal-degree bins from west to east. Values of western and eastern nuclear allele proportions per longitudinal bin are represented by red and blue solid lines, respectively. Western and eastern mitochondrial haplotype frequencies per longitudinal bin are represented by dashed red and blue lines. Points of interest (the Continental Divide and eastern extreme of the Inter-Pecos region) are highlighted with black stars, which also highlight notable shifts in nuclear or mitochondrial genetic content. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dataset, and Bayesian clustering analyses of our nuclear SNP dataset collectively highlight two distinct groups of *C. atrox*, and our estimate of nuclear F_{ST} (0.15) further argues for the distinction of these lineages. We also find evidence that these lineages have introgressed following their initial isolation during the Pleistocene. The mitochondrial zone of introgression occurs across the Inter-Pecos, a broad region stretching from the Continental Divide and the Pecos River in Texas. We find explicit evidence of admixture among individuals in our nuclear SNP data, such that the majority of individuals have partial assignment to at least two population clusters under both 2 and 4 population models. Interestingly, evidence for nuclear introgression expands beyond the Inter-Pecos region on either side of the mitochondrial introgression zone, though we find the greatest degree of admixture between western and eastern clusters closest to the Continental Divide, at the western edge of the Inter-Pecos (Fig. 5).

4.2. Pleistocene isolation, expansion and a broad zone of introgression

It is evident from our sampling that individuals belonging to both eastern and western clades do co-occur, despite evidence of historical isolation. While the Continental Divide has previously been thought to be the major barrier separating eastern and western lineages, our data show that the introgression zone between populations is much larger than previously thought. Our mitochondrial data show contrasting patterns of genetic structure in eastern and western populations, with eastern mitochondrial haplotypes tending to be more localized than western haplotypes. In contrast, we found a single western haplotype that extends from California to West Texas, consistent with recent population expansion inferred from our demographic analyses (Supplementary Online Figs. 1 and 2). Our landscape diversity estimates show particularly low genetic diversity in both lineages near the Continental Divide, and throughout much of the Inter-Pecos region (Fig. 2A and B), consistent with evidence of population expansion in both lineages (Supplementary Online Fig. 1) following Pleistocene glacial cycles.

Last Glacial Maximum (LGM) Pleistocene models suggest that eastern and western populations were separated by a wide central plateau in the Inter-Pecos region. This corroborates our inferences that the two lineages diverged in isolation for a period during the Pleistocene, prior to expanding their ranges to meet in the Inter-Pecos region since the LGM. These data together with our population genetic data indicate recent secondary contact after range expansion of both eastern and western lineages of *C. atrox* across the Inter-Pecos region. Interestingly, our LGM model predicts that the Inter-Pecos region was partially inhabited by the eastern population during the LGM. This suggests that the eastern lineage may have been present across much of the Inter-Pecos region prior to expansion of western lineages into this region, which is supported by high levels of endemic nuclear and mitochondrial diversity in some eastern portions of this region (Figs. 2B and 3D). Given that our estimate of time since population splitting greatly predates the LGM, it is possible that secondary contact between these lineages could have occurred multiple times during recessions in glaciation cycles, though this would not necessarily alter our inference of isolation followed by secondary (or tertiary) mixing.

Our results reject the hypothesis that population genetic diversity is similar between eastern and western populations of *C. atrox*. We find substantially higher levels of local genetic diversity in the eastern population relative to the western population based on nuclear SNPs (Fig. 3A–C), and mitochondrial DNA (Fig. 2A and B). It is likely that multiple factors have influenced the contrasting levels of diversity we observe, including historical demography, past and present population range size, and the diversity of habitats that occur within the distribution of each population. The eastern population is predicted to have a larger and non-linear range during the LGM, suggesting Pleistocene population sizes may have been larger in the east. Additionally, the eastern population currently occupies a region of substantially higher habitat diversity relative to the west, including forest, grasslands, and Chihuahuan Desert, in contrast to the western population that occurs primarily within arid desert habitat (Campbell and Lamar, 2004).

The oldest known fossil of *C. atrox* dates from between 3.7 and 3.2 MYA (Holman, 2000) and was found in north central Texas, while fossils found west of the Continental Divide are restricted to much later dates during the Late Pleistocene (Holman, 1995). These fossil data suggest that the eastern population may have existed in its current range for millions of years, and substantially longer than the western population has. These data may also indicate that the range currently occupied by the eastern population is the ancestral range for the species prior to isolation of western and eastern populations due to the waxing and waning of glaciation during the Pleistocene. These data also suggest that the relative age and stability of the eastern population may have contributed to its greater genetic structure and diversity.

4.3. Evidence for incipient yet failed speciation, and sex-biased gene flow

The study of genetic incompatibilities early in the speciation continuum is essential to understanding mechanisms that drive speciation because they provide greater insight into primary causative mechanisms that lead to reproductive isolation (Orr, 1995). Given broad evidence for two divergent lineages within *C. atrox* that appear to have evolved in isolation until recent secondary contact, we were interested to test for evidence of non-random gene flow that might indicate the evolution of reproductive isolation between these lineages. While our mitochondrial and nuclear data broadly agree that two distinct lineages of *C. atrox* are currently exchanging genes across a relatively large region in the center of the species' range, patterns of introgression differ between these two datasets. While the mitochondrial introgression zone appears to be flanked by two hard boundaries – the Continental Divide to the west and the Pecos River to the east – our nuclear data indicates mixing of alleles that extends far beyond these boundaries. Nuclear data also provide evidence for a remarkably steep gradient of genotypic composition at the Continental Divide.

Assuming random mating and equal fitness of offspring, we would expect that mitochondrial genotypes might predict nuclear genotypes. For example, if an individual possesses an eastern mitochondrial haplotype it would be more likely that it contained a substantial proportion of eastern nuclear alleles, and vice-versa. However, sex-biased gene flow, as well as selection against certain combinations of mitochondrial and nuclear genotypes, may alter the expected relationships between mitochondrial and nuclear genotypes. We tested this relationship and found it to be notably asymmetrical, with western mitochondrial haplotypes associated with a wide range of western nuclear allelic content, yet eastern mitochondrial haplotypes only associated with 50% or greater eastern nuclear allelic content (Fig. 6). Thus, we did not observe individuals with eastern mitochondrial haplotypes with greater than 50% western nuclear allelic content; in most cases, eastern mitochondrial haplotypes were paired with much lower levels of western nuclear alleles. These data are consistent with a low frequency of successful eastern female × western male mating, compared to high frequencies of western females mating with males from either population, which might be explained by sex-biased dispersal. Sex-biased gene flow is likely in snakes, given evidence for sex-biased dispersal (Keogh et al., 2007; Dubey et al., 2008; Lane and Shine, 2011a). In rattlesnakes, males tend to be more widely dispersing (Duvall et al., 1992), however, a recent study found higher than expected female dispersal specifically in

C. atrox (Schuett et al., 2013). Our results show relatively high frequencies of western haplotypes, yet low frequencies of western nuclear allelic content throughout much of the introgression zone. If this pattern were generated by sex-biased gene flow, it would require that western females have substantially higher dispersal capabilities than western males, and that dispersal for the eastern males and females is roughly equal.

A second intriguing possibility is that these results are due to a post-zygotic mechanism in which mito-nuclear incompatibilities lower the fitness of offspring that possess eastern mitochondrial haplotypes and a genetic background with high proportions of western nuclear alleles. Such an incompatibility might also be responsible for creating the steep geographic cline of allelic content at the Continental Divide (Fig. 3F), selecting against higher proportions of western alleles in populations with higher proportions of eastern mitochondria. Mito-nuclear incompatibilities have been shown in other species to be early causative drivers of post-zygotic genetic isolation and speciation (Ulloa et al., 1995; Bogdanova, 2007; Presgraves, 2010). It would be interesting in future studies to test competing hypotheses for what mechanisms might be driving these patterns of mito-nuclear genotypic content in *C. atrox*, to determine if there are genetic incompatibilities that limit gene flow between eastern and western *C. atrox* populations. A potential difficulty with this hypothesis is the lack of evidence for introgression driven by selection in sampled nuclear loci within the admixed *C. atrox* population. The likelihood of one of these neutral loci being physically linked to a putative locus responsible for mito-nuclear incompatibility, however, is very low and further investigation is warranted to confirm the presence and nature of such an incompatibility. The divergence of the two main lineages of *C. atrox*, together with evidence for biased gene flow and potential post-zygotic isolating mechanisms, argue that these two lineages were at some intermediate stages along the speciation continuum prior to their secondary extensive mixing.

Based on our findings, we conclude that *C. atrox* represents a single species, comprised of two divergent populations that are experiencing ongoing widespread gene flow that extends essentially to the margins of their entire distribution. Our results support the hypothesis that, at some time in the past, *C. atrox* comprised two well-differentiated lineages that represented incipient species that were each evolving in partial genetic isolation. While this extended period of mutual isolation placed these two lineages at an intermediate stage of the speciation continuum, subsequent pervasive gene flow, presumably after substantial geographic population expansion following the Pleistocene, has apparently reversed the progression of these two lineages along this continuum, and broadly mixed genetic variation between these lineages. The extent of gene flow among lineages we have observed in this study leads to the consideration of *C. atrox* as a single species based on a number of species concepts, including the biological (Mayr, 1963), the evolutionary (Simpson, 1951; Wiley, 1978), and the general lineage concept (de Queiroz, 1998). Recently, species concepts that recognize speciation with gene flow have been favored in some cases (Feder et al., 2012), and there is empirical evidence that such processes may be fairly common in nature (Nosil, 2008; Pinho and Hey, 2010). Speciation with gene flow has, however, been primarily documented in instances where gene flow is confined to localized regions of sympatry and peripatry (Leache et al., 2013; Martin et al., 2013; Osborne et al., 2013). The broad extent and penetrance of gene flow across nearly the entire range of C. atrox, together with evidence that these two lineages appear to be expanding and, if anything, mixing to a greater extent as time progresses, suggests that *C. atrox* represents an example of failed speciation rather than an example of speciation with gene flow.

Acknowledgments

We thank Jens Vindum and the California Academy of Sciences, Robert Murphy and the Royal Ontario Museum, Jesse Meik, and Corey Roelke for providing tissue samples; Jill Castoe, Rachel Wostl, and Nicole Hales for help with laboratory work; and Matthew Fujita for helpful Perl scripts. We thank Jeff Streicher and Eric Watson for comments and suggestions. Support for this work was provided by faculty startup funds from the University of Texas at Arlington to TAC and by a Phi Sigma Beta Phi Chapter research grant to DRS.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2014.12.006.

References

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. Syst. Biol. 43, 387–418.

Ashton, K.G., de Queiroz, A., 2001. Molecular systematics of the western rattlesnake, Crotalus viridis (Viperidae), with comments on the utility of the D-loop in phylogenetic studies of snakes. Mol. Phylogenet. Evol. 21, 176–189.

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise, J.C., Vrijenhoek, R.C., 1987. Mode of inheritance and variation of mitochondrial-DNA in hybridogenetic fishes of the genus Poeciliopsis. Mol. Biol. Evol. 4, 514–525.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bogdanova, V.S., 2007. Inheritance of organelle DNA markers in a pea cross associated with nuclear-cytoplasmic incompatibility. Theor. Appl. Genet. 114, 333–339.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C.D., Kageyama, M., Kitoh, A., Laine, A., Loutre, M.F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S.L., Yu, Y., Zhao, Y., 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale features. Clim. Past 3, 261–277.

Burbrink, F.T., Castoe, T.A., 2009. Molecular phylogeography of snakes. In: Seigel, R., Mullin, S. (Eds.), Snakes: Ecology and Conservation. Cornell University Press, Ithaca, NY.

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

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.

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

de Queiroz, K., 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, Oxford, England.

Douglas, M.E., Douglas, M.R., Schuett, G.W., Porras, L.W., Holycross, A.T., 2002. Phylogeography of the western rattlesnake (*Crotalus viridis*) complex, with emphasis on the Colorado Plateau. In: Schuett, G.W., Hoggren, M., Douglas, M.E., Greene, H.W. (Eds.), Biology of the Vipers. Eagle Mountain Publishing.

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

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192.

Dubey, S., Brown, G.P., Madsen, T., Shine, R., 2008. Male-biased dispersal in a tropical Australian snake (Stegonotus cucullatus, Colubridae). Mol. Ecol. 17, 3506–3514.

Duvall, D., Arnold, S.J., Schuett, G.W., 1992. Pitviper mating systems: ecological potential, sexual selection, and microevolution. In: Campbell, J.A., Brodie, E.D., (Eds.), Biology of the Pitvipers. Selva, Tyler, TX.

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

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic 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.

Gompert, Z., Buerkle, C.A., 2011. Bayesian estimation of genomic clines. Mol. Ecol. 20, 2111–2127.

Gompert, Z., Lucas, L.K., Nice, C.C., Fordyce, J.A., Forister, M.L., Buerkle, C.A., 2012a. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. Evolution 66, 2167–2181.

Gompert, Z., Parchman, T.L., Buerkle, C.A., 2012b. Genomics of isolation in hybrids. Philos. Trans. Royal Soc. B-Biol. Sci. 367, 439–450.

Good, J.M., Handel, M.A., Nachman, M.W., 2008. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. Evolution 62, 50–65.

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

Hey, J., Nielsen, R., 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D-persimilis. Genetics 167, 747–760.

Hey, J., Nielsen, R., 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. U.S.A. 104, 2785–2790.

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.

Holman, J.A., 1995. Pleistocene Amphibians and Reptiles of North America. Oxford University Press, New York, NY.

Holman, J.A., 2000. Fossil Snakes of North America: Origion, Evolution, Distribution, Paleoecology. Indiana University Press, Bloomington, IN.

Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resources 9, 1322–1332.

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

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*). Glob. Change Biol. 17, 3486–3502.

Keogh, J.S., Webb, J.K., Shine, R., 2007. Spatial genetic analysis and long-term mark-recapture data demonstrate male-biased dispersal in a snake. Biol. Lett. 3, 33–35.

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

Lane, A., Shine, R., 2011a. Intraspecific variation in the direction and degree of sex-biased dispersal among sea-snake populations. Mol. Ecol. 20, 1870–1876.

Lane, A., Shine, R., 2011b. Phylogenetic relationships within laticaudine sea snakes (Elapidae). Mol. Phylogenet. Evol. 59, 567–577.

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.

Leache, A.D., Harris, R.B., Maliska, M.E., Linkem, C.W., 2013. Comparative species divergence across eight triplets of spiny lizards (Sceloporus) using genomic sequence data. Genome Biol. Evol. 5, 2410–2419.

Lewontin, R.C., 1972. The apportionment of human diversity. In: Dobzhansky, T.H., Hecht, M.K., Steere, W.C., (Eds.) Evolutionary Biology.

Lewontin, R.C., Krakauer, J., 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74, 175–195.

Lindtke, D., Buerkle, C.A., Barbara, T., Heinze, B., Castiglione, S., Bartha, D., Lexer, C., 2012. Recombinant hybrids retain heterozygosity at many loci: new insights into the genomics of reproductive isolation in Populus. Mol. Ecol. 21, 5042–5058.

Liu, C.R., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28, 385–393.

Manni, F., Guerard, E., Heyer, E., 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Hum. Biol. 76, 173–190.

Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., Jiggins, C.D., 2013. Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res. 23, 1817–1828.

Mayr, E., 1963. Animal Species and Evolution. Belknap Press of Harvard University Press, Cambridge, MA.

Meik, J.M., Streicher, J.W., Mocino-Deloya, E., Setser, K., Lazcano, D., 2012. Shallow phylogeographic structure in the declining Mexican lance-headed rattlesnake, *Crotalus polystictus* (Serpentes: Viperidae). Phyllomedusa 11, 3–11.

Miller, M.P., 2005. Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. J. Hered. 96, 722–724.

Miller, M.P., Bellinger, M.R., Forsman, E.D., Haig, S.M., 2006. Effects of historical climate change, habitat connectivity, and vicariance on genetic structure and diversity across the range of the red tree vole (*Phenacomys longicaudus*) in the Pacific Northwestern United States. Mol. Ecol. 15, 145–159.

Nei, M., Tajima, F., 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97, 145–163.

Nosil, P., 2008. Speciation with gene flow could be common. Mol. Ecol. 17, 2103–2106.

Nosil, P., Feder, J.L., 2012. Genomic divergence during speciation: causes and consequences introduction. Philos. Trans. Royal Soc. B-Biol. Sci. 367, 332–342.

Orr, H.A., 1995. The population-genetics of speciation – the evolution of hybrid incompatibilities. Genetics 139, 1805–1813.

Osborne, O.G., Batstone, T.E., Hiscock, S.J., Filatov, D.A., 2013. Rapid speciation with gene flow following the formation of Mt. Etna. Genome Biol. Evol. 5, 1704–1715.

Palumbi, S.R., Baker, C.S., 1994. Contrasting population-structure from nuclear intron sequences and Mtdna of Humpback Whales. Mol. Biol. Evol. 11, 426–435.

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.

Pinho, C., Hey, J., 2010. Divergence with gene flow: models and data. Annu. Rev. Ecol. Evol. Syst. 41 (41), 215–230.

Presgraves, D.C., 2010. The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11, 175–180.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

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.

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138.

Schuett, G.W., Repp, R.A., Amarello, M., Smith, C.F., 2013. Unlike most vipers, female rattlesnakes (*Crotalus atrox*) continue to hunt and feed throughout pregnancy. J. Zool. 289, 101–110.

Simpson, G.G., 1951. The species concept. Evolution 5, 285–298.

Spencer, C.L., 2008. Geographic variation in western diamond-backed rattlesnake (*Crotalus atrox*) morphology. 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, pp. 55–78.

Taylor, E.B., Boughman, J.W., Groenenboom, M., Sniatynski, M., Schluter, D., Gow, J.L., 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. Mol. Ecol. 15, 343–355.

Ulloa, M., Corgan, J.N., Dunford, M., 1995. Evidence for nuclear-cytoplasmic incompatibility between Allium-Fistulosum and Allium-Cepa. Theor. Appl. Genet. 90, 746–754.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyari, A.S., Perkins, S.L., Guralnick, R.P., 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS One 2.

Watson, D.F., Philip, G.M., 1985. A refinement of inverse distance weighted interpolation. Geo-Processing 2, 315–327.

Webb, W.C., Marzluff, J.M., Omland, K.E., 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. Mol. Ecol. 20, 2390–2402.

Wiens, J.J., Engstrom, T.N., Chippindale, P.T., 2006. Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (Genus plethodon): testing the hybrid swarm hypothesis of rapid radiation. Evolution 60, 2585–2603.

Wiley, E.O., 1978. Evolutionary species concept reconsidered. Syst. Zool. 27, 17–26.

Wuster, W., da Graca Salomao, M., Quijada-Mascarenas, J.A., Thorpe, R.S., Project, B.-B.B.S., 2002. Origin and evolution of the South American pitviper fauna: evidence from mitochondrial DNA sequence data. In: Schuett, G.W., Hoggren, M., Douglas, M.E., Green, H. (Eds.), Biology of the Vipers. Eagle Mountain Publishing, Salt Lake City, UT.