Multilocus phylogeographic assessment of the California Mountain Kingsnake (*Lampropeltis zonata*) suggests alternative patterns of diversification for the California Floristic Province

E. A. MYERS,*† J. A. RODRÍGUEZ-ROBLES,‡ D. F. DENARDO,§ R. E. STAUB,¶ A. STROPOLI,** S. RUANE*† and F. T. BURBRINK*†

*Department of Biology, The Graduate School, City University of New York, NY 10016, USA, †Department of Biology, 6S-143, College of Staten Island, 2800 Victory Boulevard, Staten Island, NY 10314, USA, ‡School of Life Sciences, University of Nevada, 4505 Maryland Parkway, Las Vegas, NV 89154-4004, USA, §School of Life Sciences, Arizona State University, Tempe, AZ 85287-4501, USA, ¶332 Fiesta Avenue, Davis, CA 95616-0207, USA, **Science, Math and Engineering, Staten Island Technical High School, 485 Clawson Street, Staten Island, New York 10306, USA

**Abstract**

Phylogeographic inference can determine the timing of population divergence, historical demographic processes, patterns of migration, and when extended to multiple species, the history of communities. Single-locus analyses can mislead interpretations of the evolutionary history of taxa and comparative analyses. It is therefore important to revisit previous single-locus phylogeographic studies, particularly those that have been used to propose general patterns for regional biotas and the processes responsible for generating inferred patterns. Here, we employ a multilocus statistical approach to re-examine the phylogeography of *Lampropeltis zonata*. Using nonparametric and Bayesian species delimitation, we determined that there are two well-supported species within *L. zonata*. Ecological niche modelling supports the delimitation of these taxa, suggesting that the two species inhabit distinct climatic environments. Gene flow between the two taxa is low and appears to occur unidirectionally. Further, our data suggest that gene flow was mediated by females, a rare pattern in snakes. In contrast to previous analyses, we determined that the divergence between the two lineages occurred in the late Pliocene (c. 2.07 Ma). Spatially and temporally, the divergence of these lineages is associated with the inundation of central California by the Monterey Bay. The effective population sizes of the two species appear to have been unaffected by Pleistocene glaciation. Our increased sampling of loci for *L. zonata*, combined with previously published multilocus analyses of other sympatric species, suggests that previous conclusions reached by comparative phylogeographic studies conducted within the California Floristic Province should be reassessed.

**Keywords**: landscape genetics, niche modelling, phylogeography, reptiles, speciation

**Received 21 March 2013; revision received 23 July 2013; accepted 25 July 2013**

**Introduction**

Biologists need to identify cryptic species to properly document the Earth’s biodiversity. Although a majority of taxa are described using morphological data and techniques, molecular data are necessary for detecting cryptic species (Bickford et al. 2007). Using multiple loci in a coalescent framework provides an objective, comparable method for delimiting taxa. This approach also allows researchers to make inferences about the timing of divergence and mechanisms of speciation (Fujita et al. 2012). Additionally, studies have shown that proper enumeration of taxa is essential for ecological
and comparative evolutionary studies (Bickford et al. 2007; Smith et al. 2012). While species delimitation is not always the goal of phylogeography, the biodiversity assessments that result from these studies provide essential information for proposing evidence-based conservation strategies. Therefore, delimiting cryptic taxa is an important and urgent task, as these may be the only systematic treatment of the group of interest in the foreseeable future.

Because any single locus may be susceptible to introgression, selection or stochastic processes, multilocus data are necessary for delimiting cryptic species and determining the mechanism of speciation (Nosil 2008; Dupuis et al. 2012). Further, comparative phylogeography, which can identify biogeographic features that have structured populations of codistributed organisms (Hickerson et al. 2010), has largely relied on single locus estimates to infer community-wide patterns. However, it is not well understood how using multiple, single-locus studies from different taxa may distort interpretations of the biogeography of a region. Therefore, multi-locus data are also essential for identifying geographic regions that have driven diversification across communities of organisms.

The California Floristic Province (CFP) is a biodiversity hotspot in North America (Myers et al. 2000). Because of conservation concerns, many studies have attempted to elucidate common phylogeographic breaks and shared Pleistocene refugia in codistributed taxa in this region (Lapointe & Rissler 2005; Rissler et al. 2006; Waltari et al. 2007). Generally, these studies have indicated that there is some degree of shared phylogeographic signal (Lapointe & Rissler 2005), where divergence has been identified at the Los Angeles Basin or Transverse Mountain Ranges of California (Calsbeek et al. 2003; Feldman & Spicer 2006; Fig. 1). Because of the number of phylogeographic studies conducted in California, the region has been described as a well-developed geographic study system (Hickerson et al. 2010). However, all comparative studies have been based on single-locus data sets. Therefore, if multilocus analyses differ from previous conclusions, these comparative studies should be reassessed. For example, an early study of the Western Pond Turtle (Actinemys marmorata) using only mtDNA indicated that the Transverse and Coast Ranges were historically important in structuring populations (Spinks & Shaffer 2005), but a more recent analysis using multiple loci identified only the Monterey Bay as a biogeographic barrier (Spinks et al. 2010).

One of the earliest studies to explore population structure in the CFP used the mitochondrial ND4 gene region to show that the inland seaways of southern California caused divergence in the California Mountain

---

Fig. 1 (A) Sampling localities of the individuals included in this study; exact localities and collection numbers are given in Appendix S1 (Supporting information). Circles represent individuals assigned to the northern species (Lampropeltis zonata), triangles indicate individuals belonging to the southern species (Lampropeltis multificiata) and squares represent individuals assigned to the Peninsular Range lineage. The approximate range of L. zonata is highlighted in red, and that of L. multifasciata is highlighted in blue (modified from Stebbins 2003). Abbreviations used are BC, Baja California, CA, California, NV, Nevada, OR, Oregon, WA, Washington. (B) Lampropeltis zonata, photographed near Ashland, southwestern Oregon. Photo: A. St. John.
Kingsnake, Lampropeltis zonata (Rodriguez-Robles et al. 1999). We revisit the phylogeography of L. zonata using multiple loci to address two questions. First, do the lineages identified by Rodriguez-Robles et al. (1999) using mtDNA represent distinct species in the light of additional loci and coalescent analyses? The answer is relevant for understanding aspects of the ecology of this species and for implementing informed conservation strategies. Second, we ask what general region within the CFP has been important for driving diversification.

We first delimit species within L. zonata using a nonparametric heuristic method and then further assessed these lineages using a coalescent-based, Bayesian species delimitation approach. We subsequently estimate divergence times using a relaxed molecular clock, infer patterns of hybridization, examine fluctuations in population sizes through time and evaluate whether delimitable taxa occupy distinct ecological niches. Finally, our re-evaluation of the evolutionary history of L. zonata enables us to revisit conclusions about the phylogeographic history of biological communities within the California Floristic Province.

Materials and methods

Taxon sampling and sequencing

Genetic samples were obtained from 34 individuals from across the distribution of Lampropeltis zonata (Fig. 1; Appendix S1, Supporting information). The DNA extractions used were the same as in the earlier phylogeographic study (Rodriguez-Robles et al. 1999). The mitochondrial locus ND4 and associated tRNAs (tRNA\textsubscript{His}, tRNA\textsubscript{Ser}, tRNA\textsubscript{Leu}) were downloaded from GenBank (Appendix S1, Supporting information). Two anonymous nuclear loci (CL4, 2CL8) were amplified via PCR. Products from the PCRs were cleaned using ExoSap-IT (USB Corporation) and sequenced using 3\% TCA ATC TAG CCC ACT-3\prime and 2CL8R (5\prime-CGC CTA AAA CTA ACA GTA GG-3\prime) for the CL4 locus, and 2CL8 2CL8F (5\prime-CCC TCA ATC TAG CCC ACT-3\prime) and 2CL8R (5\prime-GAT TAG CAG GAA ACT Ct-3\prime) for the 2CL8 gene (Burbrink et al. 2011). All sequences were aligned by eye in Sequencher v4.5 (GeneCodes 2000), as no gaps were detected in any of the three loci in L. zonata. We used the program PHASE v2.1.1 (Stephens & Donnelly 2003) to determine the most probable pair of alleles for each of the two nuclear loci. PHASE was run with default parameters for 100 iterations, a thinning interval of 1 and a burn-in of 100. To check for consistency between runs, we repeated each PHASE run five times. Phased alleles were used in all analyses, unless otherwise noted.

Gene tree analyses

The most appropriate model of nucleotide substitution for each locus was determined using jModeltest (Posada 2008). Phylogeographic structure was independently assessed for each unphased locus using maximum likelihood in RAXML v7.2.8 (Stamatakis 2006), resulting in three gene trees. All trees were rooted with Cemophora coccinea, the sister taxon to the genus Lampropeltis (Rodriguez-Robles & de Jesus-Escobar 1999; Pyron & Burbrink 2009). The model GTRGAMMA was used for each locus partition. In each analysis, 1000 bootstrap replicates were performed to assess support for each node; values greater than 70\% were considered indicative of well-supported clades (Felsenstein 2004). Additionally, gene trees were estimated in a Bayesian framework in MrBayes v3.2.1 (Ronquist & Huelsenbeck 2003). Two independent runs of four Markov chains were conducted for 10 million generations sampled every 100th generation. Stationarity was assessed in Tracer v1.5 (Drummond & Rambaut 2007), and the first two million generations were discarded as burn-in. A majority rule consensus tree for each locus was generated from the post-burn-in posterior probability.

Species delimitation

Throughout this study, we adhere to the unified lineage species concept (de Queiroz 2007), which differentiates species as separately evolving metapopulations. To assess whether cryptic species are present within L. zonata, we used an approach suggested by Niemiller et al. (2012), where a nonparametric method assigns individuals to species while jointly estimating a species tree without constraining individuals to a particular taxon (O’Meara 2010). We used Bayesian species delimitation to validate these lineages by calculating the posterior probability (PP) associated with each node (Yang & Rannala 2010).

We ran the program brownie 2.0 (O’Meara 2010) using the heuristic search with the following settings: number of random starting species trees (nreps) was set to 10 000; minimum number of samples per species (minsamp) was set to two; and taxon reassignment (subsample) was set to one, so that all possible individual reassignments were searched. All other settings were left at the default options. Input trees for this analysis were those that resulted from the RAXML runs of the two nuclear genes, CL4 and 2CL8. Ten different runs were conducted to ensure consistency between analy-
ses. We used TREEANNOTATOR v1.7.2 (Drummond et al. 2012) to infer the 50% majority rule consensus tree of the output tree files from Brownie.

To infer the timing of their divergence, we treated the species estimated from Brownie as terminal taxa in the species tree inference program *BEAST, as implemented in BEAST v1.7.5 (Drummond et al. 2012). The species tree analysis was run with all three loci and with Cemophora coccinea and Lampropeltis knoblochi as outgroups (Appendix S1, Supporting information). We used a Yule model to determine tree shape; population size model was set to piecewise linear with a constant root and each locus was assigned the most appropriate model of nucleotide substitution. This analysis was run for 250 million generations and sampled every 10 000 generations. Stationarity was determined by visually inspecting the trace plots and ensuring that all ESS values were above 200 in TRACER v1.5. The first 25% of sampled genealogies were discarded as burn-in, and the most credible clade was inferred with TreeAnnotator v1.7.2 (Drummond et al. 2012).

In determining the timing of speciation within L. zonata, we followed the suggestions of Parham et al. (2012) in justifying our fossil calibration. We assigned the split of Lampropeltis from Cemophora to a lognormal distribution with a mean of 0.1 and a standard deviation of 0.9, such that the mean divergence time was 12.6 Ma. This lognormal distribution was truncated with a hard lower bound of 10.4 Ma and an upper of 24.0 Ma, giving a 95% prior credible interval of 10.56–20.25 Ma. The mean date corresponds to the oldest known kingsnake, Lampropeltis similis (Holman 1964), and to the vertebral fossils described as the holotype and paratype (University of Nebraska No. 61035 and 61036, respectively) of this extinct taxon collected from the Norden Bridge local fauna of the Valentine formation of Brown County, Nebraska, USA (Holman 1964). Lampropeltis similis has been suggested to be most similar in vertebral structure to Lampropeltis triangulum (Parmley 1994), and the diagnosis of the holotype was revised by Holman (2000) as follows: “similar to L. triangulum (Lacépède 1788), but differs in that (i) the neural arch is less depressed; (ii) the hemal keel in [sic] somewhat thinner; (iii) the centrum is not as triangular from below; and (iv) the neural canal is loaf-of-bread-shaped rather than ovoid and somewhat depressed.” Fission track dating of the Valentine Formation illustrates that this formation is Miocene in age. The stratigraphic level immediately above the formation is 10.6 ± 0.2 Ma old, whereas the lower Valentine Formation is 13.6 ± 1.3 Ma in age (Boellstorff & Skinner 1977; Wellsstead 1981). Fossil data show that during the Miocene, there was a radiation in the snake fauna of North America, during which time boid snakes were replaced by communities composed primarily of species of Colubroidea (Holman 2000). Therefore, the age of L. similis is conservatively constrained as not being younger than 10.4 Ma, and it is unlikely that the split between Lampropeltis and Cemophora occurred prior to 24 Ma, the onset of the Miocene. Additionally, the mean calibration date (12.6 Ma) is based on the mean radiometric date of the Valentine Formation.

To further test the validity of the inferred species, we used Bayesian Phylogenetics and Phylogeography (BPP v2.1; Yang & Rannala 2010). BPP implements a reversible jump Markov chain Monte Carlo to estimate a PP for a hypothesized species. All three loci were included in the BPP analysis using the guide tree generated from *BEAST. This method accommodates the species phylogeny as well as lineage sorting due to ancestral polymorphism. Algorithm 0 was implemented, and fine-tuning parameters were set so that swapping rates for each parameter ranged between the recommended values of 0.30 and 0.70. Following Leaché & Fujita (2010), we implemented three different combinations of priors for ancestral population size (θ) and the root age (τ0). In BPP, both priors are assigned a gamma $G(\alpha, \beta)$ distribution, and thus, we parameterized these priors for: very large ancestral populations and deep divergences, $θ–G(1, 10)$ and $τ_0–G(1, 10)$; small ancestral population size and shallow divergences, $θ–G(2, 2000)$ and $τ_0–G(2, 2000)$; and a more conservative prior combination that accounts for large ancestral population sizes and recent divergences, $θ–G(1, 10)$ and $τ_0–G(2, 2000)$, which may be the most biologically realistic scenario. We ran four independent analyses for each set of priors for 1 000 000 million generations, with a burn-in of 10 000, and a sampling frequency of once every five generations.

A simulation study conducted to evaluate the statistical performance of BPP showed that high support for the correct species model (i.e., avoiding false positives and false negatives) can be attained by sampling two loci with 5–10 sequences from each putative species (Zhang et al. 2011). In all additional analyses, we only used populations that adhered to the suggestions of this simulation study and therefore excluded the Peninsular populations (localities 24, 25; Fig. 1).

Incomplete lineage sorting or introgression?

The lineages identified in the mtDNA and the nuDNA gene trees are not congruent. We thus performed additional tests to determine whether discordance is more likely due to incomplete lineage sorting or hybridization. The first test, the genealogical sorting index (gsi), quantifies the degree of exclusive ancestry of specified
groups based on a gene tree (Cummings et al. 2008). The degree of exclusivity is based on a scale of 0–1, in which 1 specifies monophyly, and nonexclusivity of groups is indicated as the statistic approaches 0. The gsi statistic was calculated for the two inferred species for each locus using the gsi web server (http://www.genealogicalsorting.org ). Input trees for these analyses were the ML gene trees. P values were calculated based on 10,000 permutations.

We used DNAsp v5 (Librado & Rozas 2009) to calculate the number of haplotypes or alleles for each locus. Based on the number of observed haplotypes, we implemented a simulation developed by Rabosky et al. (2009) in R v2.15 (R Development Core Team 2006) to determine whether the observed gene tree heterogeneity represents incomplete lineage sorting or deep coalescence. The simulation calculates the probability that incomplete lineage sorting, not stochastic coalescence, is responsible for the observed discordance between mtDNA and nuDNA loci. We simulated the joint distributions of waiting times to common ancestry for the six species-locus combinations (i.e., three loci from two species each). These waiting times were scaled by relative effective population sizes of the loci (N_e = 1 for nuDNA loci, and N_e = 0.25 for the mtDNA locus). For mtDNA, we simulated 50,000 sets of waiting times and tabulated the number of simulations where the time to the most recent common ancestor of the two inferred lineages exceeded at least j nuDNA species-locus combinations (Rabosky et al. 2009). Results for j = 1 give the probability that mtDNA coalescence time exceeds at least one nuDNA locus in one species, whereas j = 4 is consistent with mtDNA coalescence times exceeding those of both nuDNA loci in the two lineages (Rabosky et al. 2009).

Historical demography

To investigate past changes in effective population sizes in the two delimited taxa, we employed the Extended Bayesian Skyline Plot (EBSP) analysis implemented in BEAST v1.6.2 (Drummond & Rambaut 2007; Heled & Drummond 2008). This test assumes panmixia within each population and therefore we excluded from the analyses any individuals of hybrid origin. For this test, time was scaled by using a substitution rate for the mtDNA locus of 1 x 10^-8 substitutions/site/year as estimated from the *BEAST* analysis. Each locus was assigned the appropriate model of nucleotide substitution and all operator parameters were changed as suggested in the EBSP manual. Each EBSP was run for 250 million generations, with a burn-in of 25 million generations. The effective sample sizes of all parameters were greater than 200 when analysed in TRACER v1.5 (Drummond & Rambaut 2007) indicating stationarity. We determined the most likely number of population size changes in TRACER by examining the frequency distribution of these changes given by the parameter demographic.populationSizeChanges. We calculated the number of polymorphic sites, π (nucleotide diversity), and Tajima’s D for each lineage using DNAsp5 (Librado & Rozas 2009).

Migration

We used the coalescence-based program MIGRATE-N 3.3.2 (Beerli 2006) to test for gene flow. Four models of migration were tested: (i) a full migration model where inferred populations are exchanging migrants; (ii) a panmictic model where individuals were sampled from a single interbreeding population where there is no lineage distinction; (iii) a model in which individuals from the northern lineage migrate into the southern lineage, but southern lineage individuals never migrate; and (iv) a model in which individuals from the southern lineage migrate into the northern lineage, but northern lineage individuals never migrate. The four models account for rates of migration as well as effective population sizes. We assessed model fit using Bayes factors (Beerli & Palczewski 2010).

Ecological niche modelling

Using climatic data and known sampling localities of individuals of *L. zonata*, we constructed ecological niche models (ENMs) for both delimited taxa within *L. zonata*. Georeferenced localities were downloaded from HerpNet (http://herpNet.org/) and assigned to the appropriate lineage based on geographic location. Individuals from southern Kern, southern Santa Cruz and northern Monterey Counties were excluded from this analysis because of uncertainty regarding lineage assignment. Duplicate and erroneous localities (i.e., individuals that were placed well outside the known distribution of *L. zonata*) were removed from the data set, resulting in 215 and 217 localities for the northern and southern lineages, respectively (Appendix S2, Supporting information). To construct ENMs, we used the 19 bioclim variables describing variation in temperature and precipitation at 30-s spatial resolution from the WorldClim data set (Hijmans et al. 2005). ENMs were reconstructed for the northern and southern lineages using MAXENT v3.3.3, with default parameter settings (Phillips et al. 2006), except the number of iterations was increased to 5000. Model performance was evaluated by examining the receiver operating characteristic curve (ROC) and the associated area under the ROC curve (AUC) statistic.
We used ENMTools (Warren et al. 2010) to test two hypotheses: the two lineages occupy identical niches (identity test), or alternatively, there are distinct environmental differences between the two distributions (range-break test; Glor & Warren 2011). The identity test is conducted by asking whether the ENMs generated for any inferred lineages are more different than ENMs calculated from pairs of samples drawn at random from a data set of all pooled samples. Two taxa have identical niches if the observed ENMs are no more different than pairs of randomly drawn samples (Glor & Warren 2011). The range-break test utilizes the same pooled data set, draws a random line through the shared geographic range, calculates ENMs for the sets of localities on either side of this line and estimates both the $I$ statistic and Schoener’s $D$ between these simulated ENMs. If the observed summary statistics fall outside 95% of the values obtained from the simulated distribution, it can be concluded that the observed region is associated with an environmental transition (Glor & Warren 2011). To test these exclusive hypotheses, the observed niche overlap between ENM models was compared to a null distribution of 100 replicates simulated in ENMTools. Statistical significance was determined using one-sample $t$-tests in R and determined using the $I$ statistic (Warren et al. 2010).

**Results**

**Phylogeographic inference and species delimitation**

The number of variable sites was greatest in ND4 (81 sites), followed by 2CL8 (15 sites) and CL4 (seven sites; Table 1). Models of nucleotide substitution for each locus in each lineage are listed in Table 1. Six samples that amplified for ND4 failed to amplify and sequence for one or both of the nuclear loci and were excluded from species delimitation analyses (Appendix S1, Supporting information). The inferred maximum-likelihood mtDNA gene tree was congruent with the previous study of genetic differentiation within *Lampropeltis zonata* (Rodriguez-Robles et al. 1999). However, the exact mtDNA structure is not observed in either of the nuDNA gene trees (Fig. 2). Bayesian estimated gene trees did not differ from those inferred from maximum likelihood and are therefore not discussed further.

Using the nuDNA gene trees as input, Brownie consistently recovered the same three putative species. These lineages are not the same as those inferred from mtDNA, but instead correspond to a northern group that occurs north of Monterey Bay in central California, a southern group distributed from the southern end of Monterey Bay to northern Baja California, Mexico, and a population from the Peninsular Ranges of southern California (Fig. 1). BPP analyses support the three lineages, with high PP under the three possible models of divergence (Table 2). *BEAST* analyses showed that the northern and southern lineages diverged at a median date of 2.07 Ma (95% HPD = 0.79–3.96 Ma), during the late Pliocene to the mid-Pleistocene.

**Incomplete lineage sorting or introgression?**

The $g$si test indicated that the northern and southern lineages are not monophyletic with respect to one another according to the ND4 gene tree (Fig. 2). In contrast, the $g$si statistic for the two delimited taxa resulted in values that showed near monophyly for the two nuDNA gene trees, as all $g$si values were significant at a $P$ value of 0.0001 (Table 3). Based on the number of alleles and the estimated effective populations sizes, the coalescent simulation developed by Rabosky et al. (2009) showed that the probability of obtaining such high values of incomplete lineage sorting for mtDNA ($j = 4$) is 0.00014 (for $j = 1$, $j = 2$ and $j = 3$ the respective probabilities are 0.10852, 0.01138 and 0.00126). These results suggest that introgressive hybridization after the divergence of the two lineages is a much more likely

---

**Table 1** Population genetic statistics for each locus for the northern and southern lineages of *Lampropeltis zonata*

<table>
<thead>
<tr>
<th>Locus</th>
<th>Lineage</th>
<th>Length (bp)</th>
<th>Sample size*</th>
<th>Polymorphic sites</th>
<th>Haplotypes/Alleles</th>
<th>$\pi$ (nucleotide diversity)</th>
<th>Tajima’s $D$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND4</td>
<td>Northern</td>
<td>786</td>
<td>19</td>
<td>57</td>
<td>14</td>
<td>0.02058</td>
<td>−0.26606**</td>
<td>GTR + G</td>
</tr>
<tr>
<td></td>
<td>Southern</td>
<td>786</td>
<td>15</td>
<td>59</td>
<td>14</td>
<td>0.02355</td>
<td>−0.14478**</td>
<td>HKY + I + G</td>
</tr>
<tr>
<td>CL4</td>
<td>Northern</td>
<td>329</td>
<td>16</td>
<td>3</td>
<td>4</td>
<td>0.00118</td>
<td>−1.19782**</td>
<td>F81</td>
</tr>
<tr>
<td></td>
<td>Southern</td>
<td>330</td>
<td>13</td>
<td>5</td>
<td>5</td>
<td>0.00233</td>
<td>−1.23002**</td>
<td>HKY</td>
</tr>
<tr>
<td>2CL8</td>
<td>Northern</td>
<td>314</td>
<td>11</td>
<td>5</td>
<td>5</td>
<td>0.00363</td>
<td>−0.30508**</td>
<td>JC</td>
</tr>
<tr>
<td></td>
<td>Southern</td>
<td>314</td>
<td>15</td>
<td>12</td>
<td>14</td>
<td>0.02025</td>
<td>−0.24623**</td>
<td>K80 + G</td>
</tr>
</tbody>
</table>

Individuals missing $\geq$15% of sequence data for CL4 or 2CL8 were excluded from the analyses.

*Samples removed: CL4, sample 29; 2CL8, samples 8, 10, 12, 17, 32, 33.

**All Tajima’s $D$ values had $P$ values $>$0.10.**

© 2013 John Wiley & Sons Ltd
Historical demography

Results from the parameter demographic.population-SizeChanges indicated that both the northern and southern species experienced one population size change. For both lineages, EBSP showed a general trend of population size increase through time, with accelerated growth during the mid-Late Pleistocene (Fig. 3). Calculated summary statistics for the number of polymorphic sites, nucleotide diversity and Tajima’s $D$ are listed in Table 1. Tajima’s $D$ values for all loci in both lineages are negative, but not significant.

Migration

Hypothesis testing in MIGRATE-N 3.3.2 (Beerli 2006) indicated that the migration model with the highest support is that in which gene flow occurs from the northern into the southern species, with no migration in the opposite direction (Table 4). This model, coupled with the distribution of mtDNA haplotypes (where northern mtDNA haplotypes are shared with the southern taxon), is suggestive of a pattern where females from the northern species have mated with males from the southern species.

Ecological niche modelling

The identity test conducted in ENMTools (Warren et al. 2010) showed that the ENMs are not identical ($P < 0.05$). The range-break test confirms these results, suggesting that the observed climatic divergence between the northern and southern lineages is greater than expected from randomized geographic breaks ($P < 0.05$). This finding indicates that there is an abrupt climatic change

Table 2 Results from Bayesian species delimitation analyses for Lampropeltis zonata assuming a 3-species model

<table>
<thead>
<tr>
<th>Priors</th>
<th>Posterior probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>q−G(1, 10), t0−G(1, 10) (Northern, (Peninsular Range, Southern)'0.96')1.0'</td>
<td></td>
</tr>
<tr>
<td>q−G(2, 2,000), t0−G(2, 2000) (Northern, (Peninsular Range, Southern)'0.98')1.0'</td>
<td></td>
</tr>
<tr>
<td>q−G(1, 10), t0−G(2, 2000) (Northern, (Peninsular Range, Southern)'0.99')1.0'</td>
<td></td>
</tr>
</tbody>
</table>

Speciation probabilities are provided for both nodes in the guide tree. All three lineages are supported with high posterior probabilities under the three sets of priors.

Table 3 Genealogical sorting index ($gs$) for the northern and southern lineages of Lampropeltis zonata

<table>
<thead>
<tr>
<th>Lineage</th>
<th>ND4</th>
<th>CL4</th>
<th>2CL8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td>0.6737*</td>
<td>1*</td>
<td>1*</td>
</tr>
<tr>
<td>Southern</td>
<td>0.5201*</td>
<td>0.8688*</td>
<td>0.8784*</td>
</tr>
</tbody>
</table>

*p = 0.0001,
associated with the area where the geographic ranges of the two lineages of *L. zonata* come into close proximity. We calculated the support for these models using the summary statistic $I$ (observed $I = 0.60$) as well as Schoener’s $D$ (observed $D = 0.32$). The results from these two statistics were indistinguishable.

**Discussion**

**Implications for comparative phylogeography**

Previous comparative phylogeographic studies based on single-locus gene trees hypothesized that communities of organisms in southern California have a shared history of vicariance that was driven by the inundation of the Los Angeles Basin or the orogeny of the Transverse Ranges (Calsbeek et al. 2003; Lapointe & Rissler 2005; Feldman & Spicer 2006). However, our results, based on multilocus coalescent approaches, indicate that the boundary between the two lineages of *Lampropeltis zonata* is further north (c. 525 km) than previously suggested, and that divergence was caused by a younger inland seaway (Dupré et al. 1991). This seaway corresponds to the inundation of the Central Valley by the Monterey Bay, an event that occurred from the late Pliocene to the mid-Pleistocene (Dupré et al. 1991). The phylogeographic history of *L. zonata* is comparable to that of the Western Pond Turtle. A recent study of this turtle (Spinks et al. 2010) shows phylogeographic structure in the mtDNA genome that is not supported by analyses of multiple nuDNA loci. The latter data set instead identifies two populations that meet at the Monterey Bay, with the evidence of hybridization (Spinks et al. 2010), a scenario similar to that of *L. zonata*. Thus, comparative studies in the CFP that examined community-level patterns of divergence may have been misled by relying on the topology of the mtDNA gene tree of *L. zonata* and perhaps other co-distributed taxa (Calsbeek et al. 2003; Lapointe & Rissler 2005; Feldman & Spicer 2006).

Increased genetic sampling indicates that *L. zonata* and *Actinemys marmorata* show a pattern differing from other sympatric taxa regarding the location of biogeographic breaks. This finding may be the result of pseudocongruence, which can occur for two reasons. First, taxa share a common barrier, but the timing of the divergence associated with this barrier is not concordant. Second, an apparent shared biogeographic pattern results from a complex mixture of processes (Soltis et al. 2006). Pseudocongruence possibly has complicated comparative studies in the CFP, which is a geologically and ecologically diverse region where several different geologic processes may be associated with phylogeographic breaks. Another pattern that could be observed in comparative analyses based on a single locus is pseudoincongruence, where seemingly discordant patterns of divergence across taxa do not reflect the true history of the communities being studied, as those communities actually share a common and temporally concordant biogeographic history. Processes that may account for such apparent discordance are differential
rates of expansion out of shared refugia, introgression or stochastic lineage sorting. Furthermore, relying on a single locus for comparative studies across taxa may result in an overinterpretation of the data (Knowles 2009). For example, if population assignments were based on mitochondrial gene trees, comparisons of the biogeographic histories of *L. zonata* and *A. marmorata* would erroneously conclude that there is not a common history. However, multilocus analyses illustrate that these two ecologically different species share a common biogeographic barrier. These results further support recent recommendations for employing multilocus analyses in comparative phylogeography (Fickerson et al. 2010) to reduce pseudocongruence and properly identify the geographic location of biogeographic barriers. We suggest that scenarios proposed for other regions (e.g., the western North American continental divide or eastern North America [Pyron & Burbrink 2010; Soltis et al. 2006]) based on a collection of single-locus studies be re-evaluated.

**Phylogeography and biogeography**

Coalescent species delimitation indicates that *L. zonata* is composed of two distinct species. Simulations testing Brownie has shown that when using empirical gene trees, this programme can potentially lump distinct species into one taxon (Rittmeyer & Austin 2012). Additionally, simulations testing *bivp* show that this method cannot properly support or collapse a node when a hypothetical taxon is sampled at only two individuals and three loci (Zhang et al. 2011). Given these caveats, recognizing two species within *L. zonata* is a conservative conclusion.

The estimated time frame (c. 2.07 Ma) for the separation of the northern and southern species of California Mountain Kingsnake is concordant with the inundation of the Central Valley (Dupré et al. 1991). Although it is possible that the divergence between these two lineages is a result of this inland seaway, results from hypothesis testing with ENMs indicate that the two species occupy distinct ecological niches. This finding suggests that climatic shifts, not necessarily vicariance, may have caused the initial divergence of the two taxa.

It is commonly believed that taxa from temperate regions experienced declines in effective population sizes at times of glacial maxima (Hewitt 2000). However, following their divergence, both the northern and southern species of *L. zonata* experienced population growth during the Pleistocene (Fig. 3). Recent studies suggest that a scenario in which effective population sizes crashed during the LGM may not be as common as previously thought. These studies indicated that populations were either stable through time or experienced growth during the Pleistocene (Feldman & Spicer 2006; Burbrink et al. 2008; Myers et al. 2013).

We found the evidence of admixture between the northern and southern species of *L. zonata* and determined that incomplete lineage sorting at the mitochondrial genome is unlikely. A more probable explanation is that postdivergence gene flow has occurred between the two lineages. Patterns of migration indicate recent gene flow from the northern species into the southern taxon. Introgression between species is relatively common in nature (Rabosky et al. 2009), and unidirectional gene flow can be caused by several reasons, including neutral and stochastic effects, selection or sex-biased processes (Petit & Excoffier 2009; Jezkova et al. 2013). In *L. zonata*, unidirectional migration, where the receiving population only retained the foreign mtDNA genome, indicates that females from the northern species dispersed into populations of the southern species. This pattern does not indicate that there have been frequent or high levels of female dispersal or males do not also disperse, yet the pattern provides the evidence of some level of female migration resulting in introgression. Mark–recapture studies (Glaudas & Rodríguez-Robles 2011) and population genetic analyses (Dubey et al. 2008; Pernetta et al. 2011) of dispersal within and between populations have repeatedly found male-biased dispersal as the norm across snakes, from boids (Rivera et al. 2006) to colubrids (Pernetta et al. 2011). To our knowledge, this is the first time that females have been indicated in dispersing between populations and causing introgression within snakes.

**Taxonomic implications**

Two distinct taxa within *L. zonata* can be delimited with high posterior probability support using coalescent approaches. These two taxa are not only diagnosable using molecular data, but also occupy distinct climatic niches. We therefore suggest recognizing both as separate species. Recognizing two species in this complex is a conservative decision, as the southern taxon could potentially be further subdivided into two separate lineages. Because our delimitations are nonsubjective, alternative hypotheses of species limits can be tested by collecting homologous genetic loci from additional individuals, and by using the data presented here. We elevate the taxon *multifasciata* to species status, as this is the oldest name associated with the holotype collected within the geographic range of the southern lineage. Conservation recommendations should reflect this revised taxonomy.
Species synonymy and distribution
Lampropeltis zonata (Lockington ex Blainville 1835) Belyphi s zonatus Lockington (1876: 53)

Type locality. Northern California. (Holotype: Lost from the California Academy of Sciences).

Lampropeltis zonata, Van Denburgh (1897: 167)
Lampropeltis pyrrhomelaena multicincta, Stejneger (1902: 153).
Lampropeltis multicincta, Blanchard (1920: 5).
Lampropeltis multicincta multifasciata, Klauber (1943: 76).
Lampropeltis zonata zonata, Klauber (1943: 76)

Distribution. Lampropeltis zonata is composed of all populations in the Sierra Nevada Mountains and the Coast Ranges north of Monterey Bay, California, north into the Klamath Mountains, in Oregon, plus an additional, disjunct population along the Columbia Gorge, in the great state of Washington (Fig. 1). These populations cluster as a lineage distinct from all other populations according to coalescent-based analysis of multiple, unlinked genetic loci.

Lampropeltis zonata, Van Denburgh (1897: 167).
Lampropeltis pyrrhomelaena multicincta, Stejneger (1902: 153).
Lampropeltis multicincta, Blanchard (1920: 5).
Lampropeltis multicincta multifasciata, Klauber (1943: 76).
Lampropeltis zonata zonata, Klauber (1943: 76)

Distribution. Lampropeltis multifasciata is composed of all populations in the Peninsular Ranges and in the Transverse Ranges, north into the Coast Ranges just south of Monterey Bay, California, including the disjunct population on Isla Sur of Islas Todos Santos, Baja California, Mexico (Fig. 1). These populations cluster as a lineage distinct from all other populations according to coalescent-based analysis of multiple, unlinked genetic loci.

Conclusion
We conducted a detailed phylogeographic assessment of Lampropeltis zonata (sensu lato) by integrating multilocus data sets, coalescent methods and ENM. Our analyses indicate that the two distinct species diverged at the Monterey Bay during the Pliocene. However, because of an abrupt environmental transition between the distributions of the species, divergence may be the result of a climatic niche shift. We found that gene flow has historically been unidirectional, from L. zonata to Lampropeltis multifasciata. These results are suggestive of female dispersal, an uncommon pattern for snakes. The two species experienced growth in effective population sizes during the last glacial maximum of the Pleistocene. Our findings indicate that researchers should exercise caution when interpreting results of single-locus studies of intraspecific genetic differentiation. By extension, conclusions about community-level phylogeographic histories based exclusively on mtDNA should be revisited using genomic scale data.

Acknowledgements
We thank D. Frost for discussion on the taxonomic history of Lampropeltis zonata, J. Heled for help with xml code for *BEAST*, B. O’Meara for a discussion on implementing Brownie, P. Beerli for correspondence on model testing in MIGRATE-N, D. Rabosky for providing his R script and A. St. John for providing the picture of L. zonata. Analyses were facilitated by a grant of computer time from the CUNY High Performance Computing Center, which is supported by US National Science Foundation Grants CNS-0855217 and CNS-0958379. We also thank the numerous individuals whom helped to collect tissue samples.

References
species delimitation in lampropeltis zonata


Klauber LM (1943) The coral king snakes of the Pacific Coast. Transactions of the San Diego Society of Natural History, 10, 75–82.


Heled J, Drummond A (2008) Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology, 8, 289.


E.A.M. and F.T.B. conceived the study; R.E.S. and D.F.D. collected the samples; E.A.M., J.A.R-R. and S.R. generated the sequence data; E.A.M. performed the phylogenetic and population genetics analyses; E.A.M. and A.S. collected the climatic data and conducted the niche modelling analyses; all authors contributed to writing the paper.

Data accessibility

GenBank accessions are listed in Appendix S1, Supporting information.

Phased, aligned nexus files for each locus are available on DRYAD entry doi:10.5061/dryad.f852.

GPS coordinates used in ENM: Supporting information.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Sampling localities for *Lampropeltis zonata*.

Appendix S2 Georeferenced localities used in ENMs.