

Historical Habitat Barriers Prevent Ring-like Genetic Continuity Throughout the Distribution of Threatened Alameda Striped Racers (*Coluber lateralis euryxanthus*)

JONATHAN Q. RICHMOND^{1,3}, DUSTIN A. WOOD¹, KAREN E. SWAIM², ROBERT N. FISHER¹, AND AMY G. VANDERGAST¹

¹ US Geological Survey, Western Ecological Research Center, 4165 Spruance Road, Suite 200, San Diego, CA 92106, USA

² Swaim Biological Incorporated, 4435 First Street PMB 312, Livermore, CA 94551, USA

ABSTRACT: We used microsatellites and mtDNA sequences to examine the mixed effects of geophysical, habitat, and contemporary urban barriers on the genetics of threatened Alameda Striped Racers (*Coluber lateralis euryxanthus*), a species with close ties to declining coastal scrub and chaparral habitat in the eastern San Francisco Bay area of California. We used cluster assignments to characterize population genetic structuring with respect to land management units and approximate Bayesian analysis to rank the ability of five alternative evolutionary hypotheses to explain the inferred structure. Then, we estimated rates of contemporary and historical migration among the major clusters and measured the fit of different historical migration models to better understand the formation of the current population structure. Our results reveal a ring-like pattern of historical connectivity around the Tri-Valley area of the East Bay (i.e., San Ramon, Amador, and Livermore valleys), with clusters largely corresponding to different management units. We found no evidence of continuous gene flow throughout the ring, however, and that the main gap in continuity is centered across the Livermore Valley. Historical migration models support higher rates of gene flow away from the terminal ends of the ring on the north and south sides of the Valley, compared with rates into those areas from western sites that border the interior San Francisco Bay. We attribute the break in ring-like connectivity to the presence of unsuitable habitat within the Livermore Valley that has been reinforced by 20th century urbanization, and the asymmetry in gene flow rates to spatial constraints on movement and east–west environmental gradients influenced by the proximity of the San Francisco Bay.

Key words: Cluster analysis; Colubridae; Conservation; Migration models; Population genetic structure; San Francisco Bay Area

PATTERNS of contemporary gene flow within a species can be highly variable at local geographic scales, particularly near urban areas where habitat fragmentation and edge effects alter historical movement patterns (Fahrig 2003). Genetic signals arising from these contemporary events inevitably overlay older signals shaped by the deeper species history, and the extent to which this history is retained within individuals is dependent on a variety of factors relating to demography, movement, and life history (Hewitt 2000; Paun et al. 2008; Pruett et al. 2008). The effects of contemporary and historical influences on population genetic structuring and diversity can be difficult to disentangle, yet the task should be integral to management plans intended to maintain population viability in threatened and endangered species. This is because the genetic signals arising from more recent disturbance events are more likely to represent imminent threats to population persistence, whereas legacy signals might provide information about natural historical barriers that may or may not be of consequence in the current landscape.

California has experienced substantial loss and modification of natural lands over the past century, with heavy impacts inflicted on coastal sage scrub, chaparral, and oak savannah grassland (Westman 1981; US Fish and Wildlife Service 2002; Ford and Hayes 2007; Riordan and Rundel 2013). One taxon particularly influenced by efforts to both conserve and develop this habitat is the Alameda Striped Racer (*Coluber lateralis euryxanthus*), a threatened subspecies that is largely restricted to Contra Costa and Alameda Counties in the eastern San Francisco Bay Area of California (hereafter, East Bay). It is one of two currently recognized subspecies of striped racers, the other of which occurs throughout much of the state and extends into northern Baja

California, Mexico (*C. l. lateralis*). Riemer (1954) described *C. l. euryxanthus* on the basis of color and pattern, namely the orange-rufous suffusion on the anterior portion of the body and increased width of the dorsolateral striping, relative to *C. l. lateralis*. Because of the distinctiveness of the snake's phenotype in the East Bay; its small distribution; and close ties to declining scrub, chaparral, oak savannah grassland, and riparian woodland habitat, *C. l. euryxanthus* was listed as threatened under the California Endangered Species Act in 1971, and later under the Federal Endangered Species Act in 1997 (US Fish and Wildlife Service 1997).

Recent challenges surrounding the snake's management involve a limited understanding of how populations are genetically structured across the landscape, and the degree to which *C. l. euryxanthus* is distinctive from neighboring populations of *C. l. lateralis*. The geographic limits of the two forms are nebulous because of a lack of fixed character differences and limited comparative studies on the morphology. Although researchers and managers generally agree on the distribution of each subspecies within the landscape (Riemer 1954; Swaim 1994; US Fish and Wildlife Service 1997; Stebbins 2003), genetic perspectives on the location of the subspecies boundaries and the extent to which the two forms may exchange genes are lacking.

In this study, we used microsatellite and mitochondrial DNA (mtDNA) sequence data to (1) characterize the landscape genetic structuring of *C. lateralis* in the East Bay, (2) infer the history of population expansion across the East Bay landscape, and, (3) build knowledge on historical movement patterns to identify areas where habitat restoration might be most important. Because of management implications and the absence of published genetic data on protected populations, we were interested in using cluster assignments to assess how genetic variation is partitioned

³ CORRESPONDENCE: e-mail, jrichmond@usgs.gov

among different land-management units and the degree to which snakes within those units are genetically admixed. We used approximate Bayesian methods to measure the fit of different historical and recent gene-flow models to better understand the formation and consequences of the current population structure. Finally, we identified the locations of major breaks in gene exchange and associated those breaks with geophysical features or other discontinuities in habitat.

MATERIALS AND METHODS

Field Sampling

Subsequent to the 1997 listing rule, the US Fish and Wildlife Service designated critical habitat areas for *C. l. euryxanthus* totaling approximately 62,659 ha (Fig. 1; US Fish and Wildlife Service 2006). These habitat areas are encompassed within six recovery units. We obtained samples from 12 locations spanning 5 of the 6 recovery units, 9 of which are within the range of *C. l. euryxanthus*, 2 from a possible intergrade area, and 1 from within the putative range of *C. l. lateralis* (Fig. 1; Jennings 1983). For clarity, we refer to habitat areas and recovery units collectively as “management units.”

The bulk of the tissue samples (5.0–10.0-mm tail clips preserved in 95% ethanol) were obtained during field surveys by Swaim Biological Inc. (SBI; Livermore, CA). All tissues from protected populations were obtained under Federal Permit TE-815537. We also obtained tissues from museum specimens, colleagues, and additional field sampling. Identification and locality data for the specimens are provided in Appendix I of the Supplemental Materials.

Genetic Data Collection

We extracted DNA using a Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA). The microsatellite library was developed specifically for *C. lateralis* using Genetic Identification Services (Chatsworth, CA), and methods for library construction, enrichment, and screening followed Jones et al. (2002).

We identified 16 variable loci and genotyped 3–36 individuals from each sampling location; 8 had sample sizes of 27–32 snakes, and 4 had fewer than 5. Repeat motifs for each locus, multiplex primer sets, and primer sequences are provided in Table S1 (Supplemental Materials). We performed 10- μ L multiplex polymerase chain reaction (PCR) reactions using a Qiagen Multiplex PCR kit and 50–100 ng of DNA template, with cycling conditions following the kit protocol. Genotyping runs were performed using an ABI 3100S genetic analyzer (Applied Biosystems, Foster City, CA) and the LIZ internal size standard at the CSUPERB microchemical core facility at San Diego State University. We scored alleles using GeneMarker v1.85 (Softgenetics LLC, State College, PA) and used Genepop on the web (available at <http://genepop.curtin.edu.au/>) to test for linkage disequilibrium across the full sample. We also used Micro-checker v2.2.3 (van Oosterhout et al. 2004) to screen for null alleles and check for scoring errors.

We sequenced a subset of individuals ($n = 3$ –7) from each sampling location for the mitochondrial NADH dehydrogenase 4 (ND4) protein-coding gene and three flanking tRNAs (tRNA-*his*, -*ser*, and -*leu*). We also included two snakes collected from sites just north of the study area in Napa

County. We designed the forward primer specifically for *C. lateralis* (ND4ML-F: GCAACGACTTTCTAAATAACCTAA) and used the previously published tRNA-*leu* as the reverse primer (Arevalo et al. 1994) to amplify 1320 nucleotides (nt) of the ND4 gene and 179 nt of the combined tRNA genes. We purified all PCR products using an UltraClean™ PCR Clean-up Kit (Mo BIO Laboratories, Inc.) and performed Sanger sequencing on an ABI 3130S capillary system at GENEWIZ, Inc. (La Jolla, CA). We edited the raw data using Sequencher v4.6 (Gene Codes Corporation, Ann Arbor, MI) and manually aligned the sequences by eye given that there were no insertions or deletions.

Population Structuring and Snake Movement

Microsatellites.—We used individual-cluster-assignment methods to ask the following three questions about *C. lateralis* in the East Bay: (1) How many clusters exist within the data set? (2) How well do clustering patterns correspond to management and taxonomic units? (3) To what extent are populations in different units admixed? We performed cluster assignments in Structure v2.3 (Pritchard et al. 2000; Falush et al. 2003) based only on the microsatellite data and in TESS v2.3.1 (Chen et al. 2007) using microsatellite and georeferenced data for each individual. We used both approaches to test for consistency in cluster membership and because the latter was specifically developed to handle data sets with strong signals of genetic isolation by distance (IBD; François and Durand 2010).

TESS and Structure implement similar Markov Chain Monte Carlo (MCMC) methods to estimate assignment coefficients, but rely on different priors for the assignment matrix. Differences between the two methods are detailed in François and Durand (2010) and the methodology for Structure has been discussed elsewhere (Pritchard et al. 2000, 2009; Falush et al. 2003; Hubisz et al. 2009; François and Durand 2010). Briefly for TESS, assignments are estimated by incorporating sample locations, spatial trends, and spatial autocorrelation in the prior distribution of the assignment coefficient matrix (Durand et al. 2009). The methods for incorporating spatial trend and autocorrelation priors in the assignment matrix are described in a separate file available with the Supplemental Materials.

Preliminary analysis of our data showed strong IBD, which can be problematic for approximating the number of clusters K (François et al. 2006; Durand et al. 2009; Pritchard et al. 2009; see the Supplemental Materials for details on the IBD analysis). TESS was developed to account for IBD, however, and provides a framework for comparing models that incorporate spatial covariates when estimating cluster membership. Model fit is inferred using the Deviance Information Criterion (DIC), which is computed during MCMC runs as the average model deviance plus a penalty term, p_D (Spiegelhalter et al. 2002). The p_D term approximates the gain in fit by increasing model complexity, with the lowest DIC score indicating the model that would best predict a replicate data set with the same structure as the observed data.

To infer cluster membership, we used admixture models with correlated allele frequencies across all sampling locations. We ran MCMC analyses for $K = 2$ –10 and calculated the mean log probability of the data [$\ln P(D|K)$]

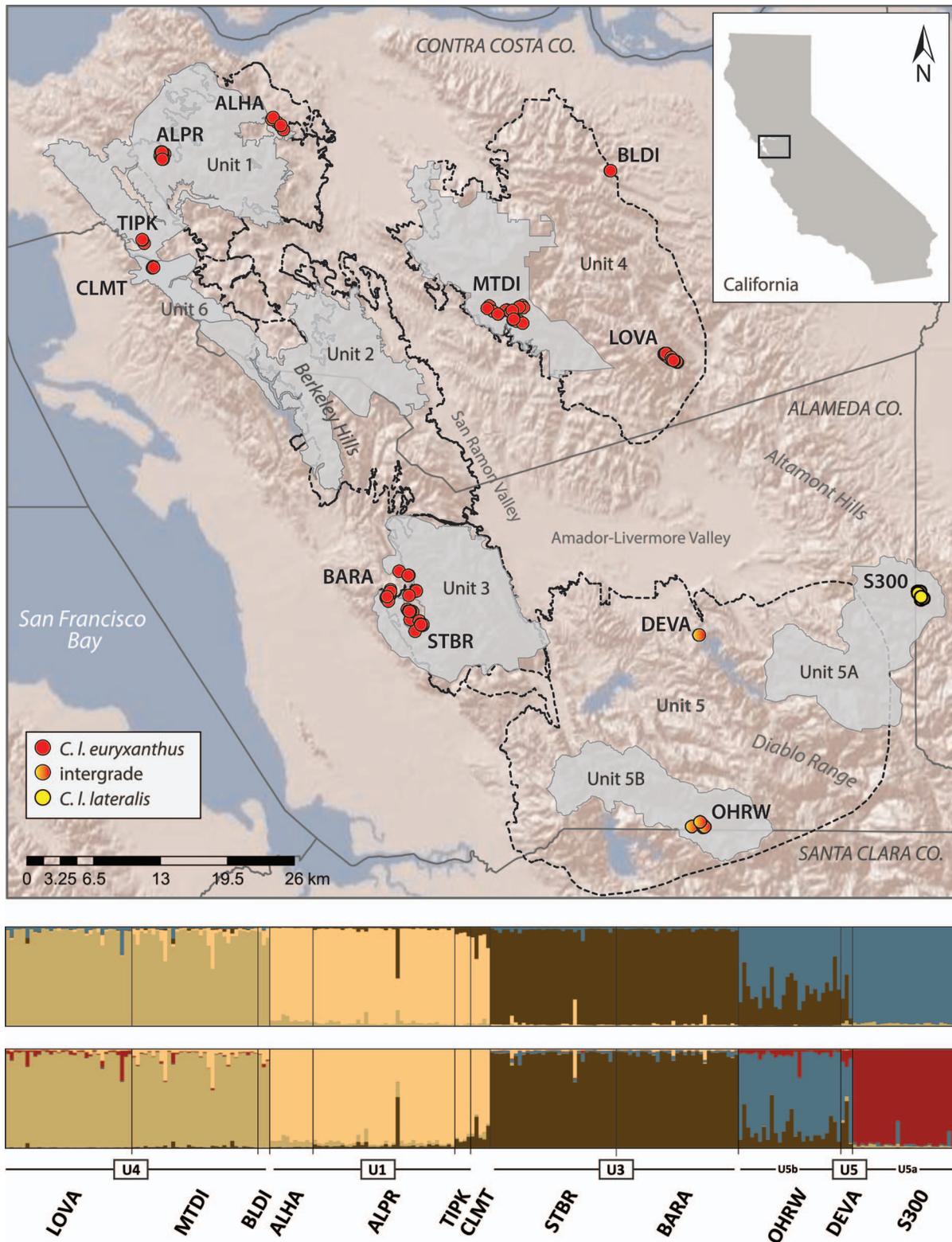


FIG. 1.—Sampling distribution (red or yellow circles), critical habitat units (light grey), and land management units (black dashed outlines) for *Coluber lateralis* in the East Bay of California (inset). Orange circles identify *C. l. euryxanthus*; yellow circles identify *C. l. lateralis*; intermediate color denotes the putative intergrade zone. Site codes and number of individuals genotyped for microsatellites/sequenced for mtDNA are as follows: LOVA = Los Vaqueros Watershed (32/5); MTDI = Mount Diablo State Park (32/4); BLDI = Black Diamond Mines Regional Park (3/3); ALHA = Alhambra (11/5); ALPR = Allen Property (36/6); TIPK = Tilden Regional Park (4/4); CLMT = Claremont Canyon Regional Preserve (5/3); STBR = Stonebrae (32/5); BARA = Bailey Ranch (31/3); OHRW = Ohlone Regional Wilderness (26/4); DEVA = Del Valle Regional Park (3/3); S300 = Site 300 (27/9). Cluster assignment plots from TESS are shown below the map for $K = 4$ (above) and $K = 5$ (below); horizontal lines above the site identifiers indicate the corresponding management unit (labeled “U#” in the boxes). A color version of this figure is available on-line.

TABLE 1.—Diversity estimates for clusters inferred in TESS based on the microsatellite genotypes ($K = 5$) of *Coluber lateralis euryxanthus*. Notations are as follows: n = sample sizes for microsatellites/mtDNA; A = mean number of observed microsatellite alleles; A_r = allelic richness, adjusted for sample size using rarefaction; A_{pr} = private allelic richness; H_o = observed frequency of heterozygotes, ranging from 0 (all individuals are homozygous) to 1 (all individuals are heterozygous); H_s = expected frequency of heterozygotes (i.e., gene diversity; Nei 1987); G_{is} = inbreeding coefficient (analog of Wright's F_{is} [1951] that relates H_o to H_s , ranging from -1 to 1); h = number of mtDNA haplotypes; h_d = mtDNA haplotype diversity; and k = average number of nucleotide differences among haplotypes.

Unit/Cluster	n	A	A_r	A_{pr}	H_o	H_s	G_{is}	h	h_d	k
Unit 1	56/20	5.17	4.66	0.28	0.50	0.55	0.08	2	0.48	0.48
Unit 3	63/8	5.67	4.96	0.49	0.44	0.47	0.07	2	0.25	0.25
Unit 4	67/12	5.67	5.15	0.31	0.52	0.51	-0.02	4	0.46	0.63
Unit 5B + Del Valle	29/7	6.08	6.35	0.94	0.48	0.52	0.07	3	0.52	1.24
Unit 5A	27/9	3.08	3.27	0.18	0.32	0.32	-0.01	3	0.61	1.18

and DIC scores (in Structure and TESS, respectively) for separate runs at each K . We plotted these scores against K and used the inflection point of the curve to approximate the number of clusters (Durand et al. 2009; Pritchard et al. 2009), but we interpreted results across a range of K because of the presence of hierarchical structure in the data and uncertainty in how the parameter is estimated (Meirmans 2015). The variance across runs tended to be greater for the DIC compared with $\ln P(D|K)$, so we ran greater numbers of shorter runs at each K in TESS and selected the 10 best scoring runs to calculate the mean DIC ($n = 25$ runs/ K , 100,000 steps; burn-in 50,000). In Structure, we performed 12 runs of 250,000 steps at each K (burn-in 100,000) and retained the 10 lowest scoring runs. To summarize the data from replicate runs, we used CLUMPP v1.1.2 to align the membership coefficient matrices at each K using the Greedy algorithm (Jakobsson and Rosenberg 2007) and Distruct v1.1 to generate assignment plots of the aligned matrices (Rosenberg 2004). We generated input files for CLUMPP using STRUCTURE HARVESTER (Earl and vonHoldt 2012).

Based on cluster assignments, we used an analysis of molecular variance (AMOVA; Excoffier et al. 1992) to examine the proportion of genetic variation explained by different hierarchical groupings of the microsatellite data. We assessed variation at several levels: (1) individuals nested within sampling locations; (2) individuals nested within sampling locations, nested within clusters (where clusters essentially corresponded to management units); and (3) individuals nested within sampling locations, nested within subspecies. For (3), we conducted analyses based on two groupings of the data to account for the ambiguity in the subspecies boundary—in the first treatment, we treated all Unit 5 samples as *C. l. lateralis* and the remaining samples as *C. l. euryxanthus*; in the second, we treated only Unit 5a samples (i.e., S300) as *C. l. lateralis*. F -statistics for all AMOVA corresponded to Weir and Cockerham's Θ (1984).

Mitochondrial DNA.—Because of the small geographic scale of this study and the limited variation in mtDNA haplotypes, we estimated a haplotype network in TCS v1.21 (Clement et al. 2000). TCS estimates the maximum number of steps among haplotypes as a result of single substitutions with 95% statistical confidence and reconstructs the network using parsimony (Templeton et al. 1992). We also grouped clades hierarchically within the network according to criteria outlined in Templeton and Sing (1993), and used GeoDis v2.6 to make inferences about demographic history (Templeton et al. 1995; Posada et al. 2000). We acknowledge the

criticisms of this method (e.g., Knowles 2008; Petit 2008), and use it here only to complement our other analyses.

Estimates of Genetic Diversity

For microsatellites, we measured standard indices of genetic diversity in GenoDive v1.2 (Meirmans and Van Tienderen 2004) and analyzed samples according to cluster identity to account for variable sample sizes at different locations (Table 1). We estimated allelic richness A_r for each management unit using rarefaction in HP-Rare v1.0 (Kalinowski 2005), and conducted a permutation test in FSTAT v2.9.3 (Goudet 2001) to compare differences in allelic richness A_r , observed heterozygosity H_o , expected frequency of heterozygotes H_s , and pairwise population differentiation F_{st} ($n = 5000$ pseudoreplicates).

Inferring Population History

To better understand the population history of *C. lateralis* in the East Bay, we used DIYABC v2.0.4 (Cornuet et al. 2014) to simulate data sets representing five different evolutionary scenarios for the inferred clusters and then compared the fit of those scenarios with the observed data using approximate Bayesian computation (ABC; Beaumont et al. 2002). Our specific questions were (1) What are the posterior probabilities for the different scenarios given equal prior probabilities? (2) How confident can we be in our estimate of the best-fit scenario? The five scenarios differed mainly in the genealogical relationships of the different clusters, with one attempting to model admixture among sites in the southern portion of the study area (Scenario 2) and others mimicking different patterns of population expansion (Fig. 2). Further descriptions on the different scenarios tested are provided in a separate file in the Supplemental Materials.

We first simulated 5×10^6 data sets based on models describing all scenarios and used those data sets to generate a reference table. Each row of the reference table consisted of the parameter values used to simulate a particular data set and the estimates of different summary statistics computed on that same data set (see Supplemental Materials). To ensure that at least some of the scenarios could produce data sets that were consistent with the real data, we first conducted a principle component analysis (PCA) in the space of summary statistics using data sets simulated from the prior probability distributions for all parameters, and verified that the observed data fell within the cloud of simulated data sets on each PCA axis.

We used the direct and multinomial logistic regression approaches to identify which scenario produced simulated

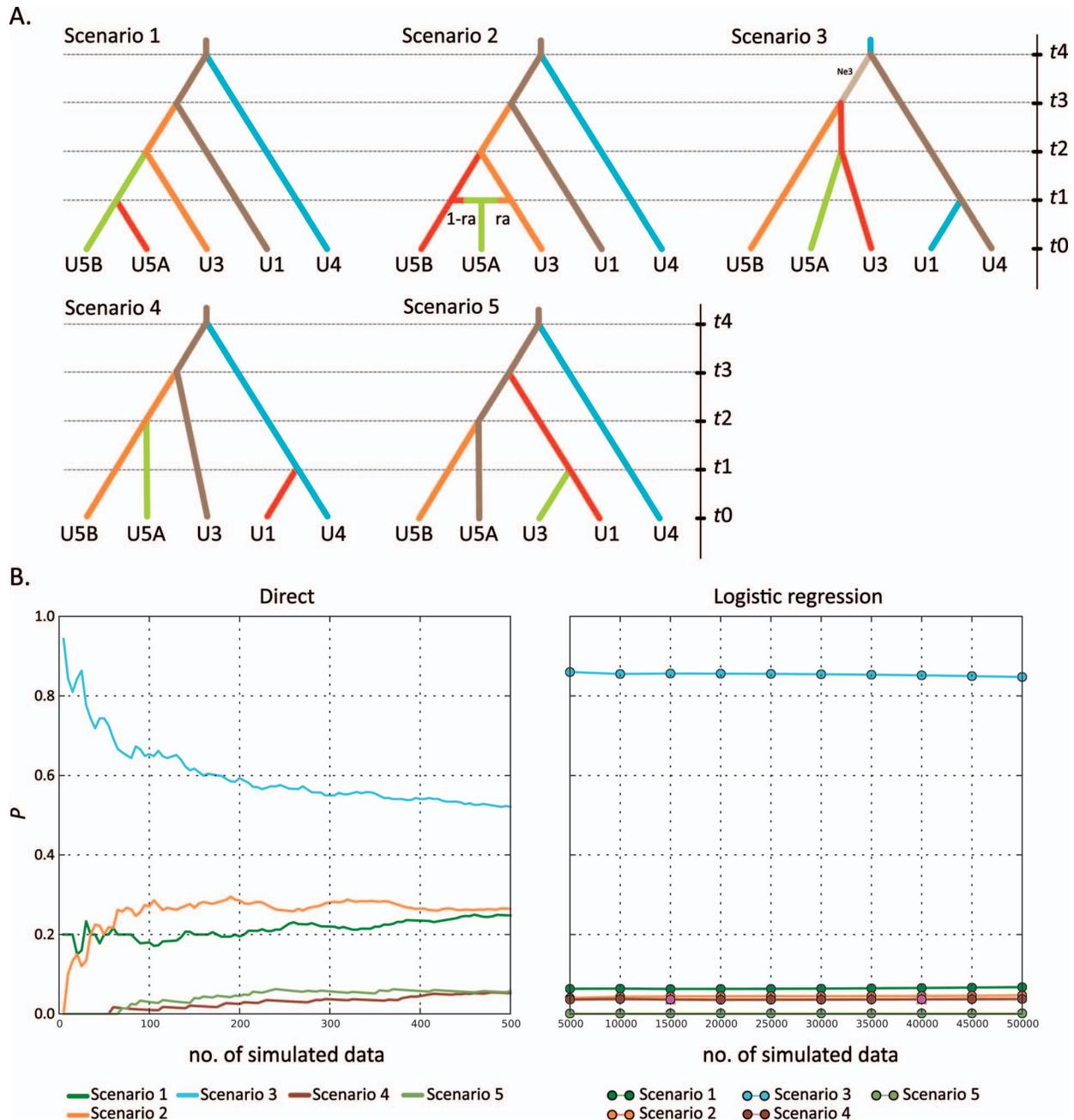


FIG. 2.—(A) Competing evolutionary scenarios used to infer patterns of population expansion in *Coluber lateralis euryxanthus*. In Scenario 2, we treated Unit 5b as a composite population with admixture rate ra from two of the geographically adjacent populations (i.e., Units 3 and 5a). Prior distributions for effective size N_e , time to coalescence t , and the admixture rate ra are described in a file available with the Supplemental Materials. Time is not drawn to scale. (B) Posterior probabilities for the different scenarios as measured by the proportion of data sets P representing a particular scenario in the n closest data sets (i.e., direct approach) and by performing a multinomial logistic regression (i.e., logistic approach). A color version of this figure is available on-line.

data sets that were closest to the observed data (reviewed in Cornuet et al. 2008). Then, we computed Type I and Type II error rates in selecting the best-fit scenario by first simulating 500 data sets and estimating parameter values for the scenario with the highest posterior probability of the 5 tested. We counted the proportion of times that the best-fit scenario did not have the highest probability over competing scenarios in these 500 data sets—this equaled the Type I error (i.e., the rate at which the best-fit scenario was not selected even though it was the “true” scenario). To compute Type II error, we simulated another 500 data sets

based on a competing scenario, and we instead counted the proportion of times the best-fit scenario had the highest posterior probability even though it was not the true scenario.

As a final assessment of the goodness-of-fit between our model parameter posterior combination and the observed data, we conducted a posterior predictive check using PCA. This analysis mirrors that described above for the model pre-evaluation step but adds a superimposed plot of data sets from the posterior predictive distribution. The model is considered an appropriate fit if the observed data are nested

TABLE 2.—Results of the AMOVA based on hierarchical groupings described in the text. For Grouping 3, which tested for differences within and between subspecies, only individuals from S300 were coded as *Coluber lateralis lateralis*—all remaining samples were coded as *C. l. eurixanthus* (similar results obtained when all Unit 5 samples were coded as *C. l. eurixanthus*). *P*-values for groupings that involve clusters are not reported because the same data were used to infer the clusters.

Source of variation	% var.	<i>F</i> -stat	<i>F</i> -value (95% CI)	<i>P</i> -value
Grouping 1				
Within individuals	79	<i>F</i> _{it}	0.21 (0.13–0.29)	—
Among individuals	1	<i>F</i> _{is}	0.01 (–0.02–0.05)	0.152
Among sample sites	20	<i>F</i> _{st}	0.20 (0.13–0.28)	0.001
Grouping 2				
Within individuals	78	<i>F</i> _{it}	0.22 (0.14–0.31)	—
Among individuals	1	<i>F</i> _{is}	0.01 (–0.02–0.05)	0.173
Among sample sites	6	<i>F</i> _{sc}	0.07 (0.05–0.09)	na
Among clusters	15	<i>F</i> _{ct}	0.15 (0.07–0.24)	na
Grouping 3				
Within individuals	72	<i>F</i> _{it}	0.28 (0.18–0.38)	—
Among individuals	15	<i>F</i> _{is}	0.17 (0.10–0.24)	0.001
Between subspecies	13	<i>F</i> _{st}	0.13 (0.07–0.20)	0.001

within the space of summary statistics inferred from both the priors and the posterior predictive distribution.

To complement results from the ABC analyses, we used Migrate-n (Beerli and Palczewski 2010) and BayesAss v3.0 (Wilson and Rannala 2003) to compare historical and contemporary migration patterns among the four land management units (Unit 1, 3, 4, and 5). Migrate-n estimates the long-term average rate of gene flow using a coalescent framework and MCMC sampling to integrate migration estimates, *m*, between the date of the sample and the most recent common ancestor (Beerli 2009). In contrast, BayesAss uses assignment tests to identify individuals with migrant ancestry from the past two generations, and then uses the frequency of those individuals to estimate *m* in a Bayesian framework with MCMC sampling (Wilson and Rannala 2003).

We developed three migration models to test different hypotheses about historical population connectivity in Migrate-n: (1) full stepping-stone model assuming connectivity in a continuous geographic ring and allowing for asymmetric gene flow rates between units; (2) incomplete stepping-stone with symmetric rates between units, but no migration between Units 4 and 5 (i.e., across the Livermore Valley); and (3) incomplete stepping-stone with asymmetric rates and no migration between Units 4 and 5. We implemented a Brownian motion mutation model and uniform priors for population size (Θ : Min = 0, Max = 10, Delta = 1.0) and migration rates (M: Min = 0.0, Max = 1000, Delta = 100) to estimate the parameter values of the best-fit model. Mutation rates were allowed to vary among each microsatellite locus. We conducted 10 replicates for each model using 4 chains (static heating set at 1.0, 1.5, 3.0, 10⁶), discarded the first 2 × 10⁵ genealogies as burn-in followed by 2 × 10⁵ steps recorded every 100 increments, resulting in 2 × 10⁸ sampled genealogies for each model. We assessed convergence by examining ESS values (>1000) and comparing the consistency of the parameter estimates from two separate runs with different initial seeds. We calculated Bayes factors from the marginal likelihoods to estimate the probabilities of the different gene flow models (Beerli and Palczewski 2010).

To assess the rate and directionality of contemporary migration among the different management units, we performed five separate analyses in BayesAss with different starting seeds to compare consistency in the estimates among runs. Each run consisted of 1 × 10⁸ iterations, with samples drawn from the posterior every 1000th iteration (burn-in = 2 × 10⁷). We adjusted the mixing parameters for allele frequencies, inbreeding coefficients, and migration rates to ensure that 40–60% of the total changes were accepted. We used Tracer v1.6 (Rambaut et al. 2013) to assess convergence of the posteriors parameter values.

RESULTS

Genetic Diversity

We genotyped 259 snakes and recovered 136 alleles across all microsatellite loci (\bar{X} = 8.50 alleles/locus). Estimates of genetic diversity for the different clusters identified in the assignment tests (*K* = 5) are provided in Table 1. Half of the total number of private alleles (*n* = 8/16) were detected in Unit 5, with 7 occurring in the Ohlone Regional Wilderness samples. We found no differences in *A_r*, *H_o*, *H_s*, or *F_{st}* among management units based on a nonparametric permutation test, and estimates of *G_{is}* revealed no indication that members of the different clusters are inbred.

We found no evidence of linkage disequilibrium among the microsatellite loci. However, Micro-checker detected three loci that had greater than expected numbers of homozygotes for certain allele size classes: D11 showed heterozygote deficiencies in 9 of 12 populations, whereas B107 and A101 showed deficiencies in 3 and 4 populations, respectively. Based on exact tests (10,000 randomizations), all three deviated from Hardy–Weinberg expectations in three or more populations; thus, we excluded them from our analyses. A fourth locus was monomorphic in all East Bay *C. lateralis* populations, but was polymorphic in *C. lateralis* from southern California (Rancho Jamul, San Diego Co.), so we eliminated it from the data set as well.

For mtDNA, we recovered 13 haplotypes in 69 individuals. The number of variable sites, parsimony informative sites, and average number of site differences was 16, 7, and 1.49 respectively, and the number of different haplotypes recovered within management units ranged from 7 in Unit 5 (5A and 5B combined) to 2 in Units 1 and 3 (Table 1). Two haplotypes, Hap-1 and 3, were recovered at higher frequency than all others and were most common in the northern part of the study area (Contra Costa Co.). We found slightly greater haplotype diversity (*h_d*) and average number of nucleotide differences (*k*) among haplotypes in the eastern section of Unit 5, which contrasted somewhat with our findings for less diversity in the microsatellites.

Analysis of Molecular Variance

For the AMOVA nested individuals within sampling locations, most of the variance (79%) was explained by within-individual and among-sampling site variation (Grouping 1; Table 2). After including cluster identity as an additional nesting group, the majority of the variation was again explained by differences among snakes within sampling locations (78%), followed by differences among the inferred clusters (15%; Grouping 2). Results of a third set of

AMOVAs (coded based on subspecies identity) were largely consistent regardless of how we classified snakes from Unit 5—among-individual variation within subspecies was equivalent to, or greater than, between subspecies (~15% of the variance explained by differences among individuals within subspecies, and ~13% explained by differences between subspecies). Although among-individual variation within and between subspecies was roughly equivalent, the differentiation between subspecies was still apparent ($F_{st} = 0.13$, 95% CI = 0.07–0.20; $P = 0.001$).

Population Structuring

Microsatellites.—In the cluster analyses, the DIC and $\ln P(D|K)$ curves had slightly different inflection points for the admixture models (Fig. S2). For Structure, the inflection of the $\ln P(D|K)$ curve was closest to $K = 5/6$, whereas in TESS the DIC curves plateaued at $K = 4/5$. DIC curves generated from analyses with and without spatial interactions (i.e., $\psi = 0.0$ vs. 0.6) were overlapping, indicating that local spatial effects did not improve model fit. There was a clear jump in the DIC between models fitting a constant trend surface vs. those fitting a linear surface (trend = 0 versus 1), with a linear surface substantially improving model fit (Fig. S2).

We found little qualitative difference in the assignments inferred using spatial priors (TESS) vs. no spatial priors (Structure), indicating that the results were driven mainly by the genetic data and not the spatial priors. Admixture plots generated from TESS and Structure analyses revealed no meaningful population structuring at $K > 5$. Because of the clear signal of IBD in the data (Fig. S3) and the ability of TESS to account for IBD in the analysis, we limit our discussion to the results generated in TESS for $K = 2$ –5 and present assignment plots for $K = 4/5$ (Fig. 1). Plots for $K = 2/3$ are shown in the Supplemental Materials (Fig. S4).

Cluster identities largely corresponded to the different management units, which form an elliptical distribution around the Tri-Valley area in the East Bay (e.g., around San Ramon, Amador, and Livermore valleys; Fig. 1). We found no evidence of overlap in the terminal portions of the ring, which were centered on the north and south sides of the Livermore Valley; populations from Unit 4 to the north of the valley always clustered separately from those to the south in Unit 5. In fact, the main division at $K = 2$ largely adhered to the east–west trajectory of the Amador–Livermore Valley. Cluster dissociation around the ring followed a nearest-neighbor pattern as we fit the data to successively higher K , beginning with the separation of Units 1 and 4 ($K = 3$) followed by the separation of Units 3 and 5 ($K = 4$). The three sites within Unit 4—Black Diamond Mines, Mt. Diablo, and Los Vaqueros—remained clustered across the full range of K values examined.

In general, there was little admixture among snakes from the different management units, with the exception of two areas. One was in the western section of Unit 5, where individuals from Del Valle and Ohlone Regional Wilderness (U-5B) were substantially admixed with samples collected on either side of the two parks in Units 3 and 5A (Fig. 1). The second area was in Tilden Regional Park and Claremont Canyon, where individuals shared a greater degree of co-ancestry with Unit 3 compared with other snakes from Unit

1. Both admixture patterns were consistent with sampling geography and IBD.

Mitochondrial DNA.—MtDNA haplotypes connected in a single most parsimonious network (17 steps, 95% connection limit; Fig. 3). A common haplotype (Hap-3) with high frequency in the northern part of the study area formed a transitional node connecting all East Bay haplotypes to others immediately north of the Bay Area in Napa County (personal observations). Only a single substitution distinguished Hap-3 from five other haplotypes in the study area, with three of the five concentrated in roughly the same area as Hap-3 and two (Hap-6 and Hap-7) extending substantially further south in the Diablo Range and well outside of the putative range of *C. l. euryxanthus*. Demographic inferences based on Templeton et al. (1995) tentatively indicate a north-to-south range expansion across the study area and isolation by distance (Table S4 in Supplemental Materials).

Historical Demographic Inference

PCA based on the 6×10^6 data sets simulated from the prior distributions of parameter values showed that the different historical demographic models could produce data sets that were consistent with the observed data (Fig. S5). We found strong statistical support for Scenario 3 as the preferred historical model (Fig. 2). This model is consistent with a ring-like pattern of north-to-south population expansion, but without closure in the terminal parts of the ring between Units 4 and 5. Using the logistic regression approach, the posterior probabilities of the different historical simulations showed that Scenario 5 (in which Units 4 and 5A on opposite sides of the Livermore Valley were treated as sister lineages) consistently had the lowest posterior probability of the different scenarios tested, and that the probability of Scenario 5 being the true scenario was near 0.0 (Fig. 2).

The Type I and Type II error rates in the choice of Scenario 3 as the preferred model were 0.10/0.07 (direct/logistic) and 0.13/0.03, respectively, indicating high confidence in Scenario 3 as the best fit among the different models tested. The posterior predictive step confirmed that the fit between the preferred model and the observed data was appropriate (Fig. S5).

Tests for recent immigration using BayesAss showed that estimated proportion of individuals assigned to the management unit in which they were sampled was consistently close to 1.0 for all units (when accounting for error), indicating that populations within these units have been genetically isolated for at least several generations (Table S5). When measuring historical gene flow in Migrate-n, goodness-of-fit tests strongly favored a stepping-stone model without ring closure across the Livermore Valley (i.e., no migration between Units 4 and 5) over one that allowed continuous gene flow in a complete ring (Table 3). Furthermore, the incomplete-ring model with different rates of multidirectional gene flow provided a better fit to the data than the alternative with fixed rates. Parameter estimates from this model indicate that gene flow has occurred at a greater rate out of Units 4 and 5 (i.e., the terminal ends of the ring) than into them; whereas, rates between units within the Berkeley Hills were similar, regardless of directionality. These results corroborate those of the assignment tests and the ABC

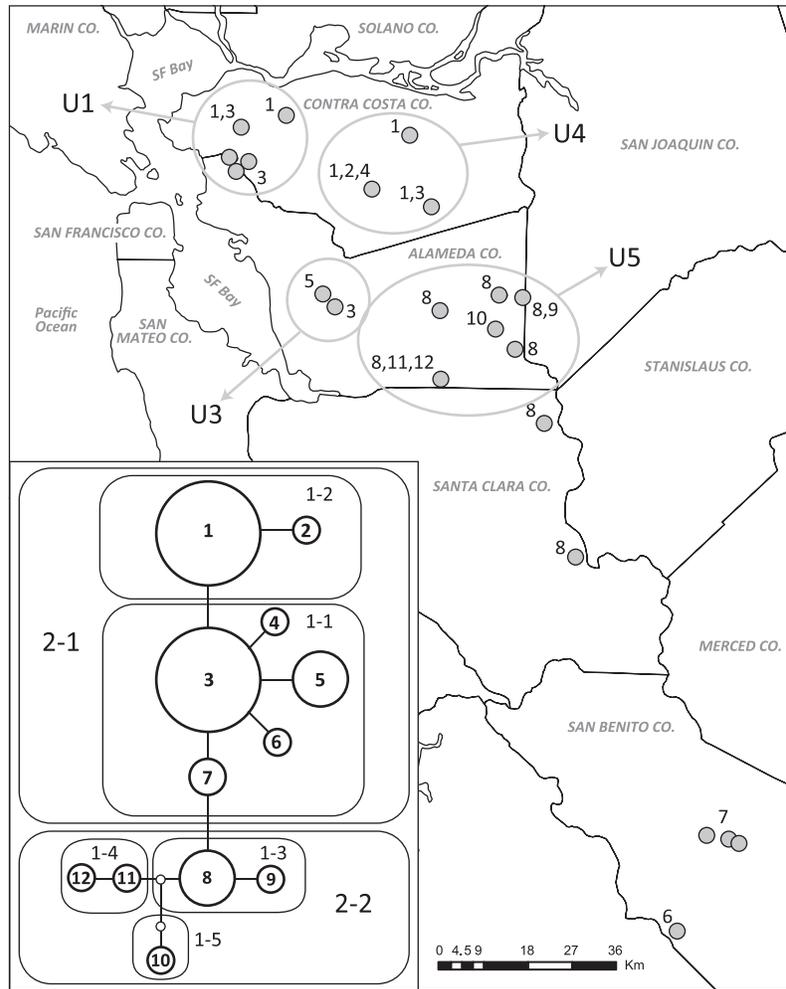


FIG. 3.—Distribution of mitochondrial DNA haplotypes for *Coluber lateralis euryxanthus* across its distribution in the East Bay region of California (U# identifies the management unit). The inset shows the fully resolved haplotype network with a 95% confidence limit. Hap-3 is the inferred ancestral haplotype.

analysis in indicating that gene flow barriers have inhibited ring-like connectivity across the Livermore Valley.

DISCUSSION

Population structuring within East Bay *C. lateralis* reveals a ring-shaped pattern encompassing the Tri-Valley area, but without terminal overlap in the tips of the ring across or around the east end of the Livermore Valley (Fig. 1). Although connectivity in this assemblage has likely waned over the past two centuries on account of the loss of contiguous habitat, our findings suggest that the lack of continuity across the Livermore Valley is probably caused by the longstanding absence of suitable habitat on the valley

floor, which in turn has been reinforced by urbanization. Furthermore, estimated rates of historical migration reveal greater movement away from the ring termini in Units 4 and 5, where habitat and climate might be less favorable to this species compared with more interior locations adjacent to the San Francisco Bay.

Patterns of Population Structuring

Spatial variation in microsatellite and mtDNA markers provide evidence that the distribution of *C. lateralis* in the East Bay initially developed via a north-to-south expansion along two geographic axes; one extends from northern Contra Costa County around the east side of the San Ramon

TABLE 3.—Results from model comparisons (listed from best supported to least supported) generated by Migrate-n for estimated rates of gene flow in *Coluber lateralis euryxanthus*. Incomplete stepping-stone models did not allow migration across the terminal portions of the ring between Units 4 and 5 (i.e., across the Livermore Valley; see Fig. 1). Values reported are the number of parameters estimated, the Bézier approximation scores of log marginal likelihoods (lmL), log Bayes Factor (LBF), and the posterior model probability.

Model description	Parameters	Bézier lmL	LBF	Probability
Incomplete stepping-stone with asymmetric rates	10 (4 Θ, 6 M)	-50,184.82	0.00	1.00
Full stepping-stone (continuous ring connectivity)	12 (4 Θ, 8 M)	-87,303.08	-37,118.26	0.00
Incomplete stepping-stone with symmetric rates	7 (4 Θ, 3 M)	-306,687.28	-256,502.46	0.00

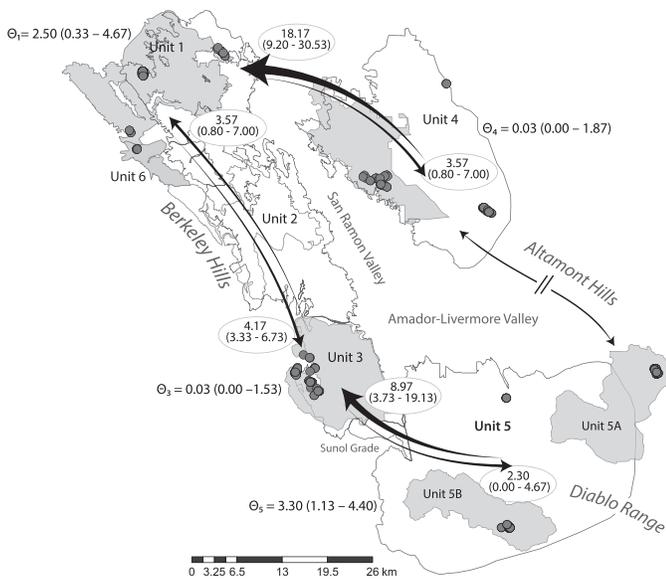


FIG. 4.—Migration rates and theta estimates (mean and 95% credible interval) for *Coluber lateralis euryxanthus* from a Migrate-n output of the incomplete stepping-stone model with asymmetric rates (see Table 3), superimposed on the outlines of different management units. Rates for each direction are shown next to the arrow heads (thicker lines and arrows indicating higher rates); theta subscripts indicate management unit; dots indicate the sampling locations, which were grouped by management unit for these analyses. Major geographic features discussed in the text are labeled. Broken arrows across the Amador–Livermore Valley and Altamont Hills reflect the model’s assumption of no migration through this area.

Valley and south into the Los Vaqueros Watershed (east axis; Fig. 1). The other extends around the west side of San Ramon Valley into the Berkeley Hills, and continues south before turning eastward to the Sunol–Cedar Mountain area of the Diablo Range (west axis). Consistent with a north-to-south expansion, the inferred ancestral mtDNA haplotype (Hap-3) occurs at a greater frequency in the northern study area compared with the south, with several one-step derivatives recovered from the same general location. Because ancestral haplotypes and their mutational derivatives are expected to occur at greater frequency near the site of origin (Templeton et al. 1995), and because Hap-3 links all East Bay haplotypes to others north of the Bay Area (personal observations), we view these patterns as evidence that *C. lateralis* has expanded southward after crossing into the East Bay from the northern Coast Ranges.

Microsatellite data support the same inference of an initial ring-like expansion and semicircular connectivity based on five additional findings: (1) no clustering ever occurred between management units or sampling sites on opposite sides of the ring (Fig. 1); (2) the IBD pattern was best described by a linear trend surface, as expected if populations were oriented in a stepping-stone pattern around a geographic ring; (3) comparing the fit of different historical models using ABC methods showed the strongest support for a history in which population differentiation mirrors the patterns predicted for a ring-like, north-to-south expansion; (4) the model in which *C. lateralis* populations in Units 4 and 5 were treated as sister lineages (i.e., Scenario 5; Fig. 2) provided a poor explanation for the observed data, as expected for populations at different ends of an incomplete ring; and (5) the best-fit model in Migrate-n assumed serial

connectivity among adjacent clusters, but without gene flow across the eastern end of the Livermore Valley.

By combining samples based on cluster identity, we inferred that the rate of historical gene flow out of Unit 5 and into Unit 3 in the southern Berkeley Hills was about 4× greater than in the reverse direction (Fig. 4). A similar trend held for the northern portion of the ring, where east-to-west movement away from sites in Unit 4 was ~5× greater than the rate into those sites (i.e., west-to-east). In contrast, the 95% credible intervals for bidirectional rates between Units 1 and 3 were broadly overlapping, indicating no bias in historical movement among populations within the Berkeley Hills. Thus, the general pattern indicates a history of greater movement away from the terminal ends of the ring in the eastern part of the study area toward localities that are either in, or closer to, the Berkeley Hills in the west. Below, we discuss different potential factors underlying these patterns, all of which point to the low-lying San Ramon and Amador–Livermore valleys as important drivers in the development of *C. lateralis* population structure in the East Bay.

Drivers of Population Structure and Historical Movement Patterns

There is essentially no suitable habitat for *C. lateralis* within the Amador–Livermore Valley at the present time—virtually all of it is urbanized, with the intersection of two major freeways creating four regional subdivisions (I-580 and I-680, respectively). Each subdivision includes a different management unit, and estimates of recent migration rates indicate that the populations in each unit are genetically isolated (Table S5). Given the location of these major freeways, the extent of urbanization, and the sensitivity of *C. lateralis* to urban edges (Mitrovich 2006), the absence of recent gene flow across Livermore Valley or among the different land-management units was not unexpected. However, several lines of evidence suggest that habitat within the Amador–Livermore Valley was probably unsuitable for *C. lateralis* long before 20th century land conversion came to fruition.

Uplift of the Berkeley Hills and the Diablo Range ~6 and 2.5 million yr ago, respectively, led to the development of the broad, topographically depressed Amador–Livermore Valley floor that allowed fluvial inputs to spread and percolate into the ground, rather than pool (Sloan 2006; Rosinski 2012). This shifting topography, combined with an east–west climate gradient driven by maritime influences from the San Francisco Bay and Pacific Ocean, gave rise to distinctive habitats in the Valley that persisted well into the 19th century (Stanford et al. 2013). In particular, the higher temperature and aridity in the eastern Livermore Valley has led to low ground water saturation and high soil alkalinity; whereas, cooler temperatures and heavier saturation from atmospheric, surface, and ground water inputs has led to a low alkaline environment in the west (US Fish and Wildlife Service 2002; Stanford et al. 2013). As a consequence, an alkali sink spanned a large area to the east and wetland–marsh predominated in the west, despite similar amounts of rainfall over the entire area (Fig. S6). These climate patterns currently persist, although neither end of the Valley bears any real resemblance to the original habitat. Most importantly, there is no documentation of *C. lateralis* occupying the alkali sink or wetland marsh.

In addition to climate and habitat factors, greater immigration and greater admixture between *C. lateralis* populations in western Alameda County might be linked to spatial arrangement alone, with populations in western Alameda County having more opportunity to exchange migrants from multiple directions (at least historically) compared with those in eastern Unit 5 at the edge of the San Joaquin Valley. Unit 5's eastern reach also encroaches on the distribution of the closely related *C. flagellum*, a snake with similar diurnal habits and diet but with greater tolerance for the xeric climate and open habitats of the San Joaquin Valley (Stebbins 2003; Mitrovich 2006). Interspecific competition, fewer options for migrant dispersal, and limits on environmental tolerance for both species might explain the higher migration rates toward lands that are close to the San Francisco Bay.

Similar factors might explain the repeated signal of greater east-to-west gene flow out of the northern ring terminus in Unit 4. Here, migrant exchangeability is limited to the northwest toward Unit 1 in the northern Berkeley Hills, where cooler climate prevails off San Pablo and San Francisco Bays. The southernmost sampling location in Unit 4, Los Vaqueros, is bounded by unsuitable habitat in the surrounding valley floors, all of which contribute to a dead end in the species' distribution on the north side of the Amador–Livermore Valley.

Migrant exchange across the Amador–Livermore Valley might have also been limited by historical habitat that varied with latitude. Sizeable estimates of seasonal and perennial wetlands (e.g., willow thicket, wet meadow, and marsh) existed on the north side of the valley as recently as the early 1800s, with dry grasslands extending across the south side (Fig. S6; Stanford et al. 2013). Although *C. lateralis* use grassland and riparian woodland, occupancy tends to be within, or adjacent to, more generalized chaparral and scrub mosaics (Swaim and McGinnis 1992; Swaim 1994; Alvarez et al. 2005; S. Bobzien, personal communication). Thus, *C. lateralis* might have always tended toward the foothills surrounding the Amador–Livermore Valley to the north, west, and south (Mt. Diablo, Berkeley Hills, and Mt. Hamilton Range, respectively), rather than to the east across the Altamont Hills (Fig. 1). Interestingly, there are also no specimen records for *C. lateralis* from the Altamont Hills, despite numerous museum specimens of other reptiles from this area.

In contrast to the north–south environmental gradients in the Amador–Livermore Valley, the ridgeline of the Berkeley Hills forms a fog shadow off the San Francisco Bay (particularly during the warmer months of the year when the snakes are active) that has led to greater habitat and climate consistency along the range's north–south axis. Given that there were fewer historical barriers preventing movement, this consistency could explain why migration rates among *C. lateralis* populations in the Berkeley Hills are not distinguishable in either direction. Unfortunately, contemporary gene flow estimates from the BayesAss analysis indicate that populations in different management units are more isolated now than in the past along this segment of the partial ring.

We note two caveats that should be considered when interpreting our evidence for ring-like population expansion in the East Bay. First, ABC requires that the data conform to

a tree-like structure to fit historical models (Beaumont et al. 2002; Cornuet et al. 2008). Although there is clear population structure with little-to-no contemporary gene flow among most of our sampling sites, the data might not be truly tree-like given the recency of habitat loss and disturbance. Second, sampling gaps in our study might have exceeded the distance at which *C. lateralis* movement would reveal recent migration between them; in turn, this might have contributed to the general lack of admixture signals between adjacent management units. Future studies should narrow these gaps and increase the number of genetic markers to better define the spatial limits of allele clines (e.g., in Unit 2, which bridges the gap between Units 6 and 3) and to reliably estimate migration at finer geographic scales.

Genetic Diversity within Management Units

Because of the protected status of *C. l. euryxanthus* and the degree of isolation among management units, we evaluated inbreeding coefficients and compared diversity measures among units to assess whether populations in different parts of the study area have less genetic variability than others. We found no differences in A_r , G_{is} , H_o , or H_s when we grouped sampling locations by management unit or cluster and no evidence that members of these groups are inbred. However, we draw attention to the finding that Unit 5A had less allelic diversity, less heterozygosity, and a smaller proportion of private alleles compared with other units (Table 1), and suggest that population dynamics at the range edge along the arid San Joaquin Valley might be different than those closer to the San Francisco Bay.

Mitochondrial haplotypes recovered near Unit 5A at the Alameda, San Joaquin, Santa Clara County, borders showed a somewhat contrasting pattern from the microsatellites in that haplotype diversity appears to be greater compared with other parts of the study area (Fig. 3). One explanation for these conflicting signals is that our sampling might have approached a major phylogeographic break, but did not extend far enough into the Diablo Range to appropriately detect it. This area roughly coincides with the putative boundary between *C. l. euryxanthus* and *C. l. lateralis*. Mismatches between mtDNA and nuclear gene boundaries are common in secondary contact zones and can be caused by different factors, including the effective size and evolutionary rates of different markers, local effects of natural selection, sex-biased dispersal, and assortative mating (reviewed in Brito 2007). Only with extended sampling to the south in the Diablo Range would we be able to confirm the pattern and explore its causes. Some of the signal conflict might also involve differences in the sampling geography between the microsatellite and mtDNA data sets—for this portion of the study area, microsatellite genotypes were limited to S300, whereas mtDNA haplotypes included S300 and several other sites to the west–southwest of S300 (location data in Appendix I of Supplemental Materials).

Edge populations, such as those in Units 4 and 5, often show less genetic diversity and greater differentiation compared with interior populations (Eckert et al. 2008; Richmond et al. 2013), and both theoretical and empirical studies have shown how environmental gradients, spatial variation in habitat suitability, realized niche limitations, and competition can contribute to population disequilibrium at

the margins of species' distributions (e.g., Sexton et al. 2009). These and other factors are likely relevant to understanding the population dynamics of *C. lateralis* in the southern study area and merit further investigation. They also indicate that reduced genetic diversity in Unit 5A might reflect evolutionary and demographic artifacts of life at the range edge, rather than a conservation concern (Eckert et al. 2008; Pearson et al. 2009).

Acknowledgments.—Funding for this project was provided by the US Fish and Wildlife Service (USFWS) Science Support Program, the USFWS's Central Valley Habitat Restoration Program, the USFWS's Central Valley Habitat Restoration Project, and the US Geological Survey Ecosystems Mission Area. We thank the East Bay Regional Parks District for land access, permission to conduct survey work, and general support of this research. The project benefited greatly from discussions, comments and insight by B. Solvesky (USFWS) and R. Stoelting. We thank J. Vindum (California Academy of Sciences) and C. Conroy (Museum of Vertebrate Zoology, U. C. Berkeley) for assistance with museum specimens; M. Westphal (Bureau of Land Management) and G. Pauly (Los Angeles County Museum of Natural History) provided additional field samples. E. Stoelting, B. Solvesky, E. Britt, A. Murphy, and four anonymous reviewers provided helpful comments on earlier versions of the manuscript. F. Wolven provided assistance with the genotyping at the CSUPERB microchemical core facility at San Diego State University. Any use of trade, product, website, or firm names in this publication is for descriptive purposes only and does not imply endorsement by the US government.

SUPPLEMENTAL MATERIAL

Supplemental material associated with this article can be found online at <http://dx.doi.org/10.1655/Herpetologica-D-15-00046.S1>

LITERATURE CITED

- Alvarez, J.A., M.A. Shea, and A.C. Murphy. 2005. A compilation of observations of Alameda whipsnakes outside of typical habitat. *Transactions of the Western Section of the Wildlife Society* 41:21–25.
- Arevalo, E., S.K. Davis, and J.W. Sites, Jr. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in Central Mexico. *Systematic Biology* 43:387–418.
- Beaumont, M.A., W. Zhang, and D.J. Balding. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035.
- Beerli, P. 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? Pp. 42–79 in *Population Genetics for Animal Conservation*, Vol. 17 (G. Bertorelle, M.W. Bruford, H.C. Hauffe, A. Rizzoli, and C. Vernesi, eds.). Cambridge University Press, UK.
- Beerli, P., and M. Palczewski. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185:313–326.
- Brito, P.H. 2007. Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology* 16:3423–3437.
- Chen, C., E. Durand, F. Forbes, and O. Francois. 2007. Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study. *Molecular Ecology Notes* 7:747–756.
- Clement, M., D. Posada, and K.A. Crandall. 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology* 10:1657–1660.
- Cornuet, J.-M., P. Pudlo, J. Veyssier, A. Delne-Garcia, M. Gautier, R. Leblois, J.-M. Marin, and A. Estoup. 2014. DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*. DOI: <http://dx.doi.org/10.1093/bioinformatics/btt763>.
- Cornuet, J.-M., F. Santos, M.A. Beaumont, C.P. Robert, J.-M. Marin, D.J. Balding, T. Guillemaud, and A. Estoup. 2008. Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24:2713–2719.
- Durand, E., F. Jay, O.E. Gaggiotti, and O. Francois. 2009. Spatial inference of admixture proportions and secondary contact zones. *Molecular Biology and Evolution* 26:1963–1973.
- Earl, D.A., and B.M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361.
- Eckert, C.G., K.E. Samis, and S.C. Loughheed. 2008. Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology* 17:1170–1188.
- Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial-DNA restriction data. *Genetics* 131:479–491.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Reviews of Ecology, Evolution, and Systematics* 34:487–515.
- Falush, D., M. Stephens, and J.M. Pritchard. 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
- Ford, L.D., and G.F. Hayes. 2007. Coastal scrub and coastal prairie. Pp. 180–207 in *Terrestrial Vegetation of California*, 3rd edition (M.G. Barbour, T. Keeler-Wolf, and A.A. Schoenherr, eds.). University of California Press, USA.
- François, O., S. Ancelet, and G. Guillot. 2006. Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* 174:805–816.
- François, O., and E. Durand. 2010. Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources* 10:773–784.
- Goudet, J. 2001. FSTAT: A program to estimate and test gene diversity and fixation indices (Version 2.9.3). Available at <http://www2.unil.ch/popgen/softwares/fstat.htm>. Archived by WebCite at <http://www.webcitation.org/6hbRhXvya> on 18 May 2016.
- Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.
- Hubisz, M.J., D. Falush, M. Stephens, and J.K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9:1322–1332.
- Jakobsson, M., and N.A. Rosenberg. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806.
- Jennings, M.R. 1983. *Coluber lateralis* (Hallowell), striped racer. *Catalogue of American Amphibians and Reptiles*. Society for the Study of Amphibians and Reptiles, USA.
- Jones, D.B., D.R. Jerry, M.I. McCormick, and L.K. Bay. 2002. Development of nine microsatellite markers for *Pomacentrus amboinensis*. *Molecular Ecology Resources* 8:1332–1334.
- Kalinowski, S.T. 2005. HP-Rare 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* 5:187–189.
- Knowles, L.L. 2008. Why does a method that fails continue to be used? *Evolution* 62:2713–2717.
- Meirmans, P.G. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* 24:3223–3231.
- Meirmans, P.G., and P.H. Van Tienderen. 2004. GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792–794.
- Mitrovich, M.J. 2006. A Case Study in Conservation Science: The Spatial Ecology and Evolutionary History of the Coachwhip Snake (*Coluber flagellum*) and Striped Racer (*C. lateralis*). Ph.D. dissertation, University of California, Davis, and San Diego State University, USA.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, USA.
- Paun, O., P. Schönswetter, M. Winkler, and A. Tribsch. 2008. Historical divergence vs. contemporary gene flow: Evolutionary history of the calcicole *Ranunculus alpestris* group (Ranunculaceae) in the European Alps and the Carpathians. *Molecular Ecology* 17:4263–4275.
- Pearson, G.A., A. Lago-Leston, and C. Mota. 2009. Frayed at the edges: Selective pressure and adaptive response to abiotic stressors are mismatched in low diversity edge populations. *Journal of Ecology* 97:450–462.
- Petit, R.J. 2008. The coup de grâce for the nested clade phylogeographic analysis? *Molecular Ecology* 17:516–518.
- Posada, D., K.A. Crandall, and A.R. Templeton. 2000. GeoDis: A program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* 9:487–488.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Pritchard, J.K., X. Wen, and D. Falush. 2010. Documentation for Structure software: Version 2.3. Available at <http://pritchardlab.stanford.edu/>

- structure.html. Archived by WebCite at <http://www.webcitation.org/6hbSG8EvT> on 18 May 2016.
- Pruett, C.L., P. Arcese, Y.L. Chan, A.G. Wilson, M.A. Patten, L.F. Keller, and K. Winkler. 2008. The effects of contemporary processes in maintaining the genetic structure of western song sparrows (*Melospiza melodia*). *Heredity* 101:67–74.
- Rambaut, A., M. Suchard, and A. Drummond. 2013. Tracer, v1.6. Available at <http://tree.bio.ed.ac.uk/software/tracer/>. Archived by WebCite at <http://www.webcitation.org/6hbScKfgf> on 18 May 2016.
- Richmond, J.Q., K.R. Barr, A.R. Backlin, A.G. Vandergast, and R.N. Fisher. 2013. Evolutionary dynamics of a rapidly receding southern range boundary in the threatened California Red-Legged Frog (*Rana draytonii*). *Evolutionary Applications* 6:808–822.
- Riemer, W.J. 1954. A new subspecies of the snake *Masticophis lateralis*. *Copeia* 1954:45–48.
- Riordan, E.C., and P.W. Rundel. 2013. The future of California coastal sage scrub in an era of increasing urbanization and global climate change. *Fremontia* 41:2–7.
- Rosenberg, N.A. 2004. DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes* 4:137–138.
- Rosinski, A.M. 2012. California geological survey zones of required investigation for liquefaction, Livermore Valley, California. Digital Mapping Techniques '10—Workshop Proceedings. Available at <http://pubs.usgs.gov/of/2012/1171/>. Archived by WebCite at <http://www.webcitation.org/6hbTE3auv> on 18 May 2016.
- Sexton, J.P., P.J. McIntyre, A.L. Angert, and K.J. Rice. 2009. Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics* 40:415–436.
- Sloan, D. 2006. Geology of the San Francisco Bay Region. University of California Press, USA.
- Spiegelhalter, D.J., N.G. Best, and B.P. Carlin. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 64:1–34.
- Stanford, B., R.M. Grossinger, J. Beagle, R. Askevold, R. Leidy, E. Beller, M. Salomon, C. Striplen, and A. Whipple. 2013. Alameda Creek Watershed Historical Ecology Study, SFEI Publication 679. San Francisco Estuary Institute, USA.
- Stebbins, R.C. 2003. Field Guide to Western Reptiles and Amphibians. Houghton Mifflin Company, USA.
- Swaim, K.E. 1994. Aspects of the Ecology of the Alameda Whipsnake (*Masticophis lateralis euryxanthus*). M.S. thesis, California State University Hayward, USA.
- Swaim, K.E., and S.M. McGinnis. 1992. Habitat associations of the Alameda whipsnake. *Transactions of the Western Section of the Wildlife Society* 28:107–111.
- Templeton, A.R., K.A. Crandall, and C.F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132:619–633.
- Templeton, A.R., E. Routman, and C.A. Philips. 1995. Separating population structure from population history: A cladistic analysis of geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782.
- Templeton, A.R., and C.F. Sing. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134:659–669.
- US Fish and Wildlife Service. 1997. Determination of endangered status for the Callippe silverspot butterfly and the Behren's silverspot butterfly and threatened status for the Alameda whipsnake. *Federal Register* 62:64306–64320.
- US Fish and Wildlife Service. 2002. Draft recovery plan for chaparral and scrub Community Species East of San Francisco Bay, California. Region 1, Portland, OR. US Department of the Interior, USA.
- US Fish and Wildlife Service. 2006. Endangered and threatened wildlife and plants: Designation of critical habitat for the Alameda whipsnake. *Federal Register* 71:58176–58230.
- Van Oosterhout, C., W.F. Hutchinson, D.P.M. Wills, and P. Shipley. 2004. Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4:535–538.
- Weir, B.S., and C.C. Cockerham. 1984. Estimating F statistics for the analysis of population structure. *Evolution* 38:1358–1370.
- Westman, W.E. 1981. Diversity relations and succession in Californian coastal sage scrub. *Ecology* 62:170–184.
- Wilson, G.A., and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177–1191.
- Wright, S. 1951. The genetical structure of populations. *Annals of Human Genetics* 15:323–354.

Accepted on 2 March 2016
Associate Editor: Adam Leache