Title
Landscape and Conservation Genetics of Amphibians and Reptiles in California

Author
Toffelmier, Erin Maurine

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Landscape and Conservation Genetics of Amphibians and Reptiles in California

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biology

by

Erin Maurine Toffelmier

2019

ABSTRACT OF THE DISSERTATION

Landscape and Conservation Genetics of

Amphibians and Reptiles in California

by

Erin Maurine Toffelmier

Doctor of Philosophy in Biology

University of California, Los Angeles, 2019

Professor Howard Bradley Shaffer, Chair

Examining patterns of diversity at fine and global spatial scales is an important component of to inferring underlying evolutionary mechanisms, understanding species distributional patterns, and informing conservation. Globally, amphibians and reptiles are among the fastest declining taxonomic groups, and now more than ever, it is necessary to quantify diversity and its spatial drivers in order to most effectively conserve species. In this dissertation, I examine the population, landscape, and conservation genomics of several species along a continuum of endangerment, from highly endangered and on the brink of extinction to widespread and abundant. Throughout, I use large-scale molecular data sets coupled with spatial

analyses to examine spatial genetic diversity in these varied species. My goals were to contribute to our understanding of how genetic diversity is distributed across a multitude of landscapes and to provide genetic context for the conservation of these species.

In Chapters 1 and 2, I examined how genetic diversity is spread across the limited ranges of two ecologically disparate species, California tiger salamanders, *Ambystoma californiense,* in Santa Barbara County, and the Panamint alligator lizard, *Elgaria panamintina,* found only in the isolated desert mountain ranges of eastern California, and found surprising parallels. In both, I found populations with exceedingly low levels of genetic diversity and genetic effective population sizes. For tiger salamanders, genetic diversity and divergence is strongly correlated with the number of suitable breeding habitats in regional neighborhoods and presence of natural vernal pools, while divergence across the range of *E. panamintina* is primarily mediated by geographic distance. In both cases, our findings have important implications for how management and mitigation efforts may more effectively assist the recovery and/or protection of these groups. In Chapter 3, I examined the drivers of spatial genetic structure in the widespread southern alligator lizard, *Elgaria multicarinata*. I found that patterns of genetic isolation are driven primarily by geographic distances, but that regional ecological niches have also diverged. Collectively, my work demonstrates the utility of integrating genetic and spatial analyses across spatial scales to help elucidate how genetic diversity is distributed across variable landscapes.

The dissertation of Erin Maurine Toffelmier is approved.

Kirk Edward Lohmueller

Peter L. Ralph

Thomas Bates Smith

Howard Bradley Shaffer, Committee Chair

University of California, Los Angeles

2019

This dissertation is dedicated to my family, for the eternal love and support, and to Robert Cooper, without whose shared love of the natural world I would be lost.

TABLE OF CONTENTS

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I would like to first thank Brad Shaffer for his mentorship and unwavering support through the years. Both through his advice and example, I have learned a great deal about what it means to be a scientist, a conservationist, and a herpetologist. I thank the members of my committee for their counsel and for setting examples of quality work. I would also like to thank my previous mentor, Tom Smith, who took a chance on someone with virtually no experience and let me learn along the way.

I am immensely grateful for the support Robert Cooper, whose enthusiasm for nature and science is infectious and has been a constant source of inspiration. I thank my parents, Cindy Combs and Mark Toffelmier, for fostering a love of knowledge and learning. They never questioned my love of creatures and dirt, and for that I am eternally grateful. I thank my brother, Dan Toffelmier, for catching me my first lizards and for setting an example of how to make the world a better place.

I have been fortunate to have spent my academic career at an institution as outstanding as UCLA. Accordingly, the list of people who have served as role models, teachers, collaborators, and friends is long. I especially thank members of the Shaffer lab, who have made this journey an adventure. Thanks to Joscha Beninde, Gary Bucciarelli, Adam Clause, Mario Colon, Robert Cooper, Natalia Gallego-García, Müge Gidiş, Jesse Grismer, Tara Luckau, Evan McCartney-Melstad, Genevieve Mount, Peter Scott, Sid Shah, Phil Spinks, Isolde van Riemsdijk, Jannet Vu, Sarah Wenner, and Ben Wielstra. I also thank the staff of the Department of Ecology and Evolutionary Biology, particularly Tessa Villaseñor and Jocelyn Yamadera, for keeping me on track.

This dissertation work was supported by a NSF Graduate Research Fellowship, a UCLA Dissertation Year Fellowship, as well as research funding from the UCLA La Kretz Center for California Science, the UCLA Department of Ecology and Evolutionary Biology, and the US Fish and Wildlife Service.

H. Bradley Shaffer was the Principle Investigator on all chapters, provided guidance on analytical approaches and framework, and edited all manuscripts. All chapters represent drafts of manuscripts intended for submission for publication. For Chapter 1, Adam G. Clause did the vast majority of the field work, I performed all lab work and analyses and wrote the manuscript. For Chapter 2, samples were contributed to H.B. Shaffer's tissue collection by many dedicated wildlife professionals over the last 30 years, I performed the laboratory work, conducted all analyses, and wrote the manuscript. For Chapter 3, DNA extracts were supplied by Chris R. Feldman for a portion of the samples and the remainder were obtained from many additional sources. I conducted all additional lab work and analyses, and wrote the manuscript.

VITA

EDUCATION

B.S. in Biology, University of California, Los Angeles, 2008

HONORS, AWARDS AND FELLOWSHIPS

2012-2015	National Science Foundation Graduate Research Fellowship Program Fellow
2013	National Academies Education Fellow in the Life Sciences
2013	La Kretz Graduate Grant
2014	Santa Monica Bay Audubon Student Grant
2014	UCLA EEB Graduate Research Award
2015	Stunt Ranch NRS Research Award (Dissertation Research)
2015	Stunt Ranch NRS Research Award (Bruin Naturalist Club Bioblitz series)
2015	UCLA EEB Graduate Research Award
2016	UCLA La Kretz Center Graduate Grant
2017	UCLA EEB Graduate Research Award; Summer
2017	UCLA EEB Graduate Research Award; Fall
2018	UCLA La Kretz Center Graduate Grant
2018	UCLA EEB Graduate Research and Travel Award, Summer
2018	UCLA Dissertation Year Fellowship

PUBLICATIONS

Fuller, T; Saatchi, S; Curd, E; **Toffelmier, E**; Thomassen, H; Buermann, W; DeSante, DF; Nott, MP; Saracco, JF; Ralph, CJ; Alexander, J; Pollinger, J; Smith, TB. Mapping the Risk of Avian Influenza in Wild Birds in the US. *BMC Infectious Diseases*. 2010. 10:187.

Njabo, K; Cornel, A; Bonneaud, C; **Toffelmier, E**; Sehgal, R; Valkiunas, G; Russell, AF; Smith, TB. Non-specific Patterns of Vector, Host and Avian Malaria Parasite Associations in a Central African Rainforest. *Molecular Ecology*. 2010. 20:5.

Curd, E; Pollinger, J; **Toffelmier, E**; Smith, TB. Rapid Influenza A Detection and Quantification using a One-Step Real-Time Reverse Transcriptase PCR and High Resolution Melting. *Journal of Virological Methods.* 2011. 176:1-2.

Shaffer, H B; Gidiş, M.; McCartney-Melstad, E; Neal, K M; Oyamaguchi, H M; Tellez, M; and **Toffelmier, E**. Conservation Genetics and Genomics of Amphibians and Reptiles. *Annual Review of Animal Biosciences*. 2015. 3 (no. 1):113–38.

Clause, AG; **Toffelmier, E.;** Shaffer, HB. "Project overview and field sampling protocol for genomic research on the panamint alligator lizard *Elgaria panamintina.*" Status review for the Pacific Southwest Regional Office of the US Fish and Wildlife Service. 2017.

Toffelmier, E; Shaffer, HB. "Preliminary Report on the Range-wide Genetics of the Panamint Alligator Lizard, *Elgaria panamintina*." Unpublished report submitted to the Carlsbad Fish and Wildlife Office by the UCLA La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, Los Angeles. 2017.

PRESENTATIONS

Toffelmier, E; Shaffer, HB. "Metapopulation dynamics of the critically endangered Santa Barbara Distinct Population Segment of the California tiger salamander, *Ambystoma californiense.*" Joint Meeting of Ichthyologists and Herpetologists. New Orleans, Louisiana. 2016. (Oral)

Toffelmier, E; Shaffer, HB. "Metapopulation dynamics of the critically endangered Santa Barbara Distinct Population Segment of the California tiger salamander, *Ambystoma californiense.*" Amphibian Population Task Force Meeting. Santa Barbara, California. 2017. (Oral)

Toffelmier, E; Shaffer, HB. "Metapopulation dynamics of the critically endangered Santa Barbara Distinct Population Segment of the California tiger salamander, *Ambystoma californiense.*" US Fish and Wildlife Service; Ventura Branch. Ventura, California. 2017. (Oral)

Toffelmier, E; Shaffer, HB. "Metapopulation dynamics of the critically endangered Santa Barbara Distinct Population Segment of the California tiger salamander, *Ambystoma californiense.*" California Tiger Salamander Working Group Meeting. Fort Ord, California. 2017. (Oral)

Toffelmier, E; Shaffer, HB. "Introgression of non-native alleles in the critically endangered Santa Barbara Distinct Population Segment of the California tiger salamander, *Ambystoma californiense.*" USFWS Hybridzation Webinar. Los Angeles, California. 2018. (Oral)

Toffelmier, E; Shaffer, HB. "Conservation and landscape genomics of the endangered California tiger salamander, *Ambystoma californiense*". Joint Meeting of Ichthyologists and Herpetologists. Rochester, NY. 2018 (Poster)

Toffelmier, E; Clause, AG.; Shaffer, HB. "Landscape genetics and connectivity in the Panamint alligator lizard, *Elgaria panamintina*". Joint Meeting of Ichthyologists and Herpetologists. Rochester, NY. 2018 (Poster)

CHAPTER 1

CONSERVATION AND LANDSCAPE GENETICS OF AN ENDANGERED VERNAL POOL AMPHIBIAN, *AMBYSTOMA CALIFORNIENSE*

Erin Maurine Toffelmier

ABSTRACT

Habitat loss and fragmentation are leading drivers of species and population declines. Changes to the arrangement and quality of important habitat features may interrupt gene flow and, in turn, lead to associated reductions in diversity and fitness if the effects of drift are strong in newly isolated populations. To better understand how changes to habitat configuration may impact populations, it is important to understand spatial population structure and how the landscape may promote (or impede) the maintenance of diversity. We used thousand of genome-wide genetic markers to examine population genetic structure, gene flow, and patterns of genetic diversity in the heavily impacted Santa Barbara County Distinct Population Segment of the California tiger salamander*, Ambystoma californiense*. We found that genetic diversity in population is among the lowest measured in any vertebrate and further that this diversity is unevenly distributed across its small range. We used environmental niche modeling and landscape resistance surfaces to identify important landscape variables that promote gene flow between breeding habitats. Across the region, proximity to potential breeding ponds was an important correlate of gene flow. When we further examined the local distribution of suitable breeding habitat, we found that local genetic diversity (i.e. of a individual vernal pool population) was positively correlated with the number of vernal pools within 2200m, the

maximum dispersal distance observed for this species. We also found that existing genetic diversity was closely correlated with the number of nearby natural vernal pools, suggesting that regional diversity is driven primarily by the presence of longstanding breeding habitat, rather than recently constructed habitat. Our results indicate that management efforts should include preservation of the few remaining natural ponds as reservoirs of diversity. Additionally, maintaining or promoting connectivity through suitable habitat will be essential to the success of conservation efforts for *A. californiense* in Santa Barbara County.

INTRODUCTION

A growing body of evidence suggests that change in habitat configuration through fragmentation is a leading driver of population and species declines (Collins and Storfer, 2003a; Cushman, 2006; DiLeo and Wagner, 2016; Grant et al., 2016). Metapopulation (Hanski, 2015; Hanski et al., 2013) and landscape genetic theory (McRae, 2006; Wright, 1943) also support the prediction that changes in habitat configuration, including disruption of connectivity and isolation, may lead to population declines. These interruptions lead to the loss of genetic diversity through isolation and drift, and may negatively impact populations by reducing fitness, limiting their ability to respond to environmental change or disease, and ultimately increasing extinction risk (DiLeo and Wagner, 2016; Frankham, 2005).

Amphibians are one of the fastest declining taxonomic groups worldwide (Collins and Storfer, 2003b; Hoffmann et al., 2010; Stuart et al., 2004). Pond breeding amphibian populations are particularly susceptible to changes in landscape configuration given their dependence on both aquatic breeding habitat and terrestrial upland habitat (and the need to migrate between the two),

relatively limited dispersal ability, and reliance on between-patch dispersal for maintenance of regional populations (Allentoft and O'Brien, 2010; Cushman, 2006; Hecnar and M'Closkey, 1996). For at-risk species, these qualities can elevate vulnerability in threat assessments, and the California Amphibian and Reptiles of Special Concern considers "population concentration/migration" to be one of eight key threats to persistence (Thomson et al., 2016). Unfortunately, it can be difficult to assess the impact of fragmentation, particularly at small spatial scales, although expanded genetic data sets now offer greater resolution to examine population structure and gene flow on large and fine spatial scales (Manel et al., 2003; McCartney-Melstad and Shaffer, 2015; Shaffer et al., 2015).

To mitigate the negative genetic consequences of isolation and loss of connectivity, an ecologically-informed understanding of how genetic diversity and connectivity are distributed across the landscape is needed (Cushman, 2006). Wetland loss and augmentation is one of the major threats to amphibian populations, and in many regions where natural ponds are rare, the majority of amphibian breeding ponds are constructed agricultural or stock (Drayer and Richter, 2016). In systems that have lost natural ponds, artificial ponds therefore often become essential elements for sustaining local populations and regional connectivity (Knutson et al., 2004; Semlitsch, 2002). Success of artificial ponds, as measured by species abundance or richness, has often been linked to the specific attributes of these ponds, such as presence of fish, water quality, and topography (Drayer and Richter, 2016; Porej and Hetherington, 2005; Shulse et al., 2010). However, constructed ponds often fall short of replicating natural ponds in achieving natural community composition (Drayer and Richter, 2016), and investigations of the population genetic impact of artificial breeding ponds has been limited (Furman et al., 2016; Wang et al., 2011).

Because the modification of habitats will undoubtedly continue, understanding the role of artificial ponds on the landscape will be an essential part of conservation management.

The Santa Barbara County distinct population segment of the California tiger salamander (*Ambstoma californiense,* "SBCTS") is an example of an amphibian on the brink of extinction. This distinct population segment (DPS) was first emergency listed as Endangered in 2000 under the Federal Endangered Species Act of 1973 (U.S. Fish and Wildlife Service, 2000a, 2000b) based on its phylogenetic and geographic distinctness, as well as immediate conservation concern (Shaffer et al., 2004; U.S. Fish and Wildlife Service, 2000a, 2000b). Later, the state of California, which does not use the DPS management approach, included the SBCTS as part of the state listing of the entire species (California Department of Fish and Game (CDFG), 2010). The decline of SBCTS populations is largely attributed to habitat loss and fragmentation (Davidson et al., 2002; U.S. Fish and Wildlife Service, 2000b). The entire range of SBCTS, particularly on the northern end of its range, has experienced extensive land conversion over the last 150 years, which impacts both the aquatic breeding and terrestrial non-breeding habitats on which the species relies. Historically, livestock ranching predominated, which frequently results in the construction of artificial, but suitable breeding ponds (U.S. Fish and Wildlife Service, 2000a). More recently, large areas have been converted to row-crops and vineyards, a process that often results in complete habitat loss (Curado et al., 2011). Combined with recent urbanization, SBCTS has seen an enormous loss of its preferred grassland habitat and vernal pool breeding sites, resulting in small, isolated patches of suitable habitat across its limited range. This has undoubtedly led to population isolation and dissociation, which will likely have long term impacts on population viability (Frankham et al., 2017).

A new and looming threat for SBCTS is potential hybridization with non-native barred tiger salamanders, *Ambystoma mavortium*. In the Central DPS of *A. californiense*, hybridization has become a significant conservation concern with respect to the genetic integrity of the species and the cascading changes in vernal pool ecosystem functionality (Fitzpatrick and Shaffer, 2007; Ryan et al., 2009; Searcy et al., 2016). Non-native genotypes have been identified in several ponds in the Santa Barbara region (U.S. Fish and Wildlife Service, 2016), although the spatial and genetic extent of introgression has not been quantified since 2011. Because tiger salamanders are capable of relatively long-distance dispersal, the existence of non-native populations poses a real and immediate threat to native populations. Identification of geographic regions where non-native introgression has occurred will therefore be invaluable in determining how management actions should proceed.

Because *A. californiense* are iteroparous with overlapping generations and relatively long lifespans, one or a few years of poor or null recruitment does not necessarily lead to local population extinction. However, several years of limited breeding or recruitment may hasten genetic drift in a population, which in turn can lead to population decline (Vucetich and Waite, 1999). These effects of extreme drift or local extinction could be ameliorated by dispersal from neighboring ponds (Frankham et al., 2017), and in many cases sets of proximal vernal pools function as a metapopulation where neighboring breeding populations exchange individuals. Based on the presumed existence of such groups of populations, the Santa Barbara DPS is managed as six units, each of which is a cluster of adjacent ponds separated by small watershed boundaries and estimates of upland population movement (U.S. Fish and Wildlife Service, 2016). Empirical research from other parts of the range suggests that *A. californiense* populations consist of multiple suitable breeding ponds with relatively high intra-pond dispersal and are not

truly independent patches as required by classic metapopulation framework (Hanski, 1994; Trenham et al., 2001). It therefore seems reasonable that the configuration and extent of these neighborhoods and the surrounding landscape should play an important role in maintaining *A. californiense* diversity.

In this study, we examine the effect of landscape configuration and the distribution of natural and artificial ponds on genetic diversity and differentiation in the Santa Barbara tiger salamander. We examine the population genomics of extant populations with the following goals: 1) quantify the genetic health, as measured by genetic variation, of populations 2) identify potential landscape correlates of genetic diversity, effective population size and divergence at the neighborhood and range-wide level, 3) assess population structure and evaluate the reality of the six-metapopulation framework, and 4) determine the geographic range of non-native genotypes across the management unit.

MATERIALS AND METHODS

Sample collection and genetic data generation

We used samples collected from 1986 to 2017 from across the range of SBCTS (total N = 471). Samples consisted of field collected larval tail tips (N = 431) and post-metamorphic tail tips from terrestrial individuals (N = 40) collected opportunistically from road kills or pitfall traps as part of local monitoring efforts. These represent 61 localities (both aquatic and terrestrial), including at least one breeding pond from each metapopulation and representatives from virtually every accessible breeding site and most of the known, remaining ponds utilized by the DPS.

We extracted genomic DNA using a salt extraction method following Sambrook and Russel (2001), and generated reduced representation exon target capture libraries following McCartney-Melsted, Mount, and Shaffer (2016). Ambystomatid salamanders have extremely large and repetitive genomes, making genomic data generation difficult with standard methods. The estimated genome size is presumably the same as the closely related *Ambystoma mexicanum*, ~32 gigabases (Keinath et al., 2015). The protocol targets 5,237 genic regions from across the genome based on expressed sequence tag sequences from the *A. mexicanum* (McCartney-Melstad et al., 2016). We sequenced samples over several Illumina HiSeq 4000 lanes at the UC Berkeley Vincent Coats Genomic Sequencing Laboratory. Raw sequence data were trimmed of barcodes and for quality (Q30) in CUTADAPT v.1.12 (Martin, 2011), and we followed the Genome Analysis Tool Kit Best Practices pipeline for sequence data preparation to align reads and call genotypes (DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2002). For alignment, we used a pseudo-reference genome that consisted of reciprocal best blast hits for the sequence capture targets, aligned with BWA-MEM (Li, 2013; McCartney-Melstad et al., 2016). We first used GATK's *genotypeGVCFs* to jointly genotype all samples. Because genomically similar, non-native barred tiger salamanders (*A. mavortium,* BTS) are known from the region, we assembled three data sets. First, we generated a single nucleotide polymorphism (SNP) data set for all samples. We then removed samples identified as BTS (see below for methods) and re-called genotypes. For these pure SBCTS samples, we generated two genetic data sets: all sequenced bases including invariant sites (with the *IncludeNonVariantSites* flag in *GenotypeGVCFs*) and a variants-only data set (*GenotypeGVCFs* with default settings). For all three datasets (all samples, SBCTS-allsites, SBCTS-variants), we retained sites that passed basic filters suggested by GATK (QualbyDepth < 2.0, RMSMappingQuality<40.0,

FisherStrand > 60.0, MappingQualityRankSum < -12.5, ReadPosRankSum < -8.0, Quality <30) and excluded insertions and deletions. To remove potentially paralogous loci, we excluded SNP variants with more than one alternate allele and loci exhibiting heterozygote excess ($p < 0.05$). We used VCFTOOLS (Danecek et al., 2011) to calculate the probability of heterozygote excess. We further filtered the SBCTS-variants dataset by selecting one SNP per locus that had less than 50% missingness across all SBCTS samples.

Non-native introgression

Introgression from non-native *A. mavortium* introduced to California in the 1950s has been identified as a potential management issue in Santa Barbara County (Fitzpatrick et al., 2010; Riley et al., 2003; U.S. Fish and Wildlife Service, 2016) and non-natives have been found in close proximity to native populations (Johnson et al., 2011). To examine introgression of non-native alleles, we used a SNP database comprised of 16,933 SNP variants, which are fixed for different alleles in pure *A. californiense* and pure BTS (E. McCartney-Melstad, pers. comm.). These fixed differences were identified from a panel that includes known non-hybrid samples from across the geographic range of *A. californiense* (including SBCTS), samples from representative regions in the native range of BTS, and samples from known introduced populations of BTS in California. For each sample, we calculated the non-native score as the total number of BTS alleles divided by the total number of alleles successfully sequenced in that individual.

Neighborhood qualities and population genetic diversity

Because adult *A. californiense* can migrate between breeding ponds over relatively long distances with moderate frequency, ponds in close proximity to each other might functionally operate together as a single population. In the Central DPS populations of *A. californiense*, adults have been documented up to 2200 meters away from the nearest breeding ponds, while 95% of the adult population is estimated to occur within 0.6 -1.8 km from the pond edge (Orloff, 2011; Searcy and Shaffer, 2011; Searcy et al., 2013; Trenham et al., 2001). Probability of dispersal between ponds is a non-linear function of intra-pond distance with a median adult migration distance of 667m (Searcy et al., 2013; Trenham et al., 2001). Dispersal in SBCTS has not been studied as extensively, although recent landscape-level mark-recapture data suggest that SBCTS exhibit similar dispersal distances to populations in the Central DPS (Searcy and Shaffer, Unpublished). We therefore examined clusters of ponds ("neighborhoods") within the known median and maximum dispersal distances of adult SBCTS (667m and 2200m, respectively). These distances also roughly correspond to the protected upland habitat requirements set out in the SBCTS Recovery plan (U.S. Fish and Wildlife Service, 2016). For each genetically sampled pond, we determined the statistical relationship between environmental attributes of potentially interacting ponds and genetic descriptors of population health.

Focal pond and neighborhood environmental attributes: Focal pond attributes included pond type (natural or artificial), surface area of focal pond, and distance to the nearest pond. We also calculated a metric of connectivity relating the number of, and distance to, neighbor ponds for each focal pond (Moilanen and Nieminen, 2002; Peterman et al., 2015). We calculated the connectivity, C_i, of pond i as:

$$C_i = \Sigma[\ exp(-kd_{ij})\ ] \quad [1]$$

where d_{ij} is the Euclidean geographic distance between focal pond i and pond j, we sum across all populations ($j \neq i$), and k is a scaling parameter equal to 1/(average dispersal distance). For each focal pond, we calculated connectivity to all ponds and connectivity to only natural ponds using an empirically derived migration distance of 667m (Moilanen and Nieminen, 2002; Peterman et al., 2015; Searcy et al., 2013). To assign ponds to neighborhoods, we calculated a 667m or 2200m buffer around each pond and assigned ponds to the same 667m or 2200m neighborhood if the buffers around adjacent ponds overlapped (Figure 1.1A). If a neighborhood did not have at least one genetically sampled representative, it was dropped as a focal pond from further analyses. For both 667m and 2200m neighborhoods, we tallied the total number of ponds in the neighborhood, and the proportion of ponds in the neighborhood that were natural, based on records provided by the US Fish and Wildlife Service (USFWS, pers. comm.).

Population Genetics: We calculated population diversity metrics for each pond to examine the relationship between environmental attributes and focal pond genetics. Most adult samples were collected within several hundred meters of a breeding pond, so we assigned adult samples to this nearest pond locality to utilize them for genetic analyses. We used the SBCTS-allsites genetic data set to calculate nucleotide diversity (π) across all samples within each pond for every sequenced base using VCFTOOLS (*--site-pi* option) and report the mean across all sequenced bases. We used the SBCTS-variants dataset to calculated genetic diversity, differentiation and effective population sizes. For each pond, we calculated observed per-SNP heterozygosity (H_o), gene diversity (H_s, also called expected heterozygosity), allelic richness (A_R), and population

inbreeding coefficient (F_{IS}) in *hierfstat* (Goudet, 2005). Pond values for $\overline{H}_o$, $\overline{H}_s$, $\bar{A}_R$, and $\overline{F}_{IS}$ are reported as the mean of the per-SNP values (excluding missing data). We calculated individual inbreeding (F_{ind}) in VCFTOOLS and calculated the mean for each pond ($\overline{F}_{ind}$). We calculated pairwise F_{ST} between all pond pairs and assessed significance by randomly permuting SNPs in 1000 bootstrap iterations. We conducted an analysis of variance (ANOVA) in *R* to test whether genetic attributes differed between the metapopulations. When the effect was significant ($P < 0.05$), we used a Tukey post-hoc test for differences between metapopulation pairs. We included the East Santa Maria (ESM), West Santa Maria (WSM), West Los Alamos (WLA), Purisima (PUR), and Santa Rita (SAR) metapopulations in the ANOVAs, but excluded the East Los Alamos (ELA) metapopulation because it is represented by a single locality. For localities with multiple years of sampling, we first estimated diversity indices per pond-year (i.e. for groups of samples collected from the same pond in the same year) and tested for significant differences across years using Wilcoxon rank sum tests in *R*. We also tested whether diversity indices across all ponds and years sampled in the DPS have changed through time with linear mixed models in *nlme* v. 3.1-99 (Pinheiro et al., 2019) setting metapopulation as a random effect. We then pooled samples across years within ponds to re-calculate genetic diversity and differentiation metrics for the neighborhood analyses. We separated samples by pond-year to calculate effective population size (N_e) in NEESTIMATOR using the molecular co-ancestry method with a minor allele frequency cutoff of 0.05 (Do et al., 2014). For ponds with multiple pond-year estimates, we tested for significant differences between time points with a Wilcoxon rank sum test and tested for changes through time across all ponds and years sampled in the DPS with a linear mixed model in *nlme*, setting metapopulation as the random effect. For ponds with multiple pond-year estimates, we generated a mean N_e to use in the neighborhoods analyses below.

Statistical approach: To test the hypothesis that pond and neighborhood attributes contribute to genetic diversity, we took an information theoretic approach to examine competing hypotheses simultaneously and pick the best set of models (Grueber et al., 2011). We fit generalized least squares (GLS) models relating pond/neighborhood environmental attributes to patterns of diversity ($\overline{H}_o$, $\overline{H}_s$, $\bar{A}_R$, π, $\bar{F}_{IS}$, $\bar{F}_{ind}$), and effective population size (N_e). Each model also included a spatial correlation structure based on the pond coordinates (longitude and latitude calculated from the centroid of each pond) to account for genetic spatial autocorrelation in the response variables. For each response, we included three types of predictor variables: 1) focal pond attributes, 2) attributes of its 667m neighborhood, and 3) attributes of its 2200m neighborhood. We also included two interaction terms that reflect the interaction between the total number of ponds and the proportion of natural ponds for each neighborhood size. Because the predictors were measured on different scales, all numeric predictors were centered on zero and then scaled by dividing by the standard deviation of each centered predictor. We used *nlme* to construct individual GLS models and *MuMIn* v. 1.43.6 (Bartoń, 2019) to compare models with all possible combinations of predictor variables. We performed model averaging and report the conditional average parameter estimates from models with $\Delta AICc < 2$ (Burnham et al., 2011; Dochtermann and Jenkins, 2011; Grueber et al., 2011).

Population structure

To identify genetic clusters and examine differentiation across the range of SBCTS, we conducted a principle components analyses in *SNPRelate* v1.6.4 (Zheng et al., 2012). We also examined DPS-wide population structure using the Bayesian method implemented in

FASTSTRUCTURE V1.0 (Raj et al., 2014). We first ran 10 replicates for each of k=1 to k=20 with the logistic model and examined the marginal likelihoods with the *chooseK.py* function to identify the best-supported value of k across replicate runs. We then iteratively re-ran FASTSTRUCTURE for the northern and southern metapopulations to examine fine-scale structure, testing values of k = 1 to k = 15 with 10 replicates each. We used the *choosek.py* function to examine marginal likelihoods.

Landscape modeling

We used species distribution modeling and landscape resistance surfaces to examine habitat suitability, gene flow and genetic differentiation across the range. First, we generated a species distribution model in MAXENT V3.4.1 with 42 landscape and environmental layers including digital elevation model ("DEM", US Geological Survey, 2013); slope, aspect, TPI (topographic position index, extracted from the DEM using the R package *raster* v2.8-19; Hijmans et al., 2019); 19 bioclimatic variables (Karger et al., 2017), 18 expanded bioclimatic and topographic variables (Title and Bemmels, 2018); land cover and canopy cover (U.S. Geological Survey, 2011a, 2011b), and the distance to nearest freshwater lake or pond ("DP"). Because SBCTS require freshwater ponds or lakes to breed, the DP layer allows us to test the hypothesis that the species breeding distribution is limited by distance to suitable water bodies. To generate the DP layer, we extracted "freshwater pond" and "lake" features from the National Wetlands Inventory (U.S. Fish and Wildlife Service, 2019) and then calculated a new raster layer as the distance of each pixel to the nearest pond or lake (Hijmans et al., 2019). All layers were normalized to 30 arc-second resolution and cropped to within 10km of the metapopulation boundaries on the eastern and northern sides of the DPS and were bounded by the Pacific Ocean

on the western and southern sides of the DPS. We used 142 locality records for this DPS, based on our own collections work and records from VertNet (www.VertNet.org), to generate 10 replicate MAXENT models, with default parameters. Each locality record is a single point from a unique geographic location, primarily representing breeding ponds. Here we report the average of these 10 models (hereafter referred to as "SDM") and rank variable contribution by permutation importance (Phillips and Dudík, 2008; Searcy and Shaffer, 2016).

To examine gene flow and differentiation in a landscape context we used *ResistanceGA* v4.0-10 in R to generate landscape resistance surfaces. This approach uses maximum likelihood population effects (MLPE) models to evaluate relative support for the fit of genetic distance matrices to a variety of surface transformations (Peterman, 2018). We used pairwise F_{ST} as the response variable, and included only ponds with $N > 2$. We calculate resistance distances as commute distance, and evaluated model fit with *AICc*. Because optimization of all possible multi-surface combinations would be computationally intractable, we first optimized resistance surfaces individually to identify candidate layers for multi-surface optimization. To further streamline optimization, we masked all environmental and landscape layers to a 10km-buffered minimum convex polygon derived from all known breeding ponds and terrestrial localities. We optimized single surfaces for each of the 42 landscape and environmental layers, the SDM and a null model of Euclidean distance with three replicates each with different random number seeds. From these single surface optimizations, we chose a subset of surfaces with $\Delta AICc < 10$ across all three replicates. We then optimized single surfaces and combinations of 2 or 3 surfaces in these subsets and conducted a pseudo-bootstrap procedure that refits the MLPE models with a random subset of localities to rank single and combined surface models by their average *AICc* among bootstrap replicates. For each repetition, we included geographic distance and an

intercept-only model for the final bootstrap analysis. We report the best-fit multi surface models within 10 *AICc* units of the best scoring model ($\Delta AICc < 10$) and for each model report the $\Delta AICc$, marginal R^2, and proportion of times the model was the top scoring among bootstrap replicates, averaged across the three replicates.

RESULTS

Genotyping and variant filtration

Sequencing yielded an average of 2,258,247 (SD 1,332,068) reads per individual. After joint genotyping and basic filtering, we recovered a total of 5,041 biallelic loci for the 471-sample data set. After removing BTS samples (see below) and conducting joint genotyping on the CTS samples, we recovered 2,076,838 mono- and bi-allelic sites across all CTS samples, with an average of 1,859,722 bases per sample. From this, we generated a variant dataset consisting of 3,811 biallelic SNPs. A single individual exhibited high individual heterozygosity (18.7% of variable sites versus ~2% in most other samples), a potential indicator of sample contamination or unexpected introgression, and was removed from further analyses.

Non-native introgression

From the panel of diagnostic SNPs, we identified non-native genotypes in 22 out of 471 samples. In each of these 22 samples, >98% of the diagnostic positions were homozygous for BTS. These represent all of the samples collected from four source ponds, and no BTS genotypes were identified at any other sites.

Neighborhood qualities and population genetic diversity

Pond and neighborhood attributes: Thirteen out of 33 genetically sampled ponds were natural and 26 out of 66 known breeding ponds (including non-sampled ponds) were natural. Surface area across all focal (genetically sampled) ponds ranged from 240 m^2 to 91,284 m^2 (median 3,184 m^2). Mean distance to nearest pond for focal ponds was 548m ± 49. Mean focal pond connectivity was 1.51 ± 1.15, and mean neighborhood size ranged from 1 to 20 known breeding ponds in 667m neighborhoods, and 1 to 23 known breeding ponds in 2200m neighborhoods. On average, 48% of all known breeding ponds were natural in 667m neighborhoods, and 55% were natural in 2200m neighborhoods.

Neighborhood trends in diversity: Across all ponds and years, $\overline{H}_S$ and F_{IS} were slightly negatively correlated with time ($\overline{H}_S$: β = -0.00012 , t-value =-2.29, p = 0.0270; F_{IS}: β = -0.003053 t-value = -2.03, p = 0.0485), while $\overline{H}_o$, $\bar{A}_R$, and π were not ($\overline{H}_o$: p = 0.42, $\bar{A}_R$: p = 0.30, π: p = 0.19). ANOVA tests of genetic variation did not differ between years in ponds with multiple years of sampling ($\overline{H}_S$: W = 69, p = 0.61, $\overline{H}_o$: W = 58, p = 0.90, $\bar{A}_R$: W = 67, p = 0.70). We therefore report diversity values for samples pooled across years for each pond. However it is measured, genetic variation was extremely low across the DPS (π: 0.00019 ± 5.95e-05; $\overline{H}_o$: 0.0162 ± 0.0061; $\overline{H}_S$: 0.162 ± 0.0067, $\bar{A}_R$: 1.02 ± 6.6e^{-3}). We report all per locality metrics in Table 1.1. Population inbreeding, measured as F_{IS}, ranged from -0.32 to 0.16 (mean: -0.04 ± 0.10; global: 0.29). Individual inbreeding averaged within ponds ($\overline{F}_{ind}$) ranged from -0.79 to 0.73 (mean: 0.17 ± 0.17. Pairwise F_{ST} ranged widely (from 1e-13 to 0.54), was generally high (mean: 0.20 ± 0.13, global: 0.30) and increased with geographic distance within metapopulations and across the DPS (Figure 1.2A). The geometric mean of pairwise F_{ST} ($\overline{F}_{ST}$) for each population

ranged from 0.02 to 0.32 (mean: 0.14 ± 0.07). Effective population size (N_e) ranged from 1.0 (CI: 0.9 -1.1) to 141.2 (CI: 23.4-362.4) (Table 2), and there was no trend in N_e over time across all populations and years ($p = 0.25$, marginal $R^2 = 0.034$) or between years for ponds with multiple years of sampling (W=14.5, $p = 0.073$). Observed and expected heterozygosity, allelic richness and π were significantly different among metapopulations (Figure 1.2B-1.2E) while π, N_e, F_{ind}, and F_{IS} were not. Attributes of the neighborhoods were generally better predicators of all diversity indices than individual pond attributes. We report all best fit models and their parameters in Table 1.3 and illustrate two example relationships in Figure 1.1B and 1.1C. For each of $\overline{H}_S$ (Figure 1.1B), $\overline{H}_o$, $\bar{A}_R$, and π (Figure 1.1C), the best-fit models included significant positive parameter estimates for the contribution of the number of ponds in the 2200m neighborhood, the proportion of natural ponds in the 2200m neighborhood, and the number of ponds in the 667m neighborhood, but negative parameter estimates for the interaction between the number of ponds and proportion of natural ponds in the 2200m neighborhood. Natural ponds had an additional positive effect on $\overline{H}_o$ and π. Pond area positively correlated with allelic richness and π. Connectivity to natural ponds had an additional positive effect on allelic richness. No variables were significant predictors of *Ne*, F_{IS} or $\overline{F}_{ind}$.

Population structure

Bayesian clustering across the DPS identified genetic clusters that are largely concordant with geography and previously identified metapopulation management units. The number of model components used to fit the data ranged from $k = 8$ to $k = 9$, with $k = 9$ the most common (8/10 replicates, Figure 1.2A). A general trend of northern (Eastern Santa Maria, Western Santa Maria and Western Los Alamos) and southern (Eastern Los Alamos, La Purisima and Santa Rita)

clusters of metapopulations emerged, with varying levels of admixture within them. When we re-ran FASTSTRUCTURE for the southern group, the number of model components used to fit the data were between $k = 3$ and $k = 4$ (Figure 1.2A). At $k = 3$ and $k = 4$, there are two primary clusters into which most samples fall (or an admixture of the two) and one or two (in $k = 3$ and $k = 4$, respectively) genetic clusters which are comprised of 3-4 samples. For the northern group (West Santa Maria, East Santa Maria and West Los Alamos), replicate runs for $k = 2$ failed to converge in the initial set of 10 replicates and in 40 additional replicates. After excluding these failed runs, the number of models components used to fit the data was always $k = 1$, suggesting very little structure in that region. Examination of the admixture proportions of additional k values reveals very little to no admixture at any level of k (Figure 1.2A).

Principal components analysis reflects similar, but not identical, population structure to FASTSTRUCTURE results. PC1 (31.9%) distinguishes a northern cluster (West and East Santa Maria) from a southern cluster (La Purisima and Santa Rita) and splits West Los Alamos localities into a north and south group (Figure 1.2B-D). West Los Alamos is isolated in PC2 (5.4%) and East and West Santa Maria are isolated PC3 (4.8%). PC2 also isolates West and East Los Alamos from all others and PC4 (3.1%) separates northern ponds in West Los Alamos. PC5 (2.9%) primarily isolates Santa Rita from the other metapopulations. In PC space, East Los Alamos clusters more closely with the West Los Alamos populations, while it clusters with the La Purisima and Santa Rita populations in Bayesian clustering at $k = 8$ or is isolated as its own cluster at $k = 9$.

Landscape modeling

Species distribution modeling with MAXENT performed well (mean AUC = 0.922 SD 0.053, Figure 1.4A). Five layers contributed > 75% of the average SDM: isothermality (37.47%), continentality (21.79%), precipitation in the coldest quarter (7.41%), temperature annual range (5.18%), and annual potential evapotranspiration (3.87%) We report values for all tested variables in Table 2.4. The best-fit resistance surface was distance to nearest pond ("DP", 54.2% of top models, $\bar{R}^2$=0.63, Figure 2.4B) followed by DP + SDM (25.37%, $\bar{R}^2$=.64, Δ*AICc*= 4.27, Figure 1.4C) and mean annual temperature + SDM (13.5%, $\bar{R}^2$=0.593, Δ*AICc*=3.97, Figure 1.4D). SDM alone performed poorly (0.10%, $\bar{R}^2$= 0.61). We report all models within 10 *AICc* units in Table 1.5.

DISCUSSION

Effective conservation management has two primary components: documenting patterns of decline, and understanding the ecological underpinnings of landscape use so that we can conserve and restore habitat and the populations that utilize it. The Santa Barbara DPS of *A. californiense* is one of the most imperiled amphibians in North America, and understanding how genetic diversity is spread across the landscape is essential to formulating management actions. As emphasized by McCartney-Melstad et al. (2018) for another member of the tiger salamander complex, meaningful estimates of population genetic parameters of relevance to conservation are particularly difficult to obtain for species with small ranges and low levels of genetic variation. The Santa Barbara distinct population segment of CTS fulfills both, with a total geographic range of roughly 35 x 20 km, and among the lowest levels of genomic variation reported for any metazoan (Robinson et al., 2016; Romiguier et al., 2014). Ongoing field studies of one

metapopulation (Purisima) confirm that both census and effective population sizes are extremely low, and may be approaching a point where intervention is necessary to avoid inbreeding depression (Searcy and Shaffer, Unpublished).

Landscape configuration contributes to genetic diversity

Previous studies of pond breeding amphibians, including ambystomatids, have demonstrated that pond placement and landscape composition are important components of species richness and density (Peterman et al., 2014; Shulse et al., 2010; da Silva et al., 2012; Wang et al., 2011). We find similar results for genetic diversity when considering the spatial distribution of ponds on the landscape for SBCTS. When we examined potential landscape correlates of diversity, we found that the number and type (natural versus artificial) of neighboring ponds are significantly correlated with genetic diversity. Heterozygosity, allelic richness, and π of individual breeding ponds were each positively correlated with the number of ponds within the 667m neighborhood, the number of ponds within the 2200m neighborhood and the proportion of ponds in the 2200m neighborhood that were natural (e.g., Figure 1.1B). Nucleotide diversity (π, Figure 1.1C) and allelic richness were additionally positively correlated with whether the focal pond itself was natural. Taken together, our results suggest that both the number of neighboring ponds and the spatial organization of those neighbors contribute to the genetic diversity in a neighborhood. This relationship of diversity to isolation (in this case measured as the number of neighbors, both total and natural) is in line with empirical (Trumbo et al., 2013) and theoretical (Wright, 1943) work that demonstrates that isolation negatively impacts genetic diversity.

While a focus on maintaining connectivity between suitable habitat patches is not novel (Cushman, 2006), our work helps define criteria for preserving and bolstering connectivity in

CTS. The number of ponds within long (2200m) and short (667m) dispersal distances was highly correlated with diversity. This suggests that these longer dispersal events play an important role in maintaining genetic diversity on this landscape. Previous studies of pond breeding amphibians have demonstrated that the negative effect of fragmentation may increase with dispersal capability (Cushman, 2006; Gibbs, 1998; Homan et al., 2004). Therefore, pond loss, particularly within longer dispersal distances, may compound this effect in the SBCTS system. Current recovery goals for SBCTS include preservation of upland habitat up to 2000m from the pond edge (U.S. Fish and Wildlife Service, 2016); our results suggest that management at a landscape level should also place a high value on connected networks of ponds within 2200m.

In this study, measures of diversity were also highly correlated with the proportion of natural ponds in 2200m neighborhoods in addition to the total number of neighboring ponds, suggesting that regional diversity is highly reflective of surviving natural ponds. Importantly, we found no significant differences in genetic diversity between natural and constructed wetlands, a finding in line with several previous studies of other pond-breeding amphibians (Furman et al., 2016). This suggests that constructed wetlands appear to help maintain existing levels of diversity, which are in turn driven by natural wetlands. Due to the ongoing loss of naturally occurring breeding ponds, artificial ponds will undoubtedly continue to be important refuges for pond breeding amphibians (Brand and Snodgrass, 2010). While construction of new ponds as mitigation for loss of existing ponds is a common management action, our results demonstrate that the simple addition of more breeding ponds to the landscape does not compensate for the effects of lost natural ponds, and that careful consideration should be given to the placement of new ponds.

Because most of the artificial ponds in our study region are constructed stock ponds, intended for agricultural use rather than as amphibian habitat, it is likely that they do not function

optimally as breeding habitat for SBCTS. However, that SBCTS utilize these ponds at all demonstrates that they do have a role in species survival. Across species and regions, there are few consistent patterns of artificial habitat use. For example, a study of amphibian utilization of artificial breeding habitats in Idaho found that several species, including *A. macrodactylum,* successfully colonized artificial ponds while other local species, including *A. tigrinum*, did not (Monello, 1999). This same study, and many others (e.g. Brand and Snodgrass, 2010; Semlitsch et al., 2015) have linked particular habitat qualities, such as slope, hydroperiod, and the presence of emergent vegetation or fish, to utilization of potential amphibian breeding habitat. We were not able to quantify specific breeding pond quality, but it is probable that it plays an important role in the promotion or maintenance of genetic diversity. Additionally, in the SBCTS system, the relative timing of loss of natural ponds and construction of artificial ponds is unknown. However, the length of time between the loss of natural ponds and the construction of artificial is likely an important factor in the rate at which diversity erodes.

Population genetic health

Overall, we found that genetic diversity in the Santa Barbara County DPS is lower than in other ambystomatid salamanders, as well as most other taxonomic groups (McCartney-Melstad et al., 2018; Robinson et al., 2016). Population estimates for π, a measure of polymorphism across the genome, was extremely low (π = 0.00008 to 0.0003) and among the most monomorphic of other imperiled vertebrate species (Robinson et al., 2016). This puts SBCTS at risk for inbreeding depression (Frankham et al., 2017) and future work should include examination of the fitness consequences of this reduction of diversity. Genetic effective population sizes were much lower than those found in CTS populations outside of the SBCTS

range. The median *Ne* for sampled SBCTS ponds was 12.15 (interquartile range: 7.9 - 20.3), while Wang et al. (2011) found effective population sizes about twice this in the more ecologically intact Central DPS in eastern Merced County (median 29, interquartile range: 16.8 - 38.35). We also failed to find any strong pond or landscape attribute correlates of effective population size in SBCTS. Wang et. al (2011) also found a strong correlation with breeding pond size and effective population size for natural ponds in Merced County, a result that is mirrored in other ambystomatids (McCartney-Melstad et al., 2018), suggesting that large natural ponds should support larger effective population sizes and therefore higher levels of genetic diversity (Storfer et al., 2007; Wang et al., 2011). That we recovered only marginal correlation of diversity with pond attributes could reflect the loss of natural ponds in this region and thus the loss of landscape-level correlates of genetic variation that require many generations to develop. There was also no discernable trend in effective population size change through time, overall or when comparing ponds with multiple years of sampling. Similarly, we found no landscape or focal pond correlates with population or individual estimates of inbreeding, but did identify a slightly negative decrease in heterozygosity through time, suggesting that diversity might be declining overall. Regionally, breeding habitat has acquired more protection over the last 30 years, but without a concomitant rise in effective population sizes or increases in diversity, direct management actions may be necessary to support genetically healthy populations.

Range wide patterns of diversity, structure and divergence

Genetic diversity varied among metapopulations: northern metapopulations exhibited higher diversity than the southern metapopulations (Figure 1.2B-D). Given that the northern metapopulations are in the most impacted part of the range, this spatial pattern of diversity is

somewhat counter intuitive. However, our results demonstrate that the larger number of surviving natural ponds there likely contributes to higher overall diversity in the north. Cumulative potential landscape conductance is moderately higher in the East and West Santa Maria metapopulations (Figure 1.4B-D), a result consistent with our observation of higher diversity in the northern populations as higher levels of gene flow likely contribute to higher levels of diversity in this region. This disparity among regions also suggests that management actions for each metapopulation should be developed on a case-by-case basis.

We found that population genetic structure reflects the metapopulation management framework fairly well. Both Bayesian and unsupervised clustering revealed major genetic clusters largely congruent with the current metapopulation management boundaries. Landscape resistance mapping, based on genetic distance matrices between breeding ponds, primarily identified distance to freshwater lakes and ponds, followed marginally by distance to freshwater lakes and ponds + suitability, as important contributors to connectivity between ponds (Figure 2.4B-2.4C). Resistance models based solely on the species distribution model performed poorly, as did models that excluded habitat suitability, indicating that habitat utilized for migration might be more flexible than habitat requirements at breeding sites (Mateo-Sánchez et al., 2015). The geographic distance-only models also performed poorly, despite an obvious linear relationship of genetic distance to geographic distance (Figure 2.2A). Distance to freshwater lakes or ponds did not significantly contribute to the species distribution model, but was included in the top resistance model. Taken together, these results suggest that population divergence is mediated by the distribution of suitable habitat and constrained by the location of suitable breeding ponds. Together, these results underscore the importance of the maintenance of (natural) breeding ponds

and suitable terrestrial habitat that surrounds them as a key to long-term population health (Wang et al., 2011).

CONCLUSIONS

Our results demonstrate the importance of examining fine scale drivers of genetic diversity in human altered landscapes (Cushman 2006). The configuration of suitable breeding ponds in regional neighborhoods appears to be an important driver of diversity and this work offers new insights into how landscape management for this DPS might proceed. This work also underscores the need to examine landscape genetic patterns at spatial scales that are biologically relevant, rather than for localities in isolation. For amphibian populations, constructed wetlands undoubtedly play important roles in maintaining diversity in impacted populations, but a primary focus of conservation efforts should be maintaining natural habitats. However, artificial ponds, especially in our study region, are generally constructed for purposes other than amphibian breeding and may represent only marginal habitat. A better understanding of the characteristics that promote wetland and upland habitat use would aid in restoration and mitigation efforts.

ACKNOWLEDGEMENTS

We thank Cat Darst and Rachel Henry at the USFWS for their ongoing support of this research, and E. McCartney-Melstad for bioinformatic and laboratory advice and encouragement. We thank Haoran Xue and Alex Zarem for assistance in the lab, and innumerable field biologists for their contributions of samples over the last 30 years. Funding for this work was provided by a US Fish and Wildlife Service Grant to HBS. Additional support for

EMT was received from the National Science Foundation Graduate Research Fellowship and the UCLA Dissertation Year Fellowship. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

FIGURES

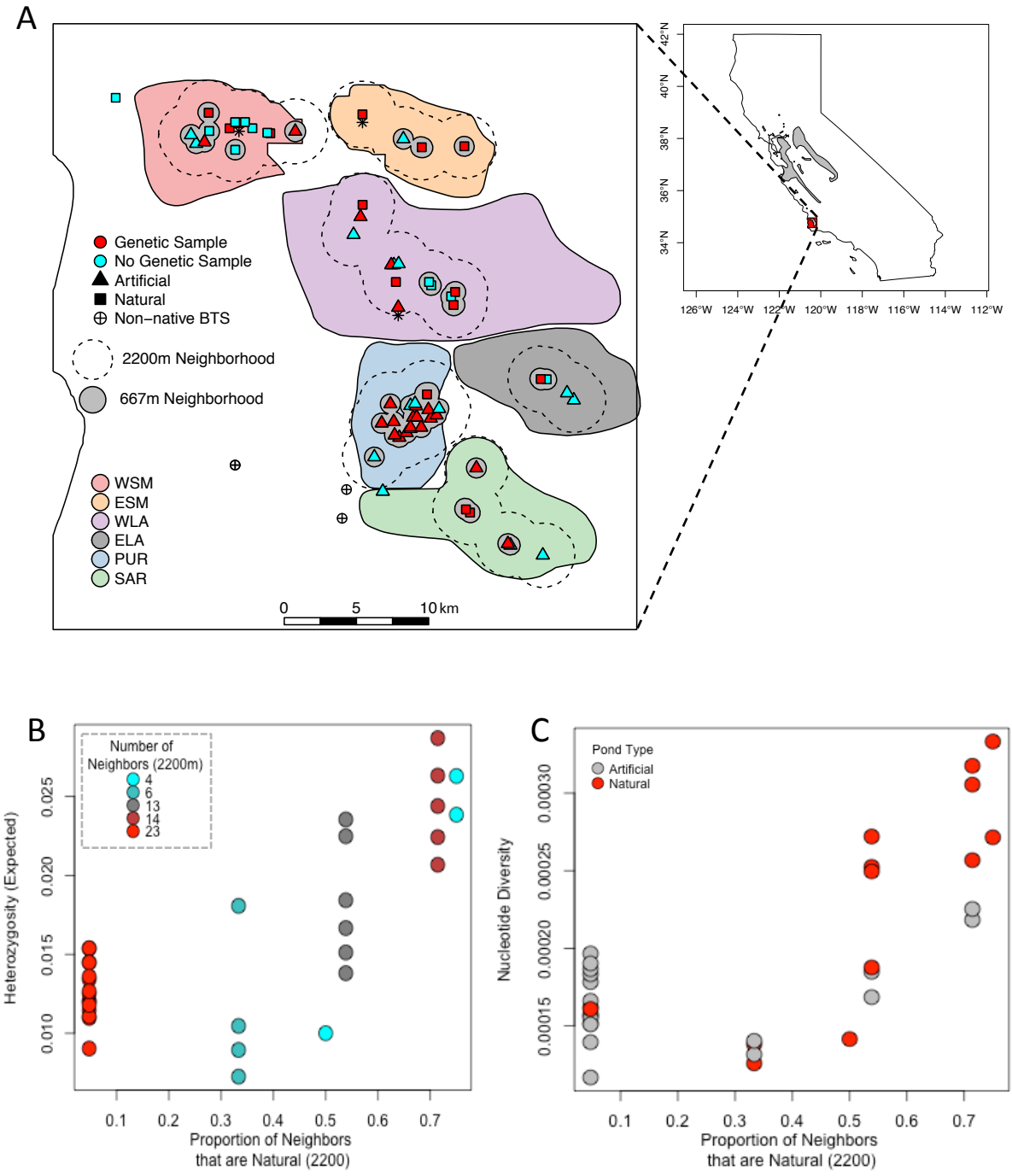
A
Genetic Sample
No Genetic Sample
Artificial
Natural
Non−native BTS
2200m Neighborhood
667m Neighborhood
WSM
ESM
WLA
ELA
PUR
SAR
0
5
10 km
42°N
40°N
38°N
36°N
34°N
126°W
124°W
122°W
120°W
118°W
116°W
114°W
112°W
B
Number of Neighbors (2200m)
4
6
13
14
23
Heterozygosity (Expected)
0.010
0.015
0.020
0.025
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Proportion of Neighbors that are Natural (2200)
C
Pond Type
Artificial
Natural
Nucleotide Diversity
0.00015
0.00020
0.00025
0.00030
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Proportion of Neighbors that are Natural (2200)

Figure 1.1. Location, type, genetic sample availability of all known breeding ponds in this region, the buffers used to generate the 667m and 2200m neighborhoods, and the presence of non-native alleles (**A**). Red points with a black asterisk (*) are ponds that had 1-2 genetic samples and as such were unsuitable for calculating population estimates of diversity, divergence and effective population size. These are not included as focal ponds in the regression analyses, but are included in Bayesian and PCA analyses. Blue points which fall outside the 667m or 2200m buffer zones are known breeding ponds that were too distant to be included in neighborhoods, but are displayed to show the full distribution of known breeding ponds in this region. Shaded grey regions in the small map of California represent the total range of *Ambystoma californiense*. Heterozygosity (*Hs*) is positively correlated with the total number of ponds in the 2200m neighborhood and the proportion of those neighbors that are natural (**B**). Nucleotide diversity is positively correlated proportion of natural neighbors in the 2200m neighborhood and the pond type of the focal pond (**C**).

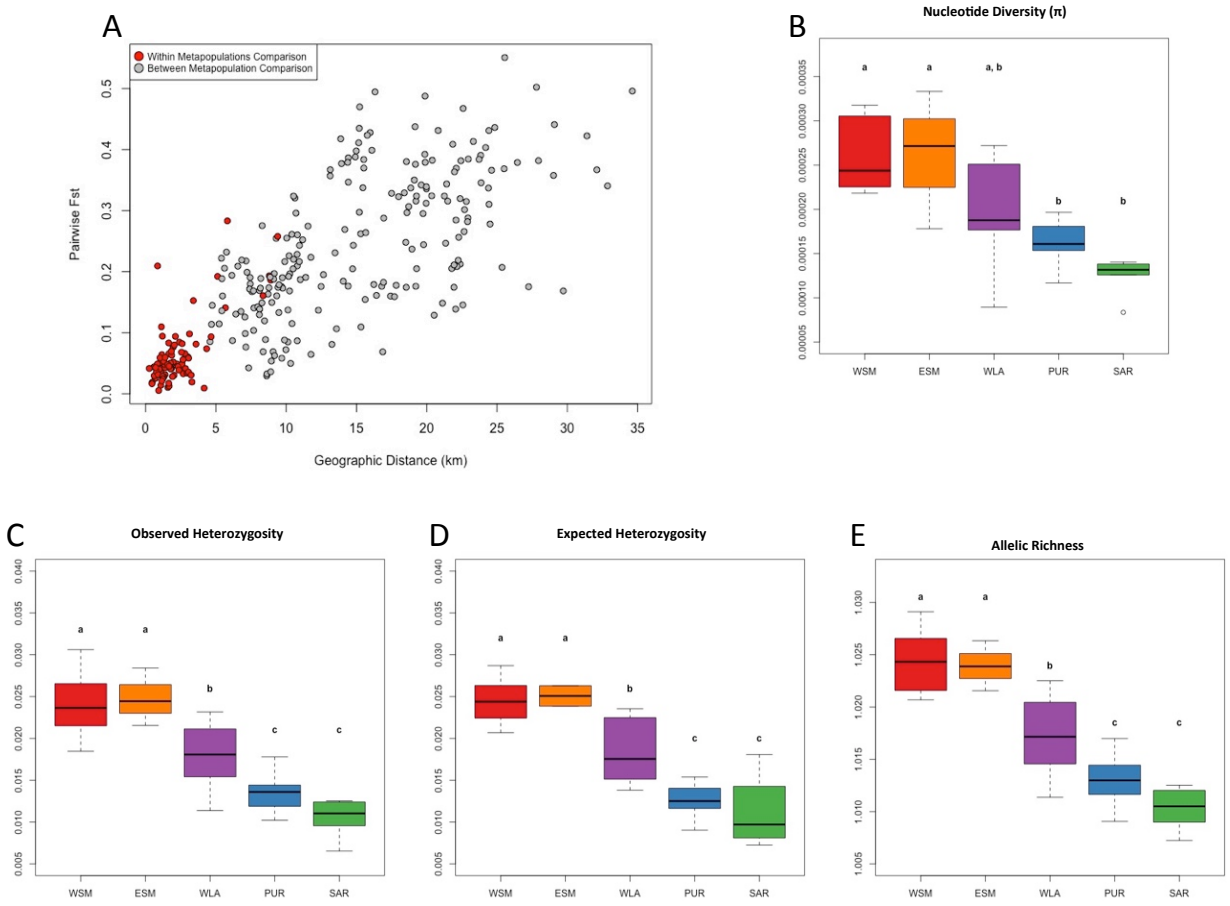

Figure 1.2. Pairwise genetic distances compared to geographic distance show a strong pattern of isolation by distance across the range (**A**). Genetic diversity is higher in northern metapopulations and significantly different between northern and southern metapopulations (**B-D**). In panels B-D, populations with the same lower case letter are not significantly different.

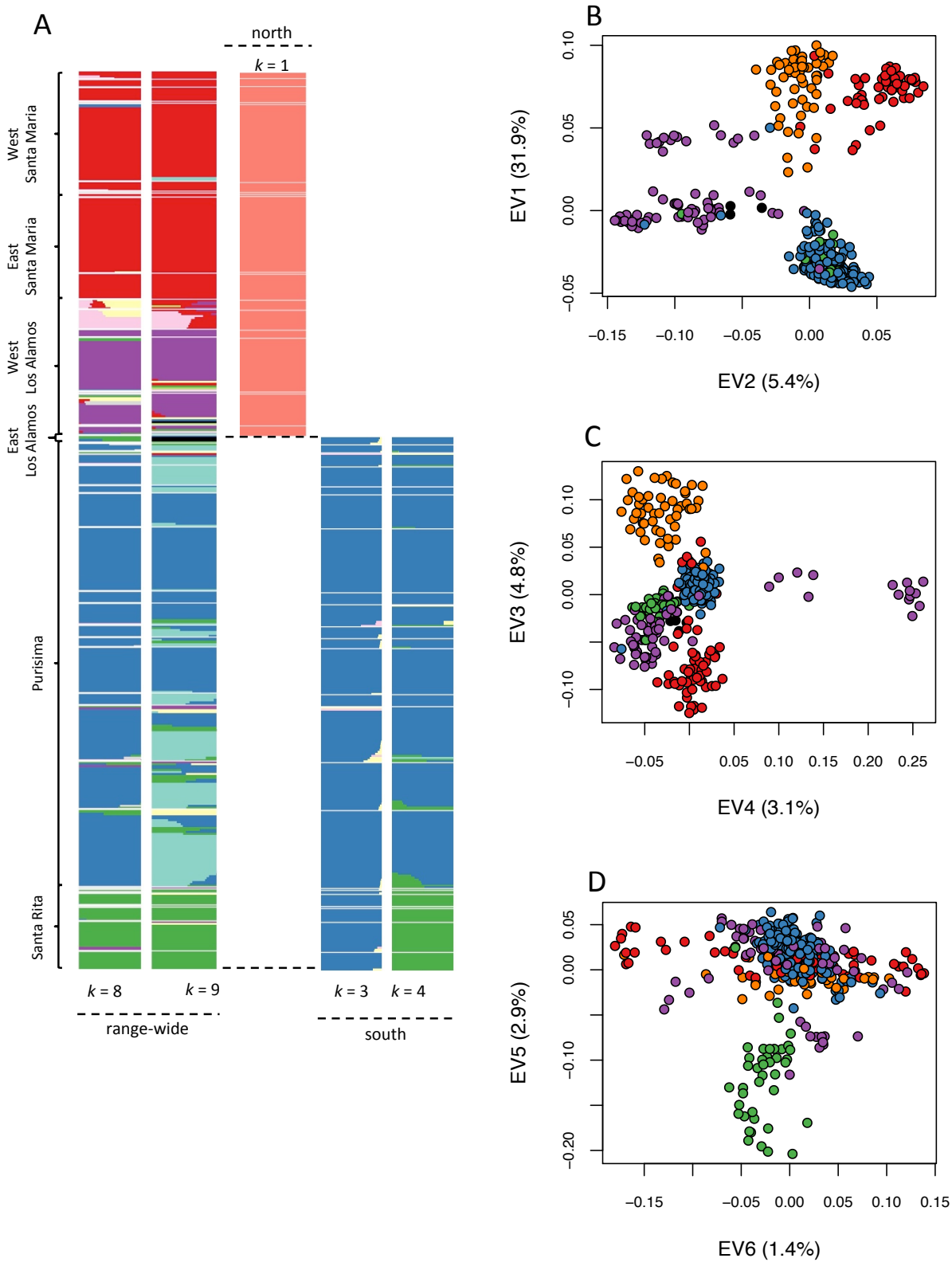
A
north
k = 1
West Santa Maria
East Santa Maria
West Los Alamos
East Los Alamos
Purisima
Santa Rita
k = 8
k = 9
k = 3
k = 4
range-wide
south
B
EV1 (31.9%)
EV2 (5.4%)
C
EV3 (4.8%)
EV4 (3.1%)
D
EV5 (2.9%)
EV6 (1.4%)

Figure 1.3. Hierarchical FASTSTRUCTURE results for all samples (range-wide), northern samples and southern samples (**A**). Across the DPS, k =8 and k = 9 were well supported, while a single population resolved for the northern three metapopulations (West and East Santa Maria and West Los Alamos) and two major genetic clusters with several individual outliers resolved in the southern group (East Los Alamos, Purisima and Santa Rita). Population clusters resulting from PCA are colored by metapopulation of origin for PC1 and PC2 (**B**), PC 3 and PC4 (**C**), and PC5 and PC6 (**D**).

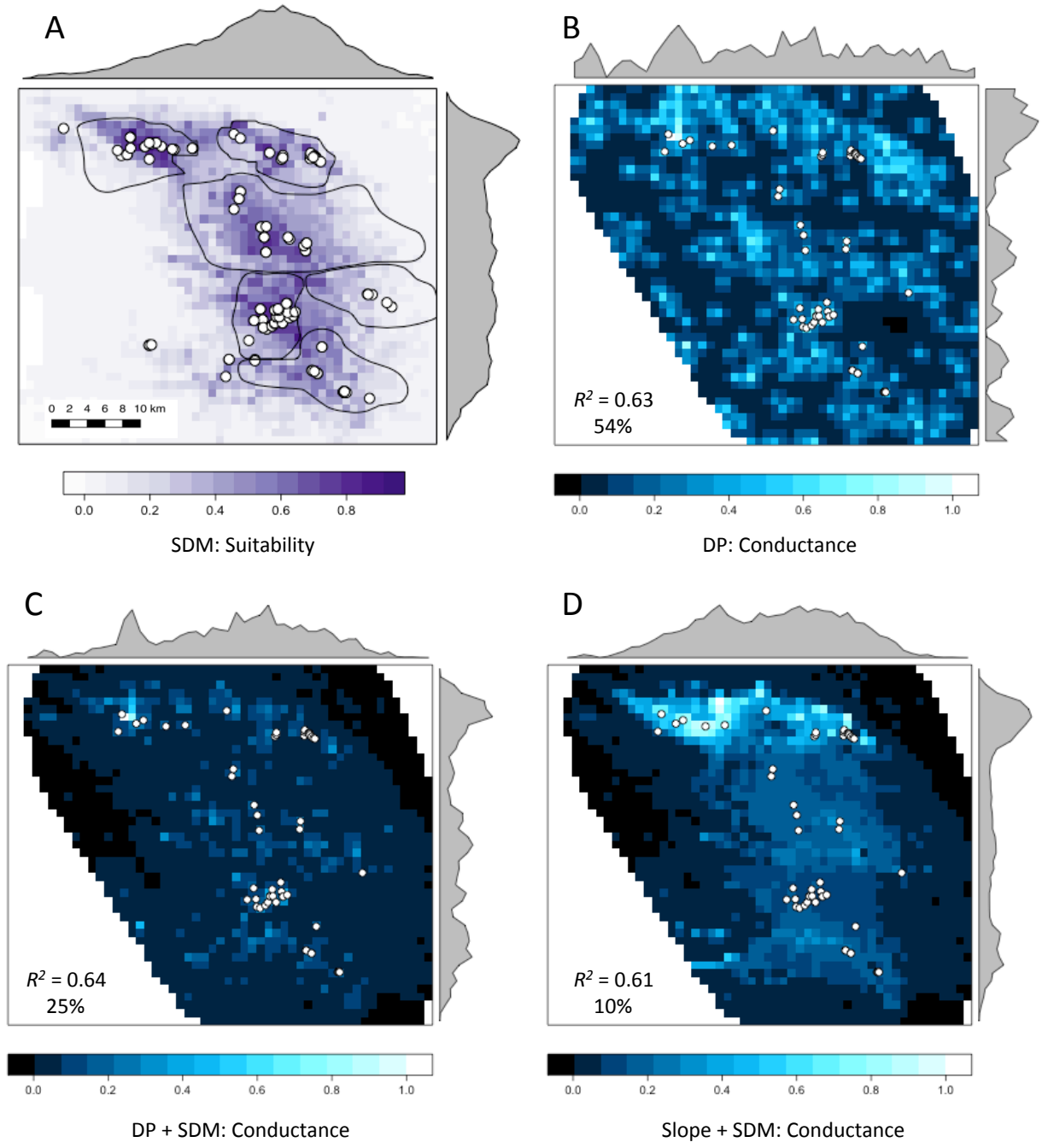

Figure 1.4: A map of the average species distribution model (**A**). The best fit resistance surface was distance to nearest freshwater lake or pond (**B**, 54.2% of top models, *R2*=0.63), and two additional composites surfaces: DP + SDM (**C**, 25.4%, *R2*=.64, Δ*AICc*= 4.27) and slope + SDM (**D**, 10.0%, *R2*=0.61, Δ*AICc*=7.07). In panels B-D, we display conductance (1/resistance) for clarity. Gray bars at the top and side of the surfaces represent the sum of the values across each column or row of cells, respectively. SDM = species distribution model, DP = Distance to nearest freshwater pond or lake.

TABLES

Table 1.1: Estimates of genetic diversity in *A. californiense* breeding ponds in Santa Barbara County.

Meta-population	*Pond*	*Year(s) sampled*	*Ho*	*Hs*	F_{IS}	A_R	π	*F (ind)*
Western Santa Maria	GUAD-2	2017	0.02	0.02	0.08	1.02	0.0002	0.14
	GUAD-3	2000	0.03	0.03	-0.07	1.03	0.0003	0.08
	SAMA-10	2001	0.02	0.02	-0.06	1.02	0.0002	0.25
	SAMA-2c	2000	0.03	*NA*	*NA*	1.03	0.0002	0.00
	SAMA-3	2000	0.02	0.02	0.02	1.02	0.0003	0.16
	SAMA-7	2001, 2017	0.03	0.03	0.08	1.03	0.0003	0.20
Eastern Santa Maria	SAMA-1	2016	0.02	*NA*	*NA*	1.02	0.0002	*NA*
	TWDA-10	1986, 1991	0.03	0.03	-0.05	1.03	0.0003	0.19
	TWDA-11*	2009, 2010, 2014, 2016, 2017	0.02	0.02	-0.03	1.02	0.0003	*NA*
West Los Alamos	ORCU-13	2017	0.02	0.02	-0.26	1.02	0.0002	0.00
	ORCU-3	2017	0.02	0.02	-0.32	1.02	0.0002	-0.17
	SISQ-1	1986, 1991	0.02	0.02	-0.03	1.02	0.0003	0.19
	SISQ-11	2009, 2017	0.01	0.01	-0.01	1.01	0.0002	0.12
	SISQ-17	2010	0.01	*NA*	*NA*	1.01	0.0001	0.00
	SISQ-3e	2001	0.02	0.02	0.16	1.02	0.0002	0.39
	SISQ-6	1986, 1991, 2000	0.02	0.02	0.06	1.02	0.0003	0.30
East Los Alamos	LOAL-18	2017	0.01	0.01	-0.17	1.01	0.0001	0.05
Purisima	LOAL-11	2009	0.01	0.01	-0.06	1.01	0.0002	0.04
	LOAL-12	2000	0.01	0.02	0.06	1.02	0.0002	0.28
	LOAL-15	2000	0.02	0.02	-0.08	1.02	0.0002	0.21
	LOAL-16	2000, 2017	0.02	0.01	-0.03	1.01	0.0002	0.26
	LOAL-17	2000, 2017, 2018	0.01	0.01	0.03	1.01	0.0002	0.16
	LOAL-32	2001	0.02	0.01	-0.14	1.01	0.0001	0.17
	LOAL-33	2001, 2017	0.01	0.01	0.00	1.01	0.0002	0.15
	LOAL-34	2001	0.01	0.01	-0.04	1.01	0.0002	0.22
	LOAL-36	2001, 2017	0.01	0.01	-0.05	1.01	0.0002	0.15
	LOAL-37	2001	0.01	0.01	0.01	1.01	0.0002	0.22
	LOAL-38	2001, 2003	0.01	0.01	-0.06	1.01	0.0002	0.14
	LOAL-39	2001, 2017	0.01	0.01	0.00	1.01	0.0002	0.17
	LOAL-47	2015	0.01	0.01	-0.08	1.01	0.0002	0.10
	LOAL-61	2003	0.01	0.01	-0.09	1.01	0.0001	0.04
	LOAL-62	2016	0.01	0.01	0.00	1.01	0.0002	0.11
Santa Rita	LOAL-2e	2008	0.01	0.01	-0.04	1.01	0.0001	0.04
	LOAL-2w*	2010	0.01	0.02	-0.15	1.01	0.0001	*NA*
	LOAL-40	2001	0.01	0.01	-0.06	1.01	0.0001	0.10
	LOAL-43*	2008	0.01	*NA*	NA	1.01	0.0001	*NA*
	LOAL-70	2008	0.01	0.01	0.15	1.01	0.0001	0.21

* includes adult samples collected nearby

Table 1.2. Molecular co-ancestry estimates of effective population size across years in A. californiense breeding ponds in Santa Barbara County.

Meta-population	*Pond*	*Year Sampled*	*Sample Size*	*Effective Population Size (range)*
Western Santa Maria	GUAD-2	2017	4	72.7 (15-175)
	GUAD-3	2000	4	21.3 (14.6-29.3)
	SAMA-10	2001	5	25.7 (15.7-38.1)
	SAMA-3	2000	7	7.8 (6.1-9.8)
	SAMA-7	2001	21	1.6 (1.3-1.8)
	SAMA-7	2017	18	9.6 (7.2-12.4)
Eastern Santa Maria	TWDA-10	1986	19	6.4 (4.6-8.4)
	TWDA-10	1991	19	10.2 (7.3-13.6)
West Los Alamos	ORCU-13	2017	5	12.2 (7.6-17.8)
	ORCU-3	2017	10	43.2 (23-69.6)
	SISQ-1	1986	13	5.2 (3.7-6.9)
	SISQ-1	1991	3	*Infinite*
	SISQ-11	2009	2	*Infinite*
	SISQ-11	2017	2	*Infinite*
	SISQ-3e	2001	5	13.9 (8.3-21)
	SISQ-6	1986	16	9.6 (6.8-12.8)
	SISQ-6	1991	7	7.8 (6-9.8)
	SISQ-6	2000	4	*Infinite*
East Los Alamos	LOAL-18	2017	4	14.8 (8.3-23.1)
Purisima	LOAL-11	2009	6	11.9 (8.5-15.8)
	LOAL-12	2000	7	0.9 (0.9-0.9)
	LOAL-15	2000	5	43.7 (15.2-87)
	LOAL-16	2000	18	8 (5.5-11)
	LOAL-16	2017	10	15.9 (8.1-26.4)
	LOAL-17	2000	5	4.5 (3.6-5.5)
	LOAL-17	2017	14	18.1 (11.5-26.2)
	LOAL-17	2018	5	19.6 (13.2-27.2)
	LOAL-32	2001	4	141.2 (23.4-362.4)
	LOAL-33	2001	14	8 (5.5-10.8)
	LOAL-33	2017	18	6.1 (4.3-8.3)
	LOAL-34	2001	6	20.7 (11.3-32.8)
	LOAL-36	2001	15	14.2 (9.9-19.1)
	LOAL-36	2017	8	8.7 (5.7-12.3)
	LOAL-37	2001	6	15.8 (9.3-24)
	LOAL-38	2001	8	3.5 (2.7-4.4)
	LOAL-38	2003	3	*Infinite*
	LOAL-39	2001	19	7 (5.1-9.2)
	LOAL-39	2017	20	15 (8.1-24)
	LOAL-47	2015	10	18.1 (12.1-25.3)
	LOAL-61	2003	17	6.8 (5-8.9)
	LOAL-62	2016	4	12.1 (8.3-16.7)
Santa Rita	LOAL-2e	2008	6	*Infinite*
	LOAL-40	2001	7	41.7 (18-75.2)
	LOAL-70	2008	9	19.1 (9.5-32)
	LOAL-70-a	2008	15	1 (0.9-1.1)

Table 1.3. Conditional averages for regression model coefficients for pond genetic metrics.

***Hs* (Gene Diversity/Expected Heterozygosity)**	Estimate	Std. Error	Adjusted SE	z value	Pr(>|z|)	
(Intercept)	0.006	0.002	0.002	3.183	**0.001**	**
Distance to Closest Pond	0.001	0.001	0.001	1.315	0.188	
Focal Pond Connectivity to All Ponds	-0.001	0.001	0.001	0.725	0.468	
Focal Pond Connectivity to Natural Ponds	0.002	0.001	0.001	1.404	0.160	
Number of Ponds (2200m)	0.004	0.002	0.002	2.641	**0.008**	**
Number of Ponds (2200m) * Prop. Natural Ponds (2200m)	-0.007	0.003	0.003	2.138	**0.032**	*
Number of Ponds (667m)	0.005	0.001	0.002	3.621	**0.00029**	***
Pond Type: Natural	0.002	0.001	0.002	1.485	0.137	
Prop. Natural Ponds (2200m)	0.026	0.006	0.006	4.300	**1.71E-05**	***
Prop. Natural Ponds (667m)	0.004	0.002	0.002	1.843	0.065	.

***Ho* (Observed Heterozygosity)**	Estimate	Std. Error	Adjusted SE	z value	Pr(>|z|)	
(Intercept)	0.006	0.001	0.001	4.016	**5.91E-05**	***
Focal Pond Connectivity to All Ponds	-0.002	0.001	0.001	1.423	0.155	
Number of Ponds (2200m)	0.007	0.001	0.001	5.783	**1.00E-08**	***
Number of Ponds (2200m) * Prop. Natural Ponds (2200m)	-0.009	0.003	0.003	3.347	**0.001**	***
Number of Ponds (667m)	0.007	0.001	0.002	4.420	**9.85E-06**	***
Pond Type: Natural	0.003	0.001	0.001	2.242	**0.025**	*
Prop. Natural Ponds (2200m)	0.026	0.006	0.006	4.352	**1.35E-05**	***

***A_R* (Allelic Richness)**	Estimate	Std. Error	Adjusted SE	z value	Pr(>|z|)	
(Intercept)	1.005	0.002	0.002	606.425	**< 2.00E-16**	***
Focal Pond Connectivity to All Ponds	-0.001	0.001	0.001	0.827	0.408	
Focal Pond Connectivity to Natural Ponds	0.002	0.001	0.001	2.359	**0.018**	*
log(Area of Focal Pond)	-0.001	0.000	0.000	2.714	**0.007**	**
Number of Ponds (2200m)	0.008	0.002	0.002	4.873	**1.10E-06**	***
Number of Ponds (2200m) * Prop. Natural Ponds (2200m)	-0.011	0.004	0.004	2.898	**0.004**	**
Number of Ponds (667m)	0.007	0.001	0.002	4.722	**2.30E-06**	***
Prop. Natural Ponds (2200m)	0.028	0.006	0.007	4.273	**1.93E-05**	***

Table 3 (Continued)

π **(Nucleotide Diversity)**	Estimate	Std. Error	Adjusted SE	z value	Pr(>|z|)	
(Intercept)	8.14E-05	1.80E-05	1.87E-05	4.35E+00	0.00001	***
Focal Pond Connectivity to All Ponds	1.70E-05	5.43E-05	5.46E-05	3.12E-01	0.755	
log(Area of Focal Pond)	-1.38E-05	3.97E-06	4.15E-06	3.33E+00	**0.001**	***
Number of Ponds (2200m)	6.73E-05	1.49E-05	1.56E-05	4.32E+00	**0.00002**	***
Number of Ponds (2200m) * Prop. Natural Ponds (2200m)	-1.03E-04	3.73E-05	3.86E-05	2.68E+00	**0.007**	**
Number of Ponds (667m)	5.74E-05	1.41E-05	1.48E-05	3.89E+00	**0.0001**	***
Pond Type: Natural	5.13E-05	2.28E-05	2.34E-05	2.19E+00	**0.028**	*
Prop. Natural Ponds (2200m)	2.68E-04	6.46E-05	6.59E-05	4.07E+00	**0.00005**	***
Prop. Natural Ponds (667m)	-4.99E-05	3.60E-05	3.70E-05	1.35E+00	0.178	

Geometric mean of Pairwise F_{ST}	Estimate	Std. Error	Adjusted SE	z value	Pr(>|z|)	
(Intercept)	0.002	0.000	0.000	7.414	<2e-16	***
Focal Pond Connectivity to Natural Ponds	-0.002	0.001	0.001	3.258	**0.001**	**
Number of Ponds (667m)	-0.001	0.000	0.000	2.767	**0.006**	**
Pond Type: Natural	0.002	0.001	0.001	3.568	**3.59E-04**	***
Prop. Natural Ponds (2200m)	-0.003	0.001	0.001	2.901	**0.004**	**
Prop. Natural Ponds (667m)	0.005	0.002	0.002	3.141	**0.002**	**

Table 1.4. Layer contributions to the species distribution model.

Layer	*Percent Contribution*	*Permutation Importance*
Isothermality	32.95	39.13
Continentality	3.59	20.48
Annual PET	9.47	6.39
Mean Diurnal Range	0.75	5.04
Precipitation of Coldest Quarter	3.19	4.31
Annual Mean Temperature	1.00	3.37
Temperature Seasonality	0.89	2.86
Aspect	2.34	2.53
Slope	5.69	2.40
Emberger Q	0.54	2.06
Precipitation of Wettest Quarter	1.96	1.71
Min Temperature Warmest Month	0.14	1.35
PET Warmest Quarter	14.29	1.15
Max Temperature of Warmest Month	2.21	1.10
Aridity	0.26	1.05
Month Count By Temperature 10	6.81	0.85
PET Seasonality	4.52	0.71
Temperature Annual Range	0.04	0.65
PET Coldest Quarter	0.04	0.55
Topographic Wetness Index	0.16	0.55
Precipitation of Driest Quarter	0.52	0.50
Precipitation of Wettest Month	0.06	0.42
Distance to Ponds and Lakes (DP)	0.91	0.34
Terrain Roughness Index	0.17	0.23
Precipitation Seasonality	0.25	0.13
Max Temperature Coldest Month	0.06	0.04
Mean Temperature of Coldest Quarter	0.04	0.03
Climatic Moisture Index	1.44	0.02
Annual Precipitation	0.09	0.02
Topographic Position Index	0.03	0.01
Elevation	0.00	0.01
PET Driest Quarter	5.45	0.00
PET Wettest Quarter	0.10	0.00
Mean Temperature of Driest Quarter	0.02	0.00
Precipitation of Driest Month	0.01	0.00
Min Temperature of Coldest Month	0.00	0.00
Mean Temperature of Warmest Quarter	0.00	0.00
Precipitation of Warmest Quarter	0.00	0.00
Mean Temperature of Wettest Quarter	0.00	0.00
Growing Deg. Days 0	0.00	0.00
Growing Deg. Days 5	0.00	0.00
Thermicity	0.00	0.00

Table 1.5. Rankings of composite and single surface models generated by *ResistanceGA* for the pairwise F_{ST} matrix. Not shown are models that were never ranked as the top model (Mean percent top replicate = 0). SDM = Species Distribution Model, DP = Distance to Nearest Freshwater Lake or Pond, CMI = Climate Moisture Index.

Surface	*Mean AICc*	*Mean R^2*	*Mean Percent Top Replicate*
DP	-262.06	0.63	54.20
DP + SDM	-257.79	0.64	25.37
Slope + SDM	-255.00	0.61	10.03
CMI + SDM	-253.31	0.60	3.93
Annual Mean Temp. + SDM	-253.29	0.60	3.70
Aspect + SDM	-250.41	0.64	2.60
SDM	-236.13	0.62	0.10
Annual Mean Temp. + DP + SDM	-244.84	0.62	0.07

REFERENCES

Allentoft, M.E., and O'Brien, J. (2010). Global Amphibian Declines, Loss of Genetic Diversity and Fitness: A Review. Diversity *2*, 47–71.

Bartoń, K. (2019). MuMIn: Multi-Model Inference.

Brand, A.B., and Snodgrass, J.W. (2010). Value of Artificial Habitats for Amphibian Reproduction in Altered Landscapes. Conserv. Biol. *24*, 295–301.

Burnham, K.P., Anderson, D.R., and Huyvaert, K.P. (2011). AIC Model Selection and Multimodel Inference in Behavioral Ecology: Some Background, Observations, and Comparisons. Behav. Ecol. Sociobiol. *65*, 23–35.

California Department of Fish and Game (CDFG), B.C. (2010). Report to the Fish and Game Commission: A Status Review of the California Tiger Salamander (*Ambystoma californiense*). Nongame Wildl. Program Rep. 2010-4 99.

Collins, J.P., and Storfer, A. (2003a). Global amphibian declines: sorting the hypotheses. Divers. Distrib. *9*, 89–98.

Curado, N., Hartel, T., and Arntzen, J.W. (2011). Amphibian pond loss as a function of landscape change – A case study over three decades in an agricultural area of northern France. Biol. Conserv. *144*, 1610–1618.

Cushman, S.A. (2006). Effects of habitat loss and fragmentation on amphibians: A review and prospectus. Biol. Conserv. *128*, 231–240.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Davidson, C., Shaffer, H.B., and Jennings, M.R. (2002). Spatial Tests of the Pesticide Drift, Habitat Destruction, UV-B, and Climate-Change Hypotheses for California Amphibian Declines. Conserv. Biol. *16*, 1588–1601.

DePristo, M.A., Banks, E., Poplin, R.E., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

DiLeo, M.F., and Wagner, H.H. (2016). A Landscape Ecologist's Agenda for Landscape Genetics. Curr. Landsc. Ecol. Rep. *1*, 115–126.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. (2014). NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour. *14*, 209–214.

Dochtermann, N.A., and Jenkins, S.H. (2011). Developing multiple hypotheses in behavioral ecology. Behav. Ecol. Sociobiol. *65*, 37–45.

Drayer, A.N., and Richter, S.C. (2016). Physical wetland characteristics influence amphibian community composition differently in constructed wetlands and natural wetlands. Ecol. Eng. *93*, 166–174.

Fitzpatrick, B.M., and Shaffer, H.B. (2007). Introduction History and Habitat Variation Explain the Landscape Genetics of Hybrid Tiger Salamanders. Ecol. Appl. *17*, 598–608.

Fitzpatrick, B.M., Johnson, J.R., Kump, D.K., Smith, J.J., Voss, S.R., and Shaffer, H.B. (2010). Rapid spread of invasive genes into a threatened native species. Proc. Natl. Acad. Sci. *107*, 3606–3610.

Frankham, R. (2005). Genetics and extinction. Biol. Conserv. *126*, 131–140.

Frankham, R., Ballou, J.D., Ralls, K., Eldridge, M., Dudash, M.R., Fenster, C.B., Lacy, R.C., and Sunnucks, P. (2017). Genetic Management of Fragmented Animal and Plant Populations (Oxford University Press).

Furman, B.L.S., Scheffers, B.R., Taylor, M., Davis, C., and Paszkowski, C.A. (2016). Limited genetic structure in a wood frog (*Lithobates sylvaticus*) population in an urban landscape inhabiting natural and constructed wetlands. Conserv. Genet. *17*, 19–30.

Gibbs, J.P. (1998). Distribution of woodland amphibians along a forest fragmentation gradient. Landsc. Ecol. *13*, 263–268.

Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. Mol. Ecol. Notes *5*, 184–186.

Grant, E.H.C., Miller, D.A.W., Schmidt, B.R., Adams, M.J., Amburgey, S.M., Chambert, T., Cruickshank, S.S., Fisher, R.N., Green, D.M., Hossack, B.R., et al. (2016). Quantitative evidence for the effects of multiple drivers on continental-scale amphibian declines. Sci. Rep. *6*, 25625.

Grueber, C.E., Nakagawa, S., Laws, R.J., and Jamieson, I.G. (2011). Multimodel inference in ecology and evolution: challenges and solutions. J. Evol. Biol. *24*, 699–711.

Hanski, I. (1994). A Practical Model of Metapopulation Dynamics. J. Anim. Ecol. *63*, 151–162.

Hanski, I. (2015). Habitat fragmentation and species richness. J. Biogeogr. *42*, 989–993.

Hanski, I., Zurita, G.A., Bellocq, M.I., and Rybicki, J. (2013). Species–fragmented area relationship. Proc. Natl. Acad. Sci. *110*, 12715–12720.

Hecnar, S.J., and M'Closkey, R.T. (1996). Regional Dynamics and the Status of Amphibians. Ecology *77*, 2091–2097.

Hijmans, R.J., Etten, J. van, Sumner, M., Cheng, J., Bevan, A., Bivand, R., Busetto, L., Canty, M., Forrest, D., Ghosh, A., et al. (2019). raster: Geographic Data Analysis and Modeling.

Hoffmann, M., Hilton-Taylor, C., Angulo, A., Böhm, M., Brooks, T.M., Butchart, S.H.M., Carpenter, K.E., Chanson, J., Collen, B., Cox, N.A., et al. (2010). The Impact of Conservation on the Status of the World's Vertebrates. Science *330*, 1503–1509.

Homan, R.N., Windmiller, B.S., and Reed, J.M. (2004). Critical Thresholds Associated with Habitat Loss for Two Vernal Pool-Breeding Amphibians. Ecol. Appl. *14*, 1547–1553.

Johnson, J.R., Thomson, R.C., Micheletti, S.J., and Shaffer, H.B. (2011). The origin of tiger salamander (*Ambystoma tigrinum*) populations in California, Oregon, and Nevada: introductions or relicts? Conserv. Genet. *12*, 355–370.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. Sci. Data *4*, 170122.

Keinath, M.C., Timoshevskiy, V.A., Timoshevskaya, N.Y., Tsonis, P.A., Voss, S.R., and Smith, J.J. (2015). Initial characterization of the large genome of the salamander *Ambystoma mexicanum* using shotgun and laser capture chromosome sequencing. Sci. Rep. *5*, 16413.

Knutson, M.G., Richardson, W.B., Reineke, D.M., Gray, B.R., Parmelee, J.R., and Weick, S.E. (2004). Agricultural Ponds Support Amphibian Populations. Ecol. Appl. *14*, 669–684.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv13033997 Q-Bio.

Manel, S., Schwartz, M.K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. *18*, 189–197.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *17*, 10–12.

Mateo-Sánchez, M.C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A., and Saura, S. (2015). Estimating effective landscape distances and movement corridors: comparison of habitat and genetic data. Ecosphere *6*, art59.

McCartney-Melstad, E., and Shaffer, H.B. (2015). Amphibian molecular ecology and how it has informed conservation. Mol. Ecol. *24*, 5084–5109.

McCartney-Melstad, E., Mount, G.G., and Shaffer, H.B. (2016). Exon capture optimization in amphibians with large genomes. Mol. Ecol. Resour. *16*, 1084–1094.

McCartney-Melstad, E., Vu, J.K., and Shaffer, H.B. (2018). Genomic data recover previously undetectable fragmentation effects in an endangered amphibian. Mol. Ecol. *27*, 4430–4443.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

McRae, B.H. (2006). Isolation by Resistance. Evolution *60*, 1551–1561.

Moilanen, A., and Nieminen, M. (2002). Simple Connectivity measures in spatial Ecology. Ecology *83*, 1131–1145.

Orloff, S.G. (2011). Movement Patterns and Migration Distances in an Upland Population of California Tiger Salamander (*Ambystoma californiense*). Herpetol. Conserv. Biol. *6*, 266–276.

Peterman, W.E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. Methods Ecol. Evol. *9*, 1638–1647.

Peterman, W.E., Anderson, T.L., Drake, D.L., Ousterhout, B.H., and Semlitsch, R.D. (2014). Maximizing pond biodiversity across the landscape: a case study of larval ambystomatid salamanders. Anim. Conserv. *17*, 275–285.

Peterman, W.E., Anderson, T.L., Ousterhout, B.H., Drake, D.L., Semlitsch, R.D., and Eggert, L.S. (2015). Differential dispersal shapes population structure and patterns of genetic differentiation in two sympatric pond breeding salamanders. Conserv. Genet. *16*, 59–69.

Phillips, S.J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography *31*, 161–175.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2019). nlme: Linear and Nonlinear Mixed Effects Models.

Porej, D., and Hetherington, T.E. (2005). Designing Wetlands for Amphibians: The Importance of Predatory Fish and Shallow Littoral Zones in Structuring of Amphibian Communities. Wetl. Ecol. Manag. *13*, 445–455.

R Core Team (2018). R: A language and environment for statistical computing.

Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. Genetics *197*, 573–589.

Riley, S.P.D., Bradley Shaffer, H., Randal Voss, S., and Fitzpatrick, B.M. (2003). Hybridization Between a Rare, Native Tiger Salamander (*Ambystoma californiense*) and Its Introduced Congener. Ecol. Appl. *13*, 1263–1275.

Robinson, J.A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B.Y., vonHoldt, B.M., Marsden, C.D., Lohmueller, K.E., and Wayne, R.K. (2016). Genomic Flatlining in the Endangered Island Fox. Curr. Biol. *26*, 1183–1189.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., et al. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature *515*, 261–263.

Ryan, M.E., Johnson, J.R., and Fitzpatrick, B.M. (2009). Invasive hybrid tiger salamander genotypes impact native amphibians. Proc. Natl. Acad. Sci. *106*, 11166–11171.

Searcy, C.A., and Shaffer, H.B. (Unpublished). Dispersal in the Santa Barbara Distinct Population Segment of the California Tiger Salamander, *Ambystoma californiense* .

Searcy, C.A., and Shaffer, H.B. (2011). Determining the migration distance of a vagile vernal pool specialist: how much land is required for conservation of California tiger salamanders? Stud. Herb. 73–87.

Searcy, C.A., and Shaffer, H.B. (2016). Do Ecological Niche Models Accurately Identify Climatic Determinants of Species Ranges? Am. Nat. *187*, 423–435.

Searcy, C.A., Gabbai-Saldate, E., and Shaffer, H.B. (2013). Microhabitat use and migration distance of an endangered grassland amphibian. Biol. Conserv. *158*, 80–87.

Searcy, C.A., Rollins, H.B., and Shaffer, H.B. (2016). Ecological equivalency as a tool for endangered species management. Ecol. Appl. *26*, 94–103.

Semlitsch, R.D. (2002). Critical Elements for Biologically Based Recovery Plans of Aquatic-Breeding Amphibians. Conserv. Biol. *16*, 619–629.

Semlitsch, R.D., Peterman, W.E., Anderson, T.L., Drake, D.L., and Ousterhout, B.H. (2015). Intermediate Pond Sizes Contain the Highest Density, Richness, and Diversity of Pond-Breeding Amphibians. PLOS ONE *10*, e0123055.

Shaffer, H.B., Pauly, G.B., Oliver, J.C., and Trenham, P.C. (2004). The molecular phylogenetics of endangerment: cryptic variation and historical phylogeography of the California tiger salamander, Ambystoma californiense. Mol. Ecol. *13*, 3033–3049.

Shaffer, H.B., Gidiş, M., McCartney-Melstad, E., Neal, K.M., Oyamaguchi, H.M., Tellez, M., and Toffelmier, E.M. (2015). Conservation Genetics and Genomics of Amphibians and Reptiles. Annu. Rev. Anim. Biosci. *3*, 113–138.

Shulse, C.D., Semlitsch, R.D., Trauth, K.M., and Williams, A.D. (2010). Influences of Design and Landscape Placement Parameters on Amphibian Abundance in Constructed Wetlands. Wetlands *30*, 915–928.

da Silva, F.R., Candeira, C.P., and de Cerqueira Rossa-Feres, D. (2012). Dependence of anuran diversity on environmental descriptors in farmland ponds. Biodivers. Conserv. *21*, 1411–1424.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L., and Waits, L.P. (2007). Putting the 'landscape' in landscape genetics. Heredity *98*, 128–142.

Stuart, S.N., Chanson, J.S., Cox, N.A., Young, B.E., Rodrigues, A.S.L., Fischman, D.L., and Waller, R.W. (2004). Status and Trends of Amphibian Declines and Extinctions Worldwide. Science *306*, 1783–1786.

Thomson, R.C., Wright, A.N., and Shaffer, H.B. (2016). California Amphibian and Reptile Species of Special Concern (University of California Press).

Title, P.O., and Bemmels, J.B. (2018). ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography *41*, 291–307.

Trenham, P.C., Koenig, W.D., and Shaffer, H.B. (2001). Spatially Autocorrelated Demography and Interpond Dispersal in the Salamander *Ambystoma californiense*. Ecology *82*, 3519–3530.

Trumbo, D.R., Spear, S.F., Baumsteiger, J., and Storfer, A. (2013). Rangewide landscape genetics of an endemic Pacific northwestern salamander. Mol. Ecol. *22*, 1250–1266.

U.S. Fish and Wildlife Service (2000a). Endangered and Threatened Wildlife and Plants; Emergency Rule To List the Santa Barbara County Distinct Population of the California Tiger Salamander as Endangered.

U.S. Fish and Wildlife Service (2000b). Endangered and threatened wildlife and plants; Final rule to list the Santa Barbara County distinct population of the California tiger salamander as endangered. Fed. Regist. *65*.

U.S. Fish and Wildlife Service (2016). Recovery plan for the Santa Barbara County Distinct Population Segment of the California tiger salamander (*Ambystoma californiense*). In U.S. Fish and Wildlife Service, Pacific Southwest Region, Ventura, California, (U.S. Fish and Wildlife Service, Pacific Southwest Region, Ventura, California), p. vi + 87 pp.

U.S. Fish and Wildlife Service (2019). National Wetlands Inventory (U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C.).

U.S. Geological Survey (2011a). NLCD 2011 Land Cover Conterminous United States (U.S. Geological Survey).

U.S. Geological Survey (2011b). NLCD2011 USFS Percent Tree Canopy (U.S. Geological Survey).

US Geological Survey (2013). USGS NED n36w121 1/3 arc-second 2013 1 x 1 degree IMG (U.S. Geological Survey).

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2002). From FastQ Data to High-

Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. In Current Protocols in Bioinformatics, (John Wiley & Sons, Inc.), p.

Vucetich, J.A., and Waite, T.A. (1999). Erosion of Heterozygosity in Fluctuating Populations. Conserv. Biol. *13*, 860–868.

Wang, I.J., Johnson, J.R., Johnson, B.B., and Shaffer, H.B. (2011). Effective population size is strongly correlated with breeding pond size in the endangered California tiger salamander, *Ambystoma californiense*. Conserv. Genet. *12*, 911–920.

Wright, S. (1943). Isolation by distance. Genetics *28*, 114.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics *28*, 3326–3328.

CHAPTER 2

GENETIC DIVERSITY AND POPULATION STRUCTURE IN A NARROW ENDEMIC, THE PANAMINT ALLIGATOR LIZARD

Erin Maurine Toffelmier

ABSTRACT

Narrow endemic species may be more susceptible to extinction or decline because the relative impacts of extrinsic (e.g. climate change or habitat loss) or intrinsic (e.g. increased drift and inbreeding leading to loss of diversity) factors may be greater than in species with greater ranges and inherently larger populations. Furthermore, these species are often poorly studied because of their limited geographic ranges, leading to a general lack of knowledge about their spatial population structure and diversity, which makes conservation management difficult. Using thousands of genome-wide single nucleotide polymorphisms and landscape genetic approaches, we explored population genetic structure and diversity in one of the least described North American squamates, the Panamint alligator lizard, *Elgaria panamintina*, a Species of Special Concern in California. The range of this species is limited to the remote desert mountain ranges of eastern California, and has been primarily documented in isolated riparian corridors. While we expected to find highly differentiated populations with limited gene flow, we found a strong general pattern of genetic isolation concordant with geographic distance and large-scale biogeographic features, and evidence to suggest on-going or recent gene flow among populations. Despite this, genetic diversity is extremely limited and population and individual

inbreeding estimates are relatively high, both of which increase extinction risk. Our work underscores the need to examine genetic patterns at fine spatial scales.

INTRODUCTION

Narrowly distributed species are often, by their very nature, targets for conservation management. Extent of occurrence, area of occupancy, population size and endemicity, are all frequently used alone or in combination as criteria for listing at most status levels on both international (e.g. IUCN RedList; IUCN, 2019) and regional (e.g. California's Amphibians and Reptiles of Special Concern; Thomson et al., 2016) management scales. Such small, isolated taxa are more susceptible to local extinctions due to extrinsic factors, including habitat loss, competition, and disease (Diamond, 1989; Sodhi et al., 2009). Coupled to these, the genetic consequences of small population size alone may push localized endemics further up the list for conservation concern and management. Reduction in genetic diversity due to inbreeding and drift may lead to loss of adaptive potential and resilience to diseases and environmental perturbations, including climate change (Frankham et al., 2017). Further, inbred populations typically exhibit higher extinction risk and island endemic populations are particularly prone to extinction associated with reduced genetic variation (Frankham, 1998). While patterns of genetic diversity within endemic species with narrow ranges vary across taxonomic groups and geographic contexts, localized endemics frequently conform to the patterns of low overall genetic diversity with highly structured populations, making them especially susceptible to drift and inbreeding (Ellstrand and Elam, 1993; Gaston, 1994; Karron, 1997; Templeton et al., 1990; Wright, 1969), although several recent studies have demonstrated high levels of diversity despite narrow ranges (Forrest et al., 2017; Mateu-Andrés and Segarra-Moragues, 2000).

In addition to population dynamics, the landscape influences genetic structure and may restrict opportunities for gene-flow (Sork and Waits, 2010; Storfer et al., 2007). This may be the case for endemics that are typically habitat specialists that utilize patchy habitats where isolation among populations is the norm (Manel et al., 2003). In order to understand population structure within narrowly distributed endemics, it is therefore important to understand how landscape and environmental features shape genetic differentiation.

The Great Basin desert, comprised of montane ranges interspersed with lowland desert, has long been a target for the study of island biogeography and phylogeography of terrestrial habitat islands (Brown, 1971; Fleishman et al., 2001; Floyd et al., 2005). As a region, the Great Basin is home to a multitude of narrow or localized endemic species. Within this complex environment, riparian habitat patches make up a tiny, yet vital, component of the desert habitat mosaic (Chambers et al., 2008; Sada et al., 2001). These specialized habitats, supported by small perennial streams sustained by local springs or snowmelt runoff, are generally small and isolated, representing habitat islands in a matrix of unsuitable or low-quality habitat (Chambers et al., 2008) (Minshall et al., 1989; Sada et al., 2001). These are areas of high biodiversity and high endemism, resulting from apparent long-term climate stability coupled with long-term isolation (Batzer and Baldwin, 2012; Chambers et al., 2008; Shepard, 1993; Stevens and Meretsky, 2008). Consequently, many of the species and/or populations present in these habitats have been identified as relictual (Hall, 1991; Stebbins, 1958). Geography and historic climate change can be isolating mechanisms: remnant populations are left behind as deserts expand and regions become drier (Riddle et al., 2014; Stevens and Meretsky, 2008). Long-term isolation, whether partial or complete, contributes to inter- and intra-specific divergence over long and short timescales, resulting in high levels of endemism in the former case and population divergence in

the latter (e.g. Houston et al., 2012; Mac et al., 1998; Sada et al., 2001; Stebbins, 1958). Conversely, there are several examples of aquatic or riparian habitat specialists which exhibit a broad pattern of isolation by distance and gene-flow among habitat patches (e.g. Lawlor, 1998; Wang, 2009).

In the Southwestern Great Basin Desert, the Panamint alligator lizard (*Elgaria panamintina*) is narrowly distributed in isolated riparian corridors within the White, Inyo, Nelson, Argus, and Panamint ranges (Stebbins, 1958, 2003; Thomson et al., 2016). The species is one of the least studied North American squamates. Since its first documented observation in 1954 and subsequent description in 1958, very little has been published on its life history or ecology (Clause et al., 2018; Stebbins, 1958; Thomson et al., 2016). Much of our understanding of its habitat use and physiological requirements is based on a handful of observations or drawn from parallels with the closely related *Elgaria multicarinata* (Mahrdt and Beaman, 2002; Stebbins, 1958; Yasuda, 2015). The species is thought to have low population densities, though it is also described frequently as "secretive" and can be nocturnal; both underscore the difficulty in collecting field data. Although it is characterized as restricted to riparian woodlands, individuals have also been found in desert scrub and Joshua tree habitat adjacent to riparian corridors, suggesting that it may be less of a habitat specialist than normally assumed. Very little beyond anecdotal evidence is known about its life history and habitat use (Clause et al., 2018; Mahrdt and Beaman, 2002; Stebbins, 1958, 2003). Due to its extremely limited and remote range, the Panamint alligator lizard has been listed as a California Species of Special Concern for the last two assessments (Jennings and Hayes, 1994; Thomson et al., 2016), and is, as of 2019, under consideration for listing as Threatened under the Federal Endangered Species Act (Adkins Giese et al., 2012).

Examining the genetic consequences of small populations has become a major tool in the species conservation framework and its implementation has grown substantially with the advent of recent DNA sequencing technologies. Additionally, advances in GIS and remote sensing technologies provide new and higher resolution data layers from which to glean information about species distributions and connectivity. Together, these approaches can contribute a wealth of information about species that are difficult to work on in the field, including *E. panamintina*. In the present study, we examine population genetic structure using thousands of genetic markers applied to the largest collection of *E. panamintina* ever assembled. We combined population genetic analysis with environmental niche and landscape resistance modeling to examine connectivity across the range, and use these data to evaluate several key issues in the molecular ecology and conservation of the species. First, given its restricted range and apparent reliance on rare, isolated habitat patches, we predicted that populations would be small and highly inbred. We further predicted that extant populations would be highly structured, and that structure would be driven by habitat-based barriers to gene flow. Taken together, these results are relevant to the impending conservation and management of the Panamint alligator lizard, and we discuss our results in the context of the conservation of this species.

MATERIALS AND METHODS

Sample collection, genotyping and data filtering

We obtained 51 tissue samples from 14 localities from across the range of *E. panamintina* (Table 2.S1). These samples consist of a combination of new tissues collected in 2015-2017, and historical museum collections (Figure 2.6B). These comprise, to the best of our knowledge, all available tissues that are suitable for genomic DNA sequencing and represent

almost all known localities, and the largest data set ever assembled for *E. panamintina*. Because these samples are somewhat geographically clustered, we refer to our sampling as: **WM** (White Mountains; *N*=40 from 9 localities), **IM** (Inyo Mountains; *N*=8 from four localities), and **SM** (southern mountains, representing three poorly sampled ranges: Panamint Mountains, *N*=1; Argus Mountain, *N*=1; and Nelson Mountains, *N*=1, each from a single locality).

We extracted genomic DNA using a salt extraction method following Green and Sambrook et al. (2012). We then generated reduced representation genomic libraries for each individual using the 3RAD protocol of Glenn et al. (2017). Briefly, high genomic weight DNA is digested with three restriction enzymes (*SphI*, *MspI* and *ClaI*; NEB; Beverly, MA, USA). We selected genomic fragments with the *MspI* cut site on the 5' end and the SphI cut site on the 3' end with 3RAD stubs designed to match the sticky end resulting from restriction enzyme digestion. The third enzyme, *ClaI*, has a complementary sticky end to *MspI* and is included to help reduce adapter dimers. Additionally, each stub contained a unique 5-8bp internal barcode to help identify individuals. Illumina-compatible libraries were completed using iTru adapaters (Glenn et al., 2019) with dual 8-bp indices using KAPA LTP reactions (KAPA Biosystems). This protocol yields quadruple indexed libraries (one internal barcode in each of the two 3RAD stubs, one external barcode in each of the two Illumina adapters). All samples were then pooled into a single equimolar pool (we also included 19 samples from *Elgaria multicarinata* and *Elgaria coerulea* for other projects) and size selected on a Pippin Prep (Sage Science, Inc.) in a window size of 350-450 base pairs. This pool was sequenced on an Illumina HiSeq 4000 (Illumina, Inc.) with 150bp paired-end sequencing.

Raw sequence reads were de-multiplexed based on the external Illumina adapters. Raw reads trimmed of the 3RAD stub sequences with CUTADAPT v. 1.12 (Martin, 2011) and read

pairs were discarded if they did not contain the correct internal 3RAD barcodes. Reads were also trimmed of low quality bases and 20 base pairs were removed from both the 5' and 3' end of both R1 and R2 reads. Sequences from each sample were clustered and genotypes were called across samples using IPYRAD 0.7.28 (Eaton, 2014). We used a clustering threshold of 0.94 for within and between sample clustering, and retained clusters with minimum read depth of 6 and maximum read depth of 10000. Within group loci were retained in IPYRAD with the following parameters: max 8 heterozygous sites/locus, max 0.5 shared heterozygous sites per locus, max 5 Ns (uncalled bases), max 10 SNPs, max 10 insertions or deletions ("in/del"), locus must be present in 6 or more individuals. We then excluded in/del variants to retain only SNPs, and removed SNPs with more than one alternate allele and those with greater than 50% missing data in VCFTOOLS v0.1.15 (Danecek et al., 2011).

Population genetics

Genetic diversity: We conducted a Hardy Weinberg Equilibrium test for each variant for in VCFTOOLS, removed loci containing an excess of heterozygotes ($p < 0.001$; a potential indicator of paralogous loci), and then randomly selected one SNP per locus using custom perl scripts to limit linkage disequilibrium between markers. From this set of SNPs, we calculated several genetic diversity metrics. For all individuals, we calculated individual inbreeding (F) for all loci in VCFTOOLS, then averaged among individuals within populations to obtain population mean individual F. For localities with $n > 2$ samples, we also calculated average observed heterozygosity (H_o), average gene diversity (H_S), population inbreeding (F_{IS}), and allelic richness (A_R) as the average of per-locus estimates within each populations. Per-locus estimates were calculated in the R package *Adegenet* v2.1.1 (Jombart and Ahmed, 2011). To assess total genetic

diversity, we calculated nucleotide diversity per site, π using *PopGenome* v2.6.1 in R (Pfeifer et al., 2014) which follows Nei's calculation (Nei, 1987). We excluded three samples that had missingness >12% and calculated π for all sequenced loci with no missing data across the 48 remaining samples. We excluded missing loci to avoid bias due to null alleles and allelic dropout, and calculated average π for each locality separately and also for all locality pairs to assess patterns of both intra- and inter-population diversity.

Effective population size: We calculated effective population (N_e) sizes for all localities with $N >$ 6 (Piute Creek, Coldwater Canyon, Silver Canyon and Toll House Spring, all in the White Mountains group). We excluded singletons and private doubletons, retained only loci with no missing data, and then randomly selected one SNP per RAD locus to limit physical linkage disequilibrium between loci. We used NEESTIMATOR (Do et al., 2014) to estimate effective population using the single sample linkage disequilibrium (LD, (Waples and Do, 2008)) method for each replicate data set. The LD method has been demonstrated to be the most robust estimator when sample size, N, is much smaller than actual N_e (Wang, 2016). We replicated this procedure of randomly selecting one SNP per locus and recalculating N_e 30 times for each locality to avoid biases that any one random set of SNPs might yield. We report the mean of these estimates for the N_e and the mean of the lower and upper bound of the confidence intervals.

Population structure: For analyses of population structure, we excluded singleton and private doubleton SNPs, which has been demonstrated to increase the accuracy of population inference (Linck and Battey, 2017), and excluded SNPs with greater than 50% missingness across all samples. To visualize and describe total variation among individuals, we used principal

components analysis (PCA) in the R package *SNPRelate* v1.14.0 (Zheng et al., 2012). Several localities consist of a single sample and are also geographically isolated (particularly the IM and SM groups, see Figure 2.1D). Range-wide PCA revealed a strong signal in these single sample/geographically distant samples so we ran three additional PCAs to examine the impact of uneven sampling: 1) PCA with all localities randomly subsampled for up to two individuals (retaining localities with only one sample) to examine sample size effects, repeated 10 times to assess the impact of random sampling individuals, 2) PCA including all White Mountains and Inyo Mountains samples (excluding Southern Mountains samples) to examine the effects of the geographic outlier samples and 3) PCA of the White Mountains samples only to examine fine-scale structure across the region with the best sampling.

We examined population structure in a Bayesian framework with the software FASTSTRUCTURE V1.0 (Raj et al., 2014). We ran 10 replicates with the logistic model for values of k =1 to k =20 to identify the number of populations and admixture in the range-wide group. We then re-ran FASTSTRUCTURE with 10 replicates each for k =1 to k =10 for the White Mountains group to examine more fine-scale structure. For each subset (range-wide and White Mountains), we used the *ChooseK.py* function to calculate the marginal likelihoods, and random seeds with the highest marginal likelihoods are presented here. We inferred an unrooted phylogentic tree for all samples based on a concatenated data set comprised of all sequenced bases for loci present in greater than 50% of samples in RAXML 8.2.4 using the GTRGAMMA model of rate heterogeneity (Leaché et al., 2015; Stamatakis, 2014). Additionally, we used TREEMIX v1.13 (Pickrell and Pritchard, 2012) to examine putative migration events among all localities. TREEMIX was run for 10 random number seeds for each of 0 to 10 migration edges. After selecting an optimal migration edge, we ran 100 replicates of this tree with the optimal

migration edge to generate a consensus tree. Because several localities consist of a single sample, we ran TREEMIX with sample size correction turned off in all cases (Pickrell and Pritchard, 2012).

We calculated global F_{ST} and pairwise population F_{ST} based on localities with $N > 2$ (Willing et al., 2010). We carried out a hierarchical analysis of molecular variation (AMOVA) in the R package *Poppr* v2.8.2, first partitioning variation among individual washes and the three major geographic groupings (WM, IM, SM), then examining variation within the WM and IM groups separately. The SM group consists of three samples, so we did not conduct an AMOVA on this group. For all AMOVAs, we assessed significance with a permutation test with 10,000 repetitions.

Landscape analyses

We also examined population structure in a landscape framework. We first used species distribution modeling (SDM) to visualize potential habitat. We then examined the relative contributions of geographic isolation and landscape resistance to genetic divergence across the range and on a fine scale in the White Mountains Region.

Species distribution models: We first developed a distribution probability based on an ecological niche model for *E. panamintina* using 41 climate and landscape variables (Table 2.1) with a resolution of ~10m. We used a presence only model implemented in MAXENT 3.4.1 because observations for this species are extremely limited and its distribution likely extends to regions that lack historical survey records (Phillips and Dudík, 2008). MAXENT has been shown to be effective for predicting environmental niche even with a small number of occurrence records

(Pearson et al., 2007). For presence points, we compiled 84 observations derived from VertNet.org records and our own sampling records. Together, these comprise virtually all geo-located sampling records for this species. To increase model accuracy, we used a target background selection approach, limiting background points to localities in the region for which observations of any reptile or amphibian species have been made (N = 7,067 points) (Merow et al., 2013; Phillips and Dudík, 2008; Searcy and Shaffer, 2016). We ran 10 replicate models and report an average among all replicates. We evaluate layer rankings by permutation importance, which has been demonstrated to more accurately reflect biologically relevant responses (Searcy and Shaffer, 2016).

Isolation by distance: We examined the relative contribution of geographic distance to genetic distance in two ways. First, we examined the relationship between Euclidean geographic distance and genetic distance between individuals with Mantel correlograms (Legendre and Fortin, 1989, 2010) implemented in *ecodist* v2.0.1 in R. We generated a 95% confidence interval with 100,000 bootstrap iterations and set breaks so that the number of sample pairs within each class was roughly equal (Diniz-Filho et al., 2013). Genetic distance and geographic distance matrices were calculated in the R packages *poppr* v2.8.2 and *fields* v9.7, respectively. We further visualized the relationship between Euclidean geographic and genetic distance with a scatterplot. Second, we used the R package *conStruct* v1.0.3 to explore individual ancestry proportions in a spatially explicit framework to isolate the geographic contribution to apparent co-ancestry (Bradburd et al., 2018). We ran *conStruct* with individuals as the sampling unit for both the spatial and non-spatial model for k = 1 to k =10. We used a cross-validation procedure with 10 replicates implemented in *conStruct* and examined predictive accuracy to determine the best model. To

avoid potential pitfalls of over-fitting the model, we also examined the contribution of all population layers to determine a biologically meaningful number of layers. We repeated the *conStruct* analysis for the White Mountains group separately to examine isolation by distance on a fine scale with *k* values 1 to 10.

Isolation by resistance: We used the R package *ResistanceGA* v4.0-10 to generate resistance surfaces to model functional connectivity based on genetic distances between localities (Peterman, 2018). The *ResistanceGA* approach generates surfaces parameterized by their concordance with genetic distances, rather than expert opinion or other parameterization schemes, which should more accurately reflect biological reality. Because the spatial distribution of our samples on the landscape is biased towards the White Mountains region, we examined resistance surfaces in the White Mountains only. Because calculation of resistance surfaces is computationally intensive, we down-sampled all layers to a resolution of ~30 meters and cropped them to the White Mountains Region. For samples with identical locality coordinates (or fell within the same 30x30m raster cell), we randomly selected one sample from that locality. For the remaining samples, we selected SNPs present in the White Mountains populations with less than 20% missingness (2097 SNPs), and calculated Nei's genetic distance (*D*) for all sample pairs in *hierfstat*. We calculated landscape distance based on commute distance, which is functionally equivalent to the resistance distance of CIRCUITSCAPE (Kivimäki et al., 2014; Marrotte and Bowman, 2017). We first singly optimized each of the 42 layers (41 climate and landscape variables plus the composite SDM). We then selected the top 10 layers with lowest *AICc* and checked for correlation within this set. From these, we selected six uncorrelated layers (VIF < 10, calculated in the R package *usdm* v1.1-18 and included the SDM, resulting in 7 total layers

(Table 2.1, Table 2.2). We optimized combinations of 2, 3 or 4 layers to evaluate the performance of composite surfaces relative to single surfaces, and included a simple model of Euclidean distance and a null model. We examined *ΔAICc*, and the variance components (marginal R^2 of the model) to select the best-fit surfaces. We also assessed the relative support of each surface via a pseudo bootstrap procedure in which we subsampled individuals and distance matrices without replacement to assess the sensitivity of the model to individual sampling points.

RESULTS

Genotyping and data filtration

Sequencing resulted in an average of 4,570,382 (SD 1,302,980) base pairs per individual. Following clustering and basic filters, the average number of clusters retained per individual was 17,633.9 (SD 2617.2) with mean depth of 15.55 (SD 10.75). This yielded 90,782 total clusters with average per sample missingness of 69% (0.00-98%).

Population genetics

Genetic diversity and effective population size: After filtering, we retained 10,726 SNPs (one SNP per locus). All population genetic values are reported in Table 2.3. Levels of genetic diversity were generally low to moderate across the range (mean population H_O: 0.12 ± 0.03; H_S: 0.14 ±0.03; A_R: 1.13 ±0.03), with moderate levels of inbreeding (population F_{IS}: 0.12 ±0.10; individual *F*: 0.41 ±0.14). To calculate π, we retained 3,589 loci comprised of 797,386 bases with only 748 variable sites for 48 samples (we excluded three samples total; one locality, Black Canyon (BCa), was excluded because it consisted of a single sample with high missingness). Total nucleotide diversity is extremely low (mean pairwise π: 8.70e-05, ± 4.06e-05, mean

population π: 2.36e-05 ± 2.35e-05). Mean estimates of effective population sizes (N_e) for Piute Creek (PCr), Coldwater Canyon (CCa), Silver Canyon (SiCa), and Tollhouse Spring (THSp) were 321.43 (LB:77.15, UB:infinite), 54.48 (LB:33.8, UB:139.14), 15.02 (LB:13.83, UB:16.38) and 18.45 (LB:16.46, UB:20.89), respectively.

Population structure: Multiple clustering methods reveal a strong signal of geography, reflecting a general pattern of differentiation among mountain ranges. The PCA of all samples clearly separates all major geographic groupings (Figure 2.1C). Principal components (PC) 1 through 4 isolate the White Mountains (WM) from the Inyo Mountains (IM) and from southern mountain ranges (SM). This pattern remained when localities were down-sampled to 1-2 individuals and PC contribution remains stable (Figure 2.2), indicating that the single sample localities do not bias this pattern. When the SM samples (which are outliers in the range-wide PCA) are withheld from the analysis, geography resolves in PC1 and PC2, isolating WM from IM populations (Figure 2.1B). Within the WM group, PC1 and PC2 also reveal a geographic gradient and separates individual localities (Figure 2.1C).

FASTSTRUCTURE for the range-wide group identified three primary groups in the total data set. FASTSTRUCTURE was generally consistent among replicates, with *k=3* model components explaining structure in the data (Figure 2.3). Marginal likelihoods modestly increased after *k*=2, and the maximum marginal likelihood was always for the largest value of *k* tested, though admixture proportions for additional *k*s are always less than 5.0×10^{-5}. For *k=3*, clusters assignments are concordant with geographic groupings. Hierarchical clustering in FASTSTRUCTURE reveals further substructure within the White Mountains group (Figure 2.3B).

Within the WM subgroup, $k = 2$ was the number of groups that explained structure in the data, while $k = 10$ (the highest number of tested groups) maximized the marginal likelihood.

The maximum likelihood tree resolves four major geographically cohesive groups with strong bootstrap support (Figure 2.4). The deepest split separates the White Mountains from the Inyo + southern mountains. A southerly group consisting of the Inyo Mountains and Southern Mountains is sister to the White Mountains. Within this group, Barrel Spring (BSp), which is the northern-most locality of the Inyo Mountains group, is sister to a group comprised of the remaining Inyo Mountains localities and the southern mountain localities.

Pairwise population F_{ST} values were moderate to high across localities, ranging from 0.08-0.15 among White Mountains population pairs and 0.18-0.30 between populations from different mountain ranges (Table 2.3). Global F_{STS} were 0.13 and 0.10 for WM and IM, respectively (and could not be calculated for SM due to low sample size). AMOVA indicates that the major source of variation (48.75%) in the total group is distributed among the major geographic clusters, with a further 37.34% partitioned among all localities within regions (Table 2.4). When the WM and IM groups were considered independently, variation within localities was 2-3 times greater than variation between localities.

Among all TREEMIX runs, major geographic clusters were consistently resolved and almost perfectly reflect geographic relationships (Figure 2.5). For example, Tollhouse Spring and the Narrows form a monophyletic group separate from the rest of the White Mountains populations, while Payson Canyon, is sister to the rest of the White Mountains. Similarly, Barrel Spring is sister to all other Inyo mountains populations plus the southern clade, while the southern clade forms a monophyletic group nested within the Inyo group. For all models with migration edges (i.e. $m > 0$), TREEMIX identified potential migration events between the

southern-most populations of the White Mountains clade and southern mountains clade, but not between the Inyo Mountains clade and anywhere else. Model likelihoods increased with each additional migration edge added, increasing sharply until five edges (Figure 2.5) were added and minimally thereafter.

Landscape analyses

Species distribution models: MAXENT models performed well (AUC of the training data = 0.993) and identified topographic roughness, seasonality of potential evapotranspiration, and mean diurnal temperature range as the most important drivers in the model based on permutation importance (Figure 2.6B, Table 2.1).

Isolation by distance and isolation by resistance: Simple tests of isolation by distance show positive correlation of genetic distances within short geographic distance (<25km), which correspond to populations within major geographic units (Figure 2.7A). At larger geographic distances, correlation with genetic distance becomes strongly negative, indicating genetic dissimilarity. Similarly, a scatter plot of the relationship between geographic distance and genetic distance reveals a relatively linear relationship (Figure 2.7B. However, pairwise genetic distances are higher among IM localities than among WM localities over similar geographic distances (Figure 2.7B). Results of the CONSTRUCT analysis demonstrate that geographic distance does contribute significantly to population structuring for all values of k (Figure 2.7C). When accounting for spatial distance, the contribution of additional layers beyond k=1 is minimal both on the range-wide and White Mountains scale (range-wide: Figure 2.7D). Cross validation also

supports higher predictive accuracy for all k-layers when comparing the spatial and non-spatial models.

For the White Mountains region, the best multi-surface optimization was Bio12+SDM+Continentality (Figure 2.6A), followed by Bio12+Continentality and Bio12+Bio09+Continentality (Table 2.2; SDM = species distribution model, Bio12 = annual precipitation, Bio09 = mean temperature of driest quarter). Bio12 was included in seven of the top 10 models. The distance only model was consistently ranked lowest of all single-surface and multi-surface models.

DISCUSSION

For narrow endemic species, understanding how populations are structured is a critical component for adequate conservation management (Frankham et al., 2017). Here, we examine population structure across the range of *E. panamintina* to help elucidate current population structure and connectivity. We found that population genetic structure strongly reflects geography and isolation by distance, genetic diversity and heterozygosity are both extremely low, and levels of individual and population inbreeding tend to be high. From a management perspective, these results are implicated in extinction risk (Frankham, 1995, 1998, 2005; Spielman et al., 2004). Several populations of *E. panamintina* exhibit elevated inbreeding coefficients on par with other species with demonstrated declines (Frankham, 1995). Isolated species and populations tend to face higher extinction risks, which is exacerbated by endemism (Frankham, 1998). Habitat specialization and limited geographic range, two hallmarks of *E. panamintina*, have also been linked to elevated extinction risk (Tingley et al., 2013). Although we find that while populations are small with extremely low genetic diversity, individual habitat

patches were not completely genetically isolated. This is somewhat surprising given the narrative surrounding the natural history of this species. Several instances of *E. panamintina* individuals have been documented well outside the riparian corridors suggesting at least the capacity to travel beyond what is traditionally thought of as suitable habitat (Hall, 1991). Our resistance surfaces based on genetic distances corroborate these observations and suggest that gene flow may occur between isolated riparian patches.

Diversity and population structure in *Elgaria panamintina*

In this study, we used a large genomic data set to examine patterns of genetic differentiation in a threatened, range-restricted species. Our results reveal a species with very low overall genetic diversity and population structure largely concordant with major geographic units. A rough comparison with other species for which similar data are available underscores the limited pool of genetic diversity in *E. panamintina,* which shows orders of magnitude less diversity (π), lower heterozygosity and higher levels of inbreeding, than similarly range-restricted taxa (Table 2.5). In fact, our data puts the Panamint alligator lizard on par with the island fox, (*Urocyon littoralis*), the most extreme example known of geographic isolation leading to genomic monomorphism (Robinson et al., 2016). A likely contributor to this pattern in *E. panamintina* is small genetic effective population sizes, ranging from ~18 to ~300 (the latter with a large confidence interval), which corroborate field observations of low-density populations (Clause et al., 2015; Mahrdt and Beaman, 2002; Stebbins, 2003; Thomson et al., 2016). Low effective population sizes can result from population bottlenecks, reproductive skew, limited connectivity between populations (Charlesworth, 2009); all of which make populations more

susceptible to environmental stochasticity which may result in extinction (Melbourne and Hastings, 2008).

Strikingly, pairwise population differentiation is low to moderate, despite overall low levels of genetic variation. Major structure is apparent at the scale of mountain ranges, with the White Mountains, Inyo Mountains and a Southern Ranges group all identified as distinct genetic clusters. Within these major geographic units, pairwise *FSTs* range from 0.08 to 0.18 (Table 2.3), a "moderate" level of differentiation (Hall and Clark 1997), while between these geographic units *FSTs* range from 0.19-0.30. Similarly, global *FSTs* indicate high population structure on the range-wide scale and moderate population substructure within geographic subunits. AMOVAs indicate that beyond the major geographic units, diversity is more partitioned within localities than between them, reflecting a weak signal of population structure at the within-mountain range spatial scale. This is corroborated by FASTSTRUCTURE results, which, on the range-wide scale, reflect major geographic units with relatively little admixture.

Maximum likelihood trees largely agree with FASTSTRUCTURE and PCA, which indicate that the deepest split is between the White Mountains populations and the rest of the range. Further, major divisions are concordant with geographic clusters with the exception of Barrel Spring, the northern-most population in the Inyo Mountains, which resolves as sister to both the Inyo Mountains and the southern mountains while PCA suggests that this locality is closer to the Inyo Mountains group. The variable position of this locality could be the result of admixture or incomplete lineage sorting. However, given our general results, this locality is likely intermediate to the north and the south due to isolation by distance and clustering assignments are unable to resolve this due to incomplete geographic sampling. Further, TREEMIX analyses suggest a signal of historical migration between the Southern Mountains and the White

Mountains group, but not with the geographically proximate Inyo Mountains, suggesting that some corridor exists or existed between the northern and southern ends of the range.

We expected gene flow to be restricted and to observe relatively higher levels of differentiation between populations due to the apparent distributional pattern and habitat restriction of *E. panamintina*. However, our results suggest some level of ongoing or historical gene flow among adjacent localities within mountain ranges, with relatively less across mountain ranges. This is consistent with other desert-refugia adapted species, such as West African crocodiles (*Crocodylus suchus*) and American pika (*Ochotona princeps*), both of which exhibit broad scale patterns of genetic structure concordant with large-scale landscape features (river drainages in the former and mountain ranges in the latter (Velo-Antón et al., 2014; Wilkening et al., 2011).

Climate fluctuations over the last 10,000 years are likely a major contributor to genetic patterns evident in extant populations. This region has experienced considerable drying since the late Pleistocene, with the reduction or loss of regional lakes and springs and overall expansion of lowland scrub habitats (Grayson, 1993, 2000; Quade et al., 1998). Large-scale population structure and low genetic diversity could therefore be a result of contracting habitat suitability and climatic oscillations leading to small population sizes and eventual drift.

Spatial patterns and connectivity: We also investigated the role of landscape features in contributing to genetic structure and find that geography and environmental features likely constrain suitable habitat for this species. Our high-resolution species distribution model partially validates the existing dogma about this species: suitable habitat within the known range of *E. panamintina* is extremely restricted, particularly by terrain roughness, seasonality in potential

evapotranspiration, diurnal temperatures (Figure 2.6B, Table 2.1). However, examination of the SDM reveals that despite apparent habitat restriction, suitable habitat exists outside of our traditional understanding of habitat usage in this species. Genetic resistance surfaces echo this finding in the White Mountains region and suggest that connectivity exists between heretofore "isolated" localities. There is a strong component of isolation by distance at small and large geographic distances, likely due to the linear arrangement of the mountain ranges and intervening washes to which this species is restricted. This is similar to results of niche modeling for montane mammals of the Great Basin which demonstrate potential available habitat area and intra-mountain connectivity is more common than previously thought (Waltari and Guralnick, 2009). A mix of terrain, thermal and hydrological climate variables contributes to habitat suitability and connectivity, which most likely reflects both physiological constraints (i.e. thermoregulation) and habitat preferences. *Elgaria* species, in general, are eurythermic and exhibit far broader and cooler operative temperature ranges than other squamates, a characteristic that is validated by their relatively higher use of shaded habitat (Kingsbury, 1993; Telemeco, 2014). While field observations of habitat use in *E. panamintina* suggest similar habitat use to other species in the genus, thermal niches have not been characterized. That this species is sometimes nocturnal suggests that daytime temperatures or water loss rates may limit activity.

CONCLUSIONS

Our results suggest that *E. panamintina* is a species with extremely limited genetic diversity and small population sizes, despite some apparent gene flow between populations. Given its geographic limitation and overall low genetic diversity, *E. panamintina* is a species whose extinction risk is elevated simply by its population parameters and distribution

(Frankham, 2005; Thomson et al., 2016). Regionally, human development in the form of infrastructure is limited, but potential human mediated threats to this species have not been adequately understood or quantified (Thomson et al., 2016). The largest single landowners in the region are the City of Los Angles and LA Department of Water and Power, which extract ground water and divert streams. Many of the tributary streams to the Owens River have been modified in some way to aid in water diversion, including channelization or diversion into pipes (US Geological Survey, 1991) which leads to degradation and loss of important riparian habitats. While the isolated nature of its distribution may partially protect this species from the pervasive influence of human-mediated habitat loss, even a small reduction of already limited riparian habitat due to water diversion may constitute a breaking point for this species. An additional regional concern is the development of new mining infrastructure. Mining has a long history in the region, but on a relatively small scale. New large-scale modern mines, such as the recently proposed lithium mine in the Panamint Valley, could pose a real threat to the persistence of this species.

ACKNOWLEDGEMENTS

We thank the Museum of Vertebrate Zoology, University of California, Berkeley, for the loan of several tissues. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. This work was supported by the USFWS, "Landscape Genomics Species Sampling" (Grant #F17AC00444) and the UCLA EEB Departmental Research Award, 2017. This research used computing resources based upon work supported by the National Science Foundation under Grant Nos. DBI-1458641 and ABI-1062432 to Indiana University.

FIGURES

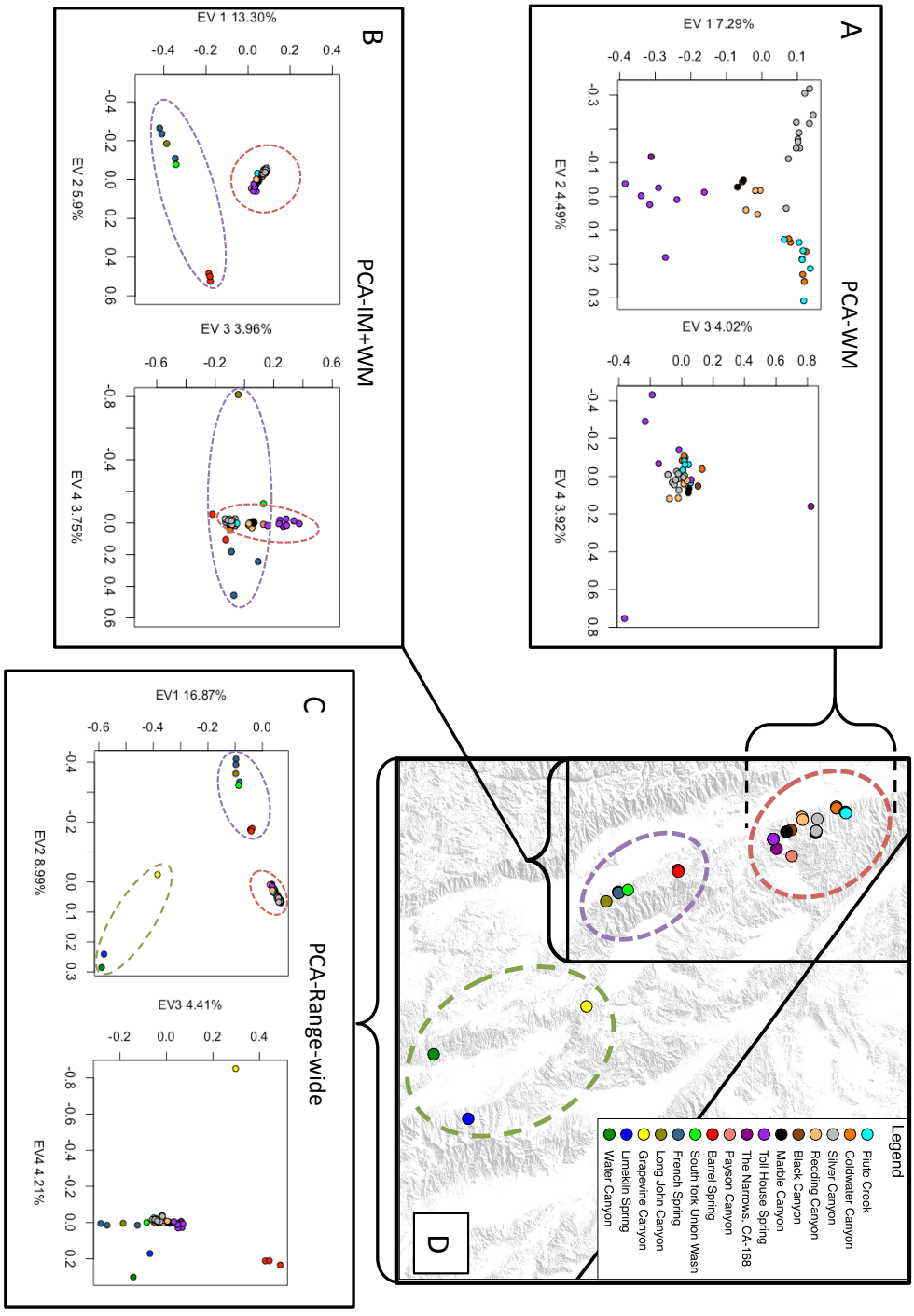
A
PCA-WM
EV 1 7.29%
EV 2 4.49%
EV 3 4.02%
EV 4 3.92%
B
PCA-IM+WM
EV 1 13.30%
EV 2 5.9%
EV 3 3.96%
EV 4 3.75%
C
PCA-Range-wide
EV1 16.87%
EV2 8.99%
EV3 4.41%
EV4 4.21%
D
Legend
Plute Creek
Coldwater Canyon
Silver Canyon
Redding Canyon
Black Canyon
Marble Canyon
Toll House Spring
The Narrows, CA-168
Payson Canyon
Barrel Spring
South fork Union Wash
French Spring
Long John Canyon
Grapevine Canyon
Limekiln Spring
Water Canyon

Figure 2.1: Principal components analysis (PCA) of the White Mountains (**A**), the Inyo + White Mountains (**B**), range-wide (**C**) all demonstrate a clear effect of large biogeographic regions, primarily defined by the White and Inyo Mountain Ranges. A map of sampling localities (**D**) demonstrates three broad scale genetic clusters identified in Bayesian clustering (dashed ellipses).

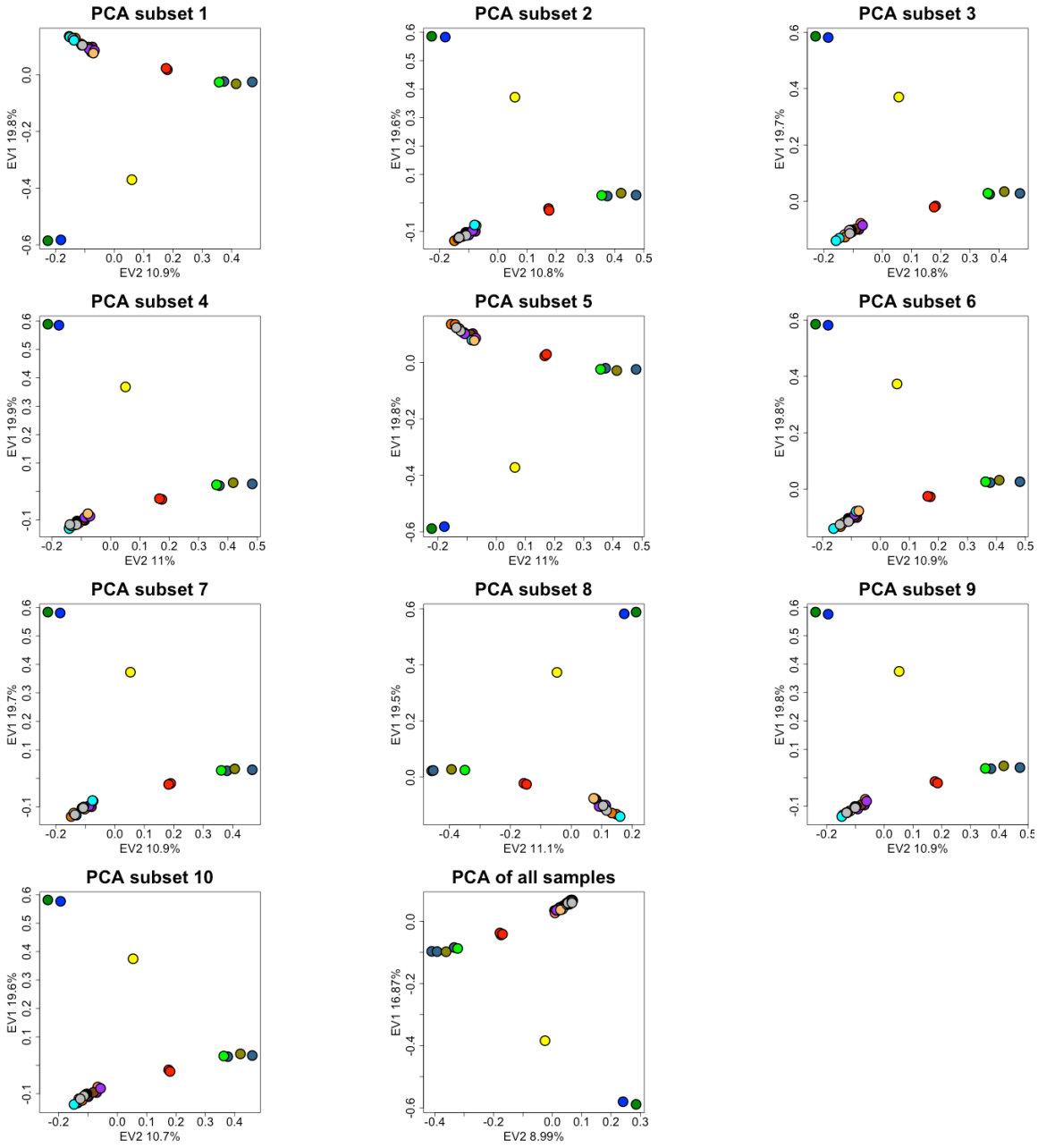

Figure 2.2: PCA replicates for sub-sampled populations in all localities (range-wide) demonstrate that broad-scale patterns of structure across the range are not driven solely by single sample localities. The final panel is the PCA of all samples for comparison.

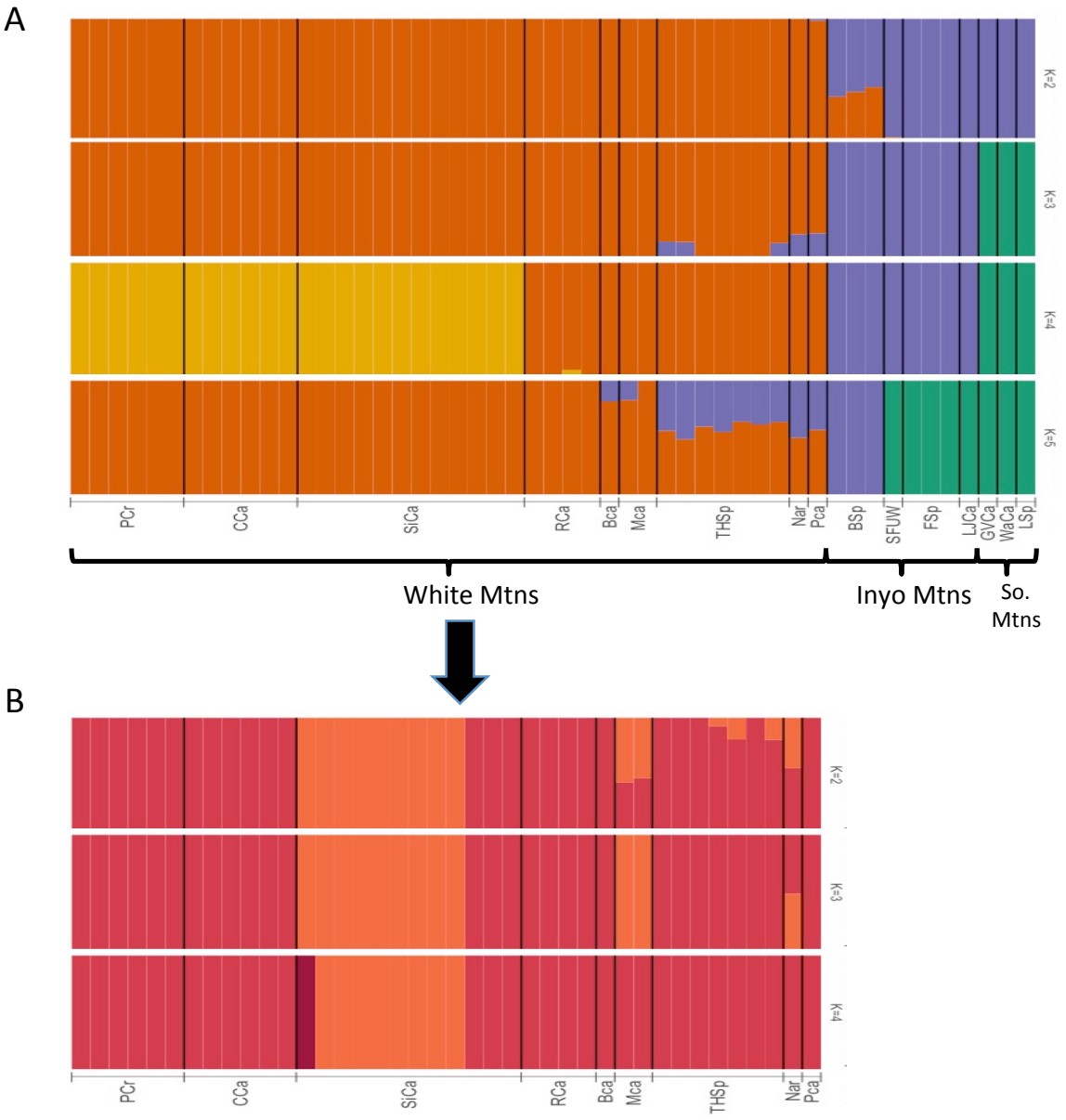

Figure 2.3: FASTSTRUCTURE results for range-wide (**A**) and the White Mountains region (**B**).

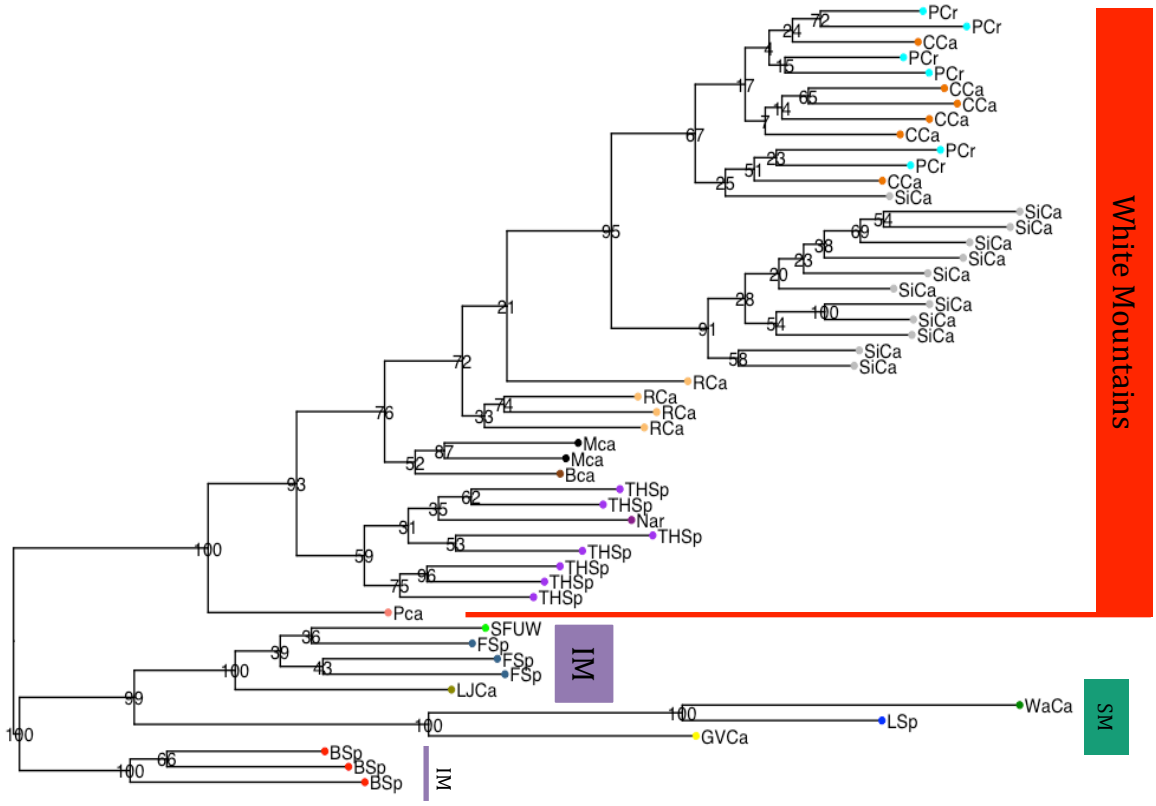

Figure 2.4: Midpoint rooted maximum likelihood tree of all samples identified four well-supported clades within *E. panamintina*.

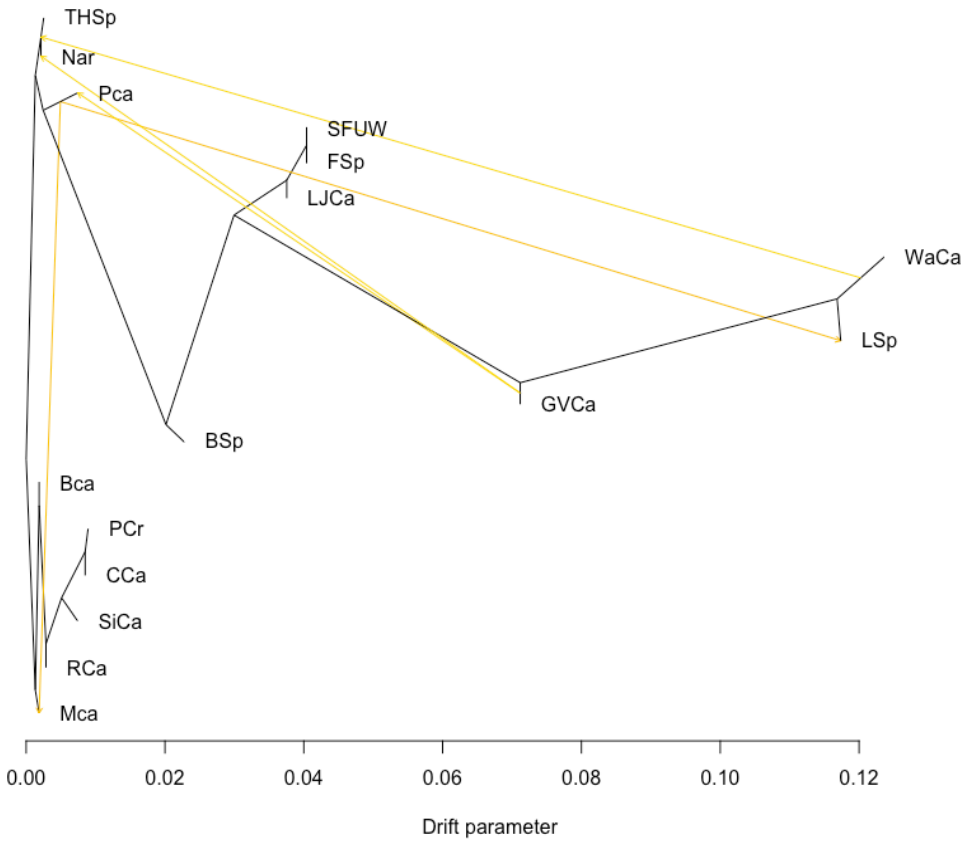

Figure 2.5: The highest likelihood tree with five migration edges suggests some level of current or historic gene flow between disjunct southern populations (GVCa, LSp, WaCa) and the southern end of the White Mountains (THSp, Nar, Pca), and between the northern (Mca) and southern White Mountains.

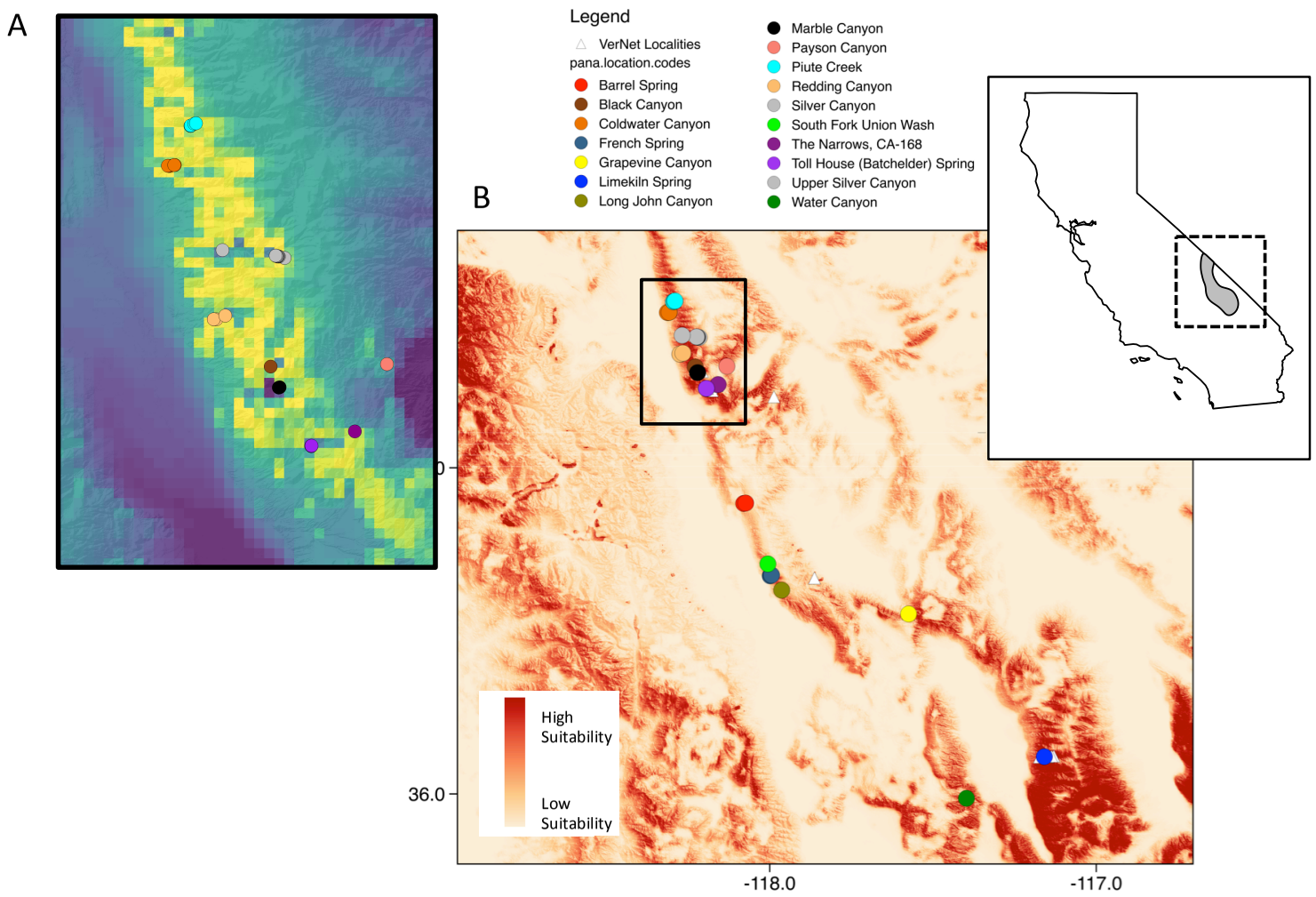

Figure 2.6: The highest likelihood resistance surface based on annual precipitation + SDM + continentality for the White Mountains Region (**A**). The average species distribution model for 10 replicates runs of MAXENT (**B**).

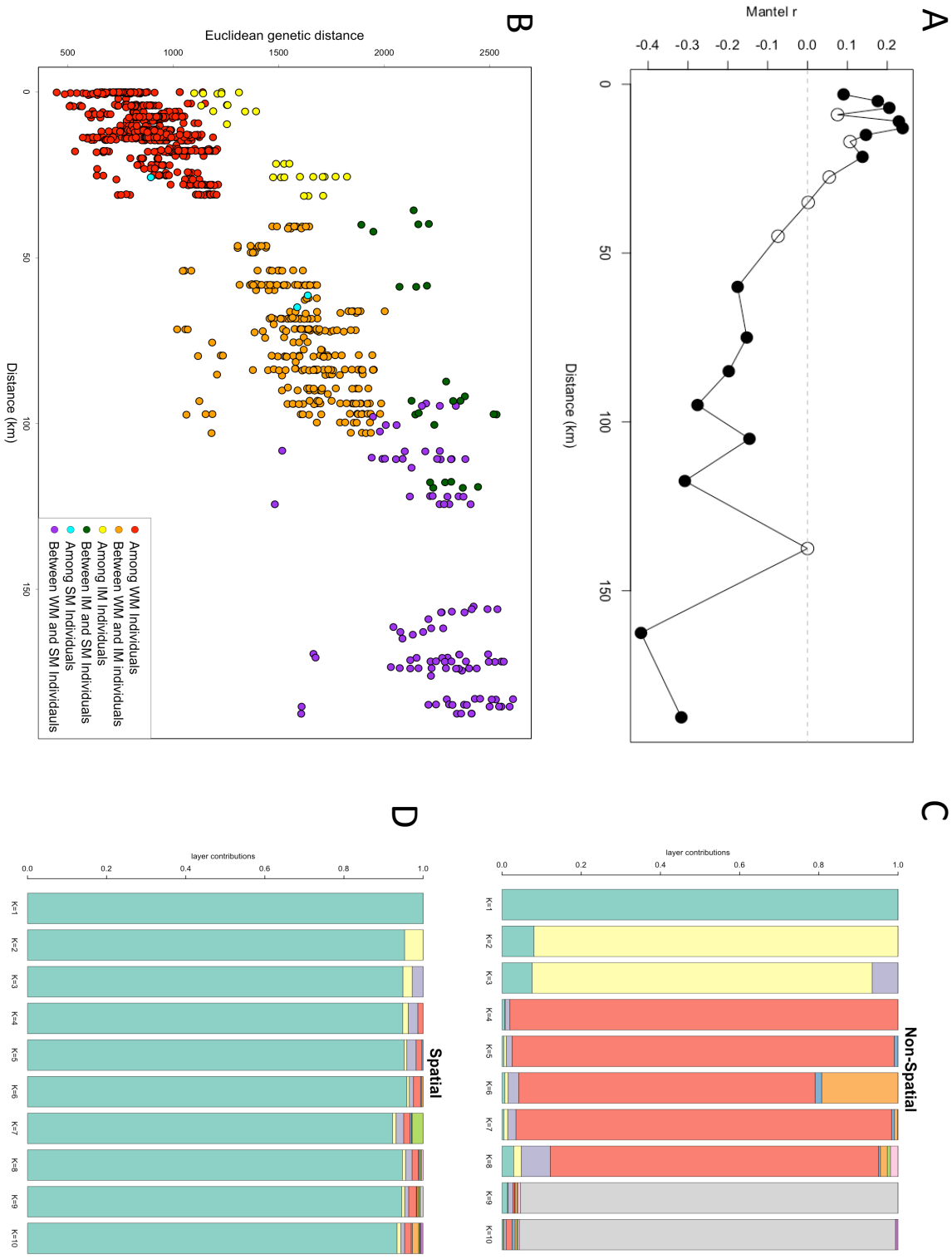
A
Mantel r
-0.4 -0.3 -0.2 -0.1 0.0 0.1 0.2
0 50 100 150
Distance (km)
B
Euclidean genetic distance
500 1000 1500 2000 2500
0 50 100 150
Distance (km)
Among WM Individuals
Between WM and IM individuals
Among IM Individuals
Between IM and SM Individuals
Among SM Individuals
Between WM and SM Individuals
C
layer contributions
0.0 0.2 0.4 0.6 0.8 1.0
K=1 K=2 K=3 K=4 K=5 K=6 K=7 K=8 K=9 K=10
Non-Spatial
D
layer contributions
0.0 0.2 0.4 0.6 0.8 1.0
K=1 K=2 K=3 K=4 K=5 K=6 K=7 K=8 K=9 K=10
Spatial

Figure 2.7: A Mantel correlogram demonstrates high positive correlation of genetic distance at short distance classes and negative correlation at large distance classes (**A**). A comparison of pairwise geographic distance to pairwise genetic distance demonstrates a strong linear effect of isolation by distance (**B**). Each point represents a comparison between two samples and points are colored based on the major geographic units of the two samples. Layer contributions of successive values of *k* in *conStruct* while ignoring (**C**) or accounting (**D**) for geographic proximity.

TABLES

Table 2.1: Permutation values for layers contributing to the averaged MAXENT model and single surface optimization results *ResistanceGA*.

Layer	*Maxent Permutation Importance*	*ResistanceGA Single Surface*	
		AICc	*R2*
Topographic Roughness Index	39.20	-	-
PET Seasonality	17.45	-	-
Mean Diurnal Range (Bio2)	11.89	-596.20	0.15
Aspect	8.04	-	-
Precipitation of Warmest Quarter (Bio18)	5.41	-	-
Emberger Q	3.99	-	-
Topographic Wetness Index	3.96	-	-
Slope	2.85	-	-
Precipitation Seasonality (Bio15)	2.15	-599.54	0.20
Topographic Position Index	1.02	-	-
Isothermality (Bio3)	0.93	-	-
Mean Temperature of Wettest Quarter (Bio8)	0.80	-	-
PET Warmest Quarter	0.59	-	-
Landcover	0.52	-	-
Precipitation of Coldest Quarter (Bio19)	0.34	-	-
Elevation	0.32	-	-
Canopy Cover	0.23	-	-
Precipitation of Wettest Month (Bio13)	0.22	-	-
Mean Temperature of Driest Quarter (Bio9)	0.04	-592.99	0.06
Annual Precipitation (Bio12)	0.03	-595.36	0.11
Growing Deg. Days 0	0.02	-	-
Precipitation of Driest Quarter (Bio17)	0.01	-	-
Continentality	0.00	-592.89	0.07
Precipitation of Driest Month (Bio14)	0.00	-	-
Mean Temperature of Warmest Quarter (Bio10)	0.00	-	-
Mean Temperature of Coldest Quarter (Bio11)	0.00	-	-
Precipitation of Wettest Quarter (Bio16)	0.00	-	-
Annual Mean Temperature (Bio1)	0.00	-	-
Temperature Seasonality (Bio4)	0.00	-	-
Max Temperature of Warmest Month (Bio5)	0.00	-	-
Min Temperature of Coldest Month (Bio6)	0.00	-	-
Temperature Annual Range (Bio7)	0.00	-	-
PET Coldest Quarter	0.00	-	-
PET Driest Quarter	0.00	-	-
PET Wettest Quarter	0.00	-	-
Annual PET	0.00	-	-
Aridity	0.00	-	-
Climatic Moisture Index	0.00	-	-
Growing Deg. Days 5	0.00	-	-
Max Temperature Coldest Month	0.00	-	-
Min Temperature Warmest Month	0.00	-	-
Month Count By Temp 10	0.00	-	-
Thermicity	0.00	-	-

Table 2.2: Best scoring resistance surfaces (single and composite). *Bio02* = mean diurnal range, *Bio09* = mean temperature of the driest quarter, *Bio12* = annual precipitation, *Bio15* = precipitation seasonality, *SDM* = species distribution model, *topowet* = topographic wetness index.

Surface	*mean AICc*	*mean R^2m*	*mean Log Likelihood*	*Percent as the top model*
Bio12 + Continentality + SDM	-384.66	0.57	197.52	18.80
Bio12 + Continentality	-384.62	0.52	196.92	13.60
Bio12 + SDM	-382.92	0.52	196.07	21.50
Bio12 + topoWet + SDM	-382.07	0.46	196.22	2.80
Bio02 + Bio12 + SDM	-381.93	0.53	196.16	3.60
Bio12 + Bio15 + Continentality + SDM	-381.62	0.52	196.81	8.90
Bio12 + Bio15 + SDM	-381.35	0.50	195.86	5.70
SDM	-381.28	0.31	195.04	4.90
Bio09 + Bio15 + SDM	-379.45	0.32	194.92	2.60
Bio02 + Bio15 + Continentality	-378.56	0.28	194.47	5.10

Table 2.3: Genetic diversity indices and effective population sizes.

Locality	*N*	*Ho*	*Hs*	*AR*	*Population π*	*Pairwise π*	*Find*	*FIS*	*FST*	*Nei's D*	*Ne (LD)*
Piute Creek	4	0.09 ± 0.18	0.12 ± 0.19	1.11 ± 0.19	0.000053	0.000071 ± 0.000053	0.53	0.19 ± 0.49	0.15 ± 0.07	0.11 ± 0.08	
Coldwater Canyon	6	0.09 ±0.18	0.12 ± 0.19	1.11 ± 0.18	0.000054	0.000069 ± 0.000054	0.52	0.19 ± 0.49	0.16 ± 0.08	0.11 ± 0.08	54.48 (33.8 - 139.14)
Silver Canyon	1	0.09 ±0.17	0.12 ± 0.17	1.12 ± 0.17	0.000053	0.000069 ± 0.000053	0.52	0.24 ± 0.47	-	0.11 ± 0.08	
Redding Canyon	12	0.11 ±0.23	0.14 ± 0.23	1.13 ± 0.22	0.000053	0.000067 ± 0.000053	0.42	0.13 ± 0.54	-	0.11 ± 0.08	15.02 (13.83 - 16.38)
Black Canyon	1	0.11 ±0.31	-	1.18 ± 0.31	-	-	0.41	-	-	0.12 ± 0.07	
Marble Canyon	2	0.11 ±0.28	0.12 ± 0.24	1.11 ± 0.26	0.000054	0.000064 ± 0.000054	0.39	-0.09 ± 0.6	-	0.12 ± 0.08	
Toll House	7	0.13 ±0.21	0.15 ± 0.21	1.15 ± 0.2	0.000053	0.000067 ± 0.000053	0.33	0.14 ± 0.48	0.15 ± 0.04	0.1 ± 0.07	18.45 (16.46 - 20.89)
The Narrows	1	0.13 ±0.33	-	1.13 ± 0.33	0.000057	0.000071 ± 0.000057	0.33	-	-	0.13 ± 0.07	
Payson Canyon	6	0.11 ±0.31	-	1.11 ± 0.31	0.000056	0.00006 ± 0.000056	0.42	-	0.17 ± 0.09	0.13 ± 0.07	321.43 (77.16 - inf.)
Barrel Spring	3	0.15 ±0.27	0.19 ± 0.27	1.11 ± 0.26	0.000053	0.000077 ± 0.000053	0.22	0.12 ± 0.55	0.21 ± 0.03	0.14 ± 0.06	
South fork Union Wash	1	0.16 ±0.37	-	1.16 ± 0.37	0.000055	0.000073 ± 0.000055	0.12	-	0.14 ± 0.06	0.15 ± 0.06	
French Spring	3	0.16 ±0.28	0.18 ± 0.27	1.18 ± 0.26	0.000053	0.000073 ± 0.000053	0.15	0.05 ± 0.56	0.25 ± 0.04	0.15 ± 0.06	
Long John Canyon	1	0.15 ±0.36	-	1.15 ± 0.36	0.000055	0.000062 ± 0.000055	0.16	-	-	0.16 ± 0.06	
Grapevine Canyon	1	0.16 ±0.36	-	1.16 ± 0.36	0.000047	0.000131 ± 0.000047	0.16	-	-	0.19 ± 0.05	
Water Canyon	1	0.07 ±0.26	-	1.07 ± 0.26	0.000054	0.000195 ± 0.000054	0.61	-	-	0.22 ± 0.07	
Limekiln Spring	1	0.1 ±0.31	-	1.1 ± 0.31	0.000045	0.000158 ± 0.000045	0.43	-	-	0.22 ± 0.07	

Table 2.4: Results of AMOVA analyses for range-wide, White Mountains and Inyo Mountains. The southern mountains group was not analyzed separately because it consisted of only 3 samples.

Range-wide

Groups = Mountain Range Group;
Samples = Locality

	Sigma	%	d.f.	Sum of squares	Mean sq.	Obs.	Std.Obs	*P*
Variations Between Groups	119.53	**48.76**	2	2514.88	1257.44	119.53	8.26	1.00E-04
Variations Between samples Within Groups	34.09	**13.91**	13	2563.94	197.23	34.09	12.02	1.00E-04
Variations Within samples	91.54	**37.34**	35	3203.79	91.54	91.54	-9.19	1.00E-04
Total Variations	245.16	**100**	50	8282.61	165.65			

Within the White Mountains

Groups = Localities

	Sigma	%	d.f.	Sum of squares	Mean sq.	Obs.	Std.Obs	*P*
Variations Between Groups	22.48	**22.72**	7	1246.55	178.08	22.48	15.33	1.00E-04
Variations Within Groups	76.47	**77.28**	31	2370.43	76.47			
Total Variations	98.95	**100**	38	3616.97	95.18			

Within the Inyo Mountains

Groups = Localities

	Sigma	%	d.f.	Sum of squares	Mean sq.	Obs.	Std.Obs	*P*
Variations Between Groups	101.17	**32.63**	4	1510.11	377.53	101.17	3.05	101.17
Variations Within Groups	208.92	**67.37**	4	835.67	208.92			
Total Variations	310.08	**100**	8	2345.78	293.22			

Table 2.5: Comparison estimates of genetic diversity from various other taxa demonstrate that *E. panamintina* exhibits extremely low diversity.

Organism	*data type*	H_O	H_E	A_R	F_{IS}	π	*Reference*
AustraliansSmall mammals	RAD	0.14-0.31	0.13-034	1.13-1.33			White et al., 2018
American pika, *Ochotona princeps*	RAD	0.30-0.361	0.29-0.33		-0.079		Waterhouse et al., 2018
Foothill yellow-legged frog, *Rana boylii*	RAD					0.001-0.003	McCartney-Melstad et al., 2017
leguminous shrub, *Caragana microphylla*	GBS	0.19-0.28					Xu et al., 2017
mayflies, *Baetis spp.*	RAD	0.12-0.18	0.25-0.28				Polato et al., 2017
Desert iguana, *Disosaurus dorsalis*	cytb, MC1R and Rag1					0.0005-0.0028	Hague et al., 2016
Chuckwalla, *sauromalus ater*	cytb, MC1R and Rag1					0.001-0.003	Hague et al., 2016
Zebratail lizards, *Calisaurus spp.*	cytb, MC1R and Rag1					0.003-0.006	Hague et al., 2016
Black toad, *Bufo exul*	microsatellite	0.518-0.558	0.484 - 0.539	3.00-3.909	-0.11 - 0.071		Wang 2009
Panamint alligator lizard, *Elgaria panamintina*	**RAD**	**0.12 ± 0.03**	**H_s: 0.14 ±0.03**	**1.13 ±0.03**	**0.12 ±0.10**	**0.000024 ± 0.0000235**	Present Study

Table 2.S1: Catalog number and locality information for samples used.

Catalog Number	*Location*
TC2975	Piute Creek
AGC1258	Piute Creek
AGC1257	Piute Creek
TC2974	Piute Creek
AGC1259	Piute Creek
EMT057	Piute Creek
AGC1251	Coldwater Canyon
AGC1234	Coldwater Canyon
AGC1239	Coldwater Canyon
AGC1232	Coldwater Canyon
EMT051	Coldwater Canyon
AGC1238	Coldwater Canyon
MVZ269776	Silver Canyon
TC2986	Silver Canyon
TC3035	Silver Canyon
TC3037	Silver Canyon
TC2957	Silver Canyon
TC2972	Silver Canyon
TC3036	Silver Canyon
TC3034	Silver Canyon
TC2959	Silver Canyon
TC2976	Silver Canyon
TC2971	Silver Canyon
TC2996	Silver Canyon
TC2985	Redding Canyon
AGC1245	Redding Canyon
AGC1246	Redding Canyon
AGC1244	Redding Canyon
AGC1223	Payson Canyon
AGC1247	Black Canyon
AGC1253	Marble Canyon
AGC1254	Marble Canyon
AGC1222	The Narrows
AGC1224	Toll House
AGC1260	Toll House
AGC1262	Toll House
TC2941	Toll House
TC2955	Toll House
AGC1261	Toll House
AGC1221	Toll House
TC2080	Water Canyon
TC2077	Limekiln Spring
TC2063	South fork Union Wash
MVZ191076	Grapevine Canyon
AGC1230	Barrel Spring
AGC1229	Barrel Spring
AGC1267	Barrel Spring
TC1577	French Spring
AGC1162	French Spring
AGC1163	French Spring
TC1580	Long John Canyon

REFERENCES

Adkins Giese, C.L., Greenwald, D.N., and Curry, T. (2012). Petition to List 53 Amphibians and Reptiles in the United States as Threatened or Endangered Species Under the Endangered Species Act. Cent. Biol. Divers.

Batzer, D.P., and Baldwin, A.H. (2012). Wetland Habitats of North America: Ecology and Conservation Concerns (University of California Press).

Bradburd, G.S., Coop, G.M., and Ralph, P.L. (2018). Inferring Continuous and Discrete Population Genetic Structure Across Space. Genetics *210*, 33–52.

Brown, J.H. (1971). Mammals on Mountaintops: Nonequilibrium Insular Biogeography. Am. Nat. *105*, 467–478.

Chambers, J.C., Devoe, N., and Evenden, A. (2008). Collaborative management and research in the Great Basin - examining the issues and developing a framework for action (Ft. Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station).

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. *10*, 195–205.

Clause, A.G., Cunningham, L., and Emmerich, K. (2015). Public Comments for the USFWS Status Review of the Panamint Alligator Lizard *Elgaria panamintina*.

Clause, A.G., Norment, C.J., Cunningham, L., Emmerich, K., Buckmaster, N.G., Nordin, E., and Hansen, R.W. (2018). What is the Best Available Science?: Conservation Status of Two California Desert Vertebrates. BioRxiv 342048.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Diamond, J.M. (1989). The present, past and future of human-caused extinctions. Philos. Trans. R. Soc. Lond. B Biol. Sci. *325*, 469–477.

Diniz-Filho, J.A.F., Soares, T.N., Lima, J.S., Dobrovolski, R., Landeiro, V.L., de Campos Telles, M.P., Rangel, T.F., and Bini, L.M. (2013). Mantel test in population genetics. Genet. Mol. Biol. *36*, 475–485.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. (2014). NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour. *14*, 209–214.

Eaton, D.A.R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics *30*, 1844–1849.

Ellstrand, N.C., and Elam, D.R. (1993). Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. Annu. Rev. Ecol. Syst. *24*, 217–242.

Fleishman, E., Austin, G.T., and Murphy, D.D. (2001). Biogeography of Great Basin butterflies: revisiting patterns, paradigms, and climate change scenarios. Biol. J. Linn. Soc. *74*, 501–515.

Floyd, C.H., Vuren, D.H.V., and May, B. (2005). Marmots on Great Basin Mountaintops: Using Genetics to Test a Biogeographic Paradigm. Ecology *86*, 2145–2153.

Forrest, A., Escudero, M., Heuertz, M., Wilson, Y., Cano, E., and Vargas, P. (2017). Testing the hypothesis of low genetic diversity and population structure in narrow endemic species: the endangered *Antirrhinum charidemi* (Plantaginaceae). Bot. J. Linn. Soc. *183*, 260–270.

Frankham, R. (1995). Inbreeding and Extinction: A Threshold Effect. Conserv. Biol. *9*, 792–799.

Frankham, R. (1998). Inbreeding and Extinction: Island Populations. Conserv. Biol. *12*, 11.

Frankham, R. (2005). Genetics and extinction. Biol. Conserv. *126*, 131–140.

Frankham, R., Ballou, J.D., Ralls, K., Eldridge, M., Dudash, M.R., Fenster, C.B., Lacy, R.C., and Sunnucks, P. (2017). Genetic Management of Fragmented Animal and Plant Populations (Oxford University Press).

Gaston, K.J. (1994). Causes of rarity. In Rarity, (Dordrecht: Springer Netherlands), pp. 114–135.

Glenn, T.C., Nilsen, R.A., Kieran, T.J., Sanders, J.G., Bayona-Vásquez, N.J., Finger, J.W., Pierson, T.W., Bentley, K.E., Hoffberg, S.L., Louha, S., et al. (2019). Adapterama I: Universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). BioRxiv 049114.

Grayson, D.K. (1993). The desert's past: a natural prehistory of the Great Basin (Smithsonian Inst Pr).

Grayson, D.K. (2000). Mammalian responses to Middle Holocene climatic change in the Great Basin of the western United States. J. Biogeogr. *27*, 181–192.

Hall, C.A. (1991). Natural History of the White-Inyo Range, Eastern California (University of California Press).

Houston, D.D., Evans, R.P., and Shiozawa, D.K. (2012). Evaluating the genetic status of a Great Basin endemic minnow: the relict dace (*Relictus solitarius*).

IUCN (2019). The IUCN Red List of Threatened Species. Version 2019-1.

Jennings, M.R., and Hayes, M.P. (1994). Amphibian and reptile species of special concern in California. Calif. Dep. Fish Game Inland Fish. Div. Rancho Cordova.

Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics *27*, 3070–3071.

Karron, J.D. (1997). Genetic consequences of different patterns of distribution and abundance. In The Biology of Rarity, W.E. Kunin, and K.J. Gaston, eds. (Dordrecht: Springer Netherlands), pp. 174–189.

Kingsbury, B.A. (1993). Thermoregulatory Set Points of the Eurythermic Lizard Elgaria multicarinata. J. Herpetol. *27*, 241.

Kivimäki, I., Shimbo, M., and Saerens, M. (2014). Developments in the theory of randomized shortest paths with a comparison of graph node distances. Phys. Stat. Mech. Its Appl. *393*, 600–616.

Lawlor, T.E. (1998). Biogeography of Great Basin Mammals: Paradigm Lost? J. Mammal. *79*, 1111–1130.

Leaché, A.D., Banbury, B.L., Felsenstein, J., Oca, A.N.-M. de, and Stamatakis, A. (2015). Short Tree, Long Tree, Right Tree, Wrong Tree: New Acquisition Bias Corrections for Inferring SNP Phylogenies. Syst. Biol. syv053.

Legendre, P., and Fortin, M.J. (1989). Spatial pattern and ecological analysis. Vegetatio *80*, 107–138.

Legendre, P., and Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol. Ecol. Resour. *10*, 831–844.

Linck, E.B., and Battey, C.J. (2017). Minor allele frequency thresholds strongly affect population structure inference with genomic datasets.

Mac, M.J., Opler, P.A., Haecker, C.E., and Doran, P.D. (1998). Status and Trends of the Nation's Biological Resources, Volume 2, (U.S. Geological Survey).

Mahrdt, C., and Beaman, K. (2002). Panamint alligator lizard, *Elgaria panamintina*. Species account for the West Mojave Management Plan, Riverside, California.

Manel, S., Schwartz, M.K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. *18*, 189–197.

Marrotte, R.R., and Bowman, J. (2017). The relationship between least-cost and resistance distance. PLOS ONE *12*, e0174212.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *17*, 10–12.

Mateu-Andrés, I., and Segarra-Moragues, J.G. (2000). Population subdivision and genetic diversity in two narrow endemics of Antirrhinum L. Mol. Ecol. *9*, 2081–2087.

Melbourne, B.A., and Hastings, A. (2008). Extinction risk depends strongly on factors contributing to stochasticity. Nature *454*, 100–103.

Merow, C., Smith, M.J., and Silander, J.A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography *36*, 1058–1069.

Minshall, G.W., Jensen, S.E., and Platts, W.S. (1989). The Ecology of Stream and Riparian Habitats of the Great Basin Region: A Community Profile. US Fish Wildl. Serv. Biol. Rep. *85*, 142.

Nei, M. (1987). Molecular Evolutionary Genetics (Columbia University Press).

Pearson, R.G., Raxworthy, C.J., Nakamura, M., and Peterson, A.T. (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeogr. *34*, 102–117.

Peterman, W.E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. Methods Ecol. Evol. *9*, 1638–1647.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. (2014). PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. Mol. Biol. Evol. *31*, 1929–1936.

Phillips, S.J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography *31*, 161–175.

Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet *8*, e1002967.

Quade, J., Forester, R.M., Pratt, W.L., and Carter, C. (1998). Black Mats, Spring-Fed Streams, and Late-Glacial-Age Recharge in the Southern Great Basin. Quat. Res. *49*, 129–148.

Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. Genetics *197*, 573–589.

Riddle, B.R., Jezkova, T., Hornsby, A.D., and Matocq, M.D. (2014). Assembling the modern Great Basin mammal biota: insights from molecular biogeography and the fossil record. J. Mammal. *95*, 1107–1127.

Robinson, J.A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B.Y., vonHoldt, B.M., Marsden, C.D., Lohmueller, K.E., and Wayne, R.K. (2016). Genomic Flatlining in the Endangered Island Fox. Curr. Biol. *26*, 1183–1189.

Sada, D.W., Williams, J.E., Silvey, J.C., Halford, A., Ramakka, J., Summers, P., and Lewis, L. (2001). Riparian area management: A guide to managing, restoring, and conserving springs in the Western United States. Technical Reference 1737-17. Bureau of Land Management, Denver, Colorado (BLM/ST/ST-01/001+ 1737).

Searcy, C.A., and Shaffer, H.B. (2016). Do Ecological Niche Models Accurately Identify Climatic Determinants of Species Ranges? Am. Nat. *187*, 423–435.

Shepard, W.D. (1993). Desert springs-both rare and endangered. Aquat. Conserv. Mar. Freshw. Ecosyst. *3*, 351–359.

Sodhi, N.S., Brook, B.W., and Bradshaw, C.J.A. (2009). Causes and Consequences of Species Extinctions. In The Princeton Guide to Ecology, (Princeton University Press), p. pp.514-520.

Sork, V., and Waits, L. (2010). Contributions of landscape genetics–approaches, insights, and future potential. Mol. Ecol. *19*, 3489–3495.

Spielman, D., Brook, B.W., and Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proc. Natl. Acad. Sci. U. S. A. *101*, 15261–15264.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Stebbins, R.C. (1958). A new alligator lizard from the Panamint Mountains, Inyo County, California. Am. Mus. Novit. 1–27.

Stebbins, R.C. (2003). A Field Guide to Western Reptiles and Amphibians (Houghton Mifflin Harcourt).

Stevens, L.E., and Meretsky, V.J. (2008). Aridland Springs in North America: Ecology and Conservation (University of Arizona Press).

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L., and Waits, L.P. (2007). Putting the 'landscape' in landscape genetics. Heredity *98*, 128–142.

Telemeco, R.S. (2014). Immobile and Mobile Life-History Stages Have Different Thermal Physiologies in a Lizard. Physiol. Biochem. Zool. *87*, 203–215.

Templeton, A.R., Shaw, K., Routman, E., and Davis, S.K. (1990). The Genetic Consequences of Habitat Fragmentation. Ann. Mo. Bot. Gard. *77*, 13–27.

Thomson, R.C., Wright, A.N., and Shaffer, H.B. (2016). California Amphibian and Reptile Species of Special Concern (University of California Press).

Tingley, R., Hitchmough, R.A., and Chapple, D.G. (2013). Life-history traits and extrinsic threats determine extinction risk in New Zealand lizards. Biol. Conserv. *165*, 62–68.

US Geological Survey (1991). Geology and Water Resources of Owens Valley, California.

Velo-Antón, G., Godinho, R., Campos, J.C., and Brito, J.C. (2014). Should I Stay or Should I Go? Dispersal and Population Structure in Small, Isolated Desert Populations of West African Crocodiles. PLOS ONE *9*, e94626.

Waltari, E., and Guralnick, R.P. (2009). Ecological niche modelling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. J. Biogeogr. *36*, 148–161.

Wang, I.J. (2009). Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). Mol. Ecol. *18*, 3847–3856.

Wang, J. (2016). A comparison of single-sample estimators of effective population sizes from genetic marker data. Mol. Ecol. *25*, 4692–4711.

Waples, R.S., and Do, C. (2008). ldne: a program for estimating effective population size from data on linkage disequilibrium. Mol. Ecol. Resour. *8*, 753–756.

Wilkening, J.L., Ray, C., Beever, E.A., and Brussard, P.F. (2011). Modeling contemporary range retraction in Great Basin pikas (*Ochotona princeps*) using data on microclimate and microhabitat. Quat. Int. *235*, 77–88.

Willing, E.-M., Bentzen, P., Van Oosterhout, C., Hoffmann, M., Cable, J., Breden, F., Weigel, D., and Dreyer, C. (2010). Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. Mol. Ecol. *19*, 968–984.

Wright, S. (1969). Evolution and the Genetics of Populations, Volume 2: Theory of Gene Frequencies (University of Chicago Press).

Yasuda, C.M. (2015). Ecology and Distribution of the Panamint Alligator Lizard (*Elgaria panamintina*). Masters Thesis. California State University, Chico.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics *28*, 3326–3328.

CHAPTER 3

GEOGRAPHIC AND ECOLOGICAL ISOLATION CONTRIBUTE TO RANGE-WIDE PATTERNS IN POPULATION STRUCTURE IN A WIDESPREAD LIZARD, *ELGARIA MULTICARINATA*

Erin Maurine Toffelmier

ABSTRACT

Understanding how the landscape influences patterns of population differentiation is an important step in examining how diversity is maintained across species' ranges. This is particularly true for species with large and heterogeneous ranges because regional evolutionary and demographic dynamics may change across the large spatial scale, depending on the local landscape context. We examined the interplay of isolating mechanisms in the wide-ranging southern alligator lizard, *Elgaria multicarinata*. This species is distributed through much of Western North America and is found in many terrestrial habitat types, thus providing an opportunity to examine how range-wide and regional dynamics contribute to divergence on large and small spatial scales. We used thousands of genetic markers to assess population structure, the relative importance of isolation by distance and environment, and constructed environmental niche models to assess ecological niche divergence among genetic clusters. We found that genetic divergence across the range of *E. multicarinata* is primarily driven by habitat suitability based distances, but that ecological divergence has occurred among genetic lineages. Comparisons of projected historical to current niches suggests that the total range of *E. multicarinata* has expanded since the Last Glacial Maximum and that current ecological divergence may have occurred following this range expansion.

INTRODUCTION

Environmental and geographic heterogeneity contribute to genetic differentiation across species' ranges (Avise, 2000; Jenkins et al., 2010; Ortego et al., 2012; Sork and Waits, 2010). This may be particularly true for widespread species, or for species with heterogeneous ranges, when both geographic and environmental distances are large. Across ranges, spatial patterns in population differentiation emerge due to the interplay of of genetic drift, selection, and gene flow (Rousset, 1997; Slatkin, 1985). Examining genetic differentiation with respect to geographic and environmental distance is a first step to identifying underlying evolutionary processes (e.g. natural selection, gene flow and drift) and examining spatial patterns in divergence has become a cornerstone of landscape genetic approaches (Holderegger and Wagner, 2008; Manel et al., 2003; Sork and Waits, 2010).

Two commonly quantified patterns are isolation by distance ("IBD"; Rousset, 1997; Wright, 1943) and isolation by environment ("IBE"; Sexton et al., 2014; Wang and Bradburd, 2014). IBD is the correlation of genetic distance with geographic distance: populations that are farther apart exchange relatively fewer migrants than populations that are close together (Duforet-Frebourg and Slatkin, 2016; Wright, 1943). While IBD is thought to be a relatively common phenomenon, landscape complexity and environmental differences likely contribute to divergence in many species (Crispo et al., 2006; Lee and Mitchell-Olds, 2011; Wang and Bradburd, 2014). IBE occurs when genetic divergence is correlated with environmental dissimilarity rather than geographic distance, and may result from divergent selection within different environments or biased dispersal between them (Crispo et al., 2006; Sexton et al., 2014; Wang and Bradburd, 2014; Wang and Summers, 2010; Wang et al., 2013). Parapatric lineages

that exhibit niche divergence may be a result of historical geographic isolation followed be niche divergence (Coyne and Orr, 1998; Mayr, 1963). Alternatively, geographic isolation may occur due to niche divergence followed by distributional shifts (Coyne and Orr, 1998; Wang and Bradburd, 2014). Moreover, geography and environment likely operate in tandem, reinforcing divergence at several levels (Nosil et al., 2005; Thorpe et al., 2008; Wang et al., 2013). It is therefore important to consider the effects of both geographic and ecological isolation when examining spatial patterns in phylogeographic structure.

Spatial patterns in population structure and divergence may also change across a species' range and it is important to consider both spatial scale and geographic range when making landscape inferences (Cushman and McGarigal, 2002; Trumbo et al., 2013). Because evolutionary processes may act differentially among scales and regions, conclusions drawn from a particular region or scale might not be applicable to other levels (Cushman and McGarigal, 2002; Wiens, 1989). Few studies examine the influence of landscape variation on genetic structure in more than one region and across spatial scales, despite the fact that many species exist in heterogeneous habitats.

Here we explore the roles of geography, and current and past climates in shaping regional and range-wide population structure in *Elgaria multicarinata*, the southern alligator lizard. This species ranges from southern Washington state to northern Baja, Mexico and inhabits chaparral, grasslands, oak woodlands and pine forests (Fitch, 1934, 1938; Stebbins, 2003). That it encompasses wide geographic and climatic gradients makes it an ideal target to examine how landscape heterogeneity acts differentially across species ranges. Previous analyses of mitochondrial and nuclear DNA revealed a deep north-south genetic split within *E. multicarinata* and further north-south clade subdivision largely concordant with geography (Feldman and

Spicer, 2006; Leavitt et al., 2017), suggesting that at least large scale landscape features are important drivers of structure in this species. However, because of the large heterogeneous range of *E. multicarinata*, ecologically mediated processes are likely to be significant contributors to range-wide substructure, though to date, no study has explicitly examined the role of biotic and abiotic factors in contributing to spatial structure in this species. We combine genome-wide genetic markers with ecological niche modeling and landscape genetics approaches to quantify the relative importance of IBD, IBE, and environmental niche in shaping spatial genetic structure in *E. multicarinata*.

MATERIALS AND METHODS

Laboratory methods

We obtained 208 samples (Table 3.S1) from five of the six species in the genus *Elgaria* (198 samples from across the range of *E. multicarinata*, six samples from across the range of *E. panamintina*, and one representative each of *E. coerulea, E. cedrosensis, E. paucicarinata,* and *E. kingii*). Genomic DNA for 99 samples was provided by C. Feldman that was extracted by phenol-chloroform purification (Feldman and Spicer, 2006; Maniatas et al., 1982). Forty-five of the samples from the current study are a subset of those used for previous genetic work on *E. multicarinata* in Feldman and Spicer (2006). We used salt extraction to extract genomic DNA from the remaining 109 samples following Sambrook and Russel (2001). For all samples, we generated reduced representation libraries with a three-enzyme restriction site associated DNA protocol (3RAD; Glenn et al., 2017) using the enzymes *MspI, ClaI* and *SphI*. Each individual sample was quadruple barcoded with unique 5' and 3' internal RAD-stub barcodes and external Illumina indices. Samples were aggregated in equimolar ratios into four pools ranging from 135

to 150 samples per pool (we included 294 samples for other projects). Each pool was size selected for 450-550bp fragments on a Pippin Prep (Sage Science, Beverly, MA) and sequenced individually on 150bp paired Illumina HiSeq 4000 lanes at the Vincent J. Coats Genomic Sequencing Laboratory. Following sequencing, samples were de-multiplexed based on external Illumina indices. We discarded reads that did not contain the correct internal RAD-stub barcodes, trimmed 20 base pairs from both the 3' and 5' ends of each read, and removed remaining low quality bases from the ends of reads using CUTADAPT V1.12 (Martin, 2011). We used IPYRAD v0.7.28 (Eaton, 2014) to *de novo* cluster reads within samples and align consensus sequences among samples using a clustering threshold value of 0.94 and all other default IPYRAD settings. We constructed data sets for 1) all species and samples and 2) *E. multicarinata* and *E. panamintina* together. For each set, we used VCFTOOLS v0.1.15 (Danecek et al., 2011) to select single nucleotide polymorphism (SNP) variants with 50% or less missing data, minor allele count greater than two, which did not exhibit an excess of heterozygotes ($p < 0.05$), and which were present in more than one individual.

Population structure and phylogenetic reconstruction

We examined broad scale patterns in population structure with two methods. First, we used a Bayesian approach to estimate individual ancestry coefficients in a spatially constrained framework implemented in the R program *tess3r* (Caye et al., 2016). We ran *tess3r* using a projected least squares algorithm with *k* values from 1 to 50, a maximum of 200 iterations per run, with 10 repetitions of each *k*. We masked 10% of genotypes to calculate the cross-entropy criterion to choose *k.* For visualization, we used a Kriging model to interpolate a projected map of genetic clusters. Second, we conducted maximum likelihood phylogenetic analysis in RAXML

v8.2.4 using the GTRGAMMA model of rate heterogeneity and 100 rapid bootstrap searches to assess node confidence (Leaché et al., 2015; Stamatakis, 2014). We included all samples from all species, setting *E. coerulea* as the outgroup, and included all SNPs present in a locus.

Genetic analyses

We excluded samples from *E. coerulea, E. cedrosensis, E. paucicarinata,* and *E. kingii* for all further genetic analyses. For each of the primary genetic clusters identified in clustering Bayesian clustering and phylogenetic reconstruction, we calculated Tajima's *D* as the average across loci for each genetic cluster in VCFTOOLS. To match the spatial resolution of the landscape layers we used in subsequent analyses, we thinned samples to one randomly selected sample per 5 km raster cell ($N = 124$) and calculated individual pairwise Euclidean genetic distance in *adegenet* v2.1.1 (Jombart and Ahmed, 2011). We used analysis of variance (ANOVA) implemented in R to compare the distributions of within-clade genetic distances.

Environmental niche modeling

To explore current and past distributional patterns, we estimated the geographic range and environmental niche of *E. multicarinata* with ecological niche models (ENM). We collected geo-referenced presence localities for *E. multicarinata* from the Global Biodiversity Information Facility (GBIF.org), including museum records from several collections and research grade records from the community generated database iNaturalist.org. We excluded records if their estimated coordinate error was greater than 1000m. To build the ENMs, we selected background points with a target group selection (TGS) approach based on locality records for all other species of reptiles that fell within 200km of the suspected range of *E. multicarinata* (Phillips and

Dudík, 2008; Searcy and Shaffer, 2016). We thinned both the presence localities and TGS localities to one locality per ~5km grid cell to match the resolution of the landscape and environmental layers (N = 8110). We collected a set of 19 bioclimatic variables (Karger et al., 2017), 18 expanded bioclimatic and topographic variables (Title and Bemmels, 2018), and three topological variables (slope, aspect, topographic position index (TPI)) extracted from a digital elevation model (US Geological Survey, 2013) using the *terrain* function in the R package *raster* v2.8-19 (Hijmans et al., 2019). To reduce multicollinearity among predictor variables, we used a stepwise procedure to remove layers with a variance inflation factor of 10 or greater and used the remaining variables (N= 17) to generate an average suitability model based on 10 replicates based on current climate and landscape variables in MAXENT v3.3.3 (Phillips and Dudík, 2008; Searcy and Shaffer, 2014). We obtained climate data from the Last Glacial Maximum (LGM; 21,000 years before present) generated from the Community Climate System Model version 4 based on PMIP3 data for the same bioclimatic and topographic variables (Karger et al., 2017; Title and Bemmels, 2018). All layers were standardized to ~5 km resolution using the RASTER package. We used the current ENM model to generate suitability layers for the current climate (range-ENM_{CUR}) and climate at LGM (range-ENM_{LGM}). Global ENMs may not fully capture the nuances of regional distributions because they estimate suitability based on representative localities from potentially very ecologically different parts of the range (Searcy and Shaffer, 2014). We therefore we also generated ENMs for each genetic cluster separately to examine regional distributional patterns. We defined the range of each cluster by generating a 1km-buffered minimum convex polygon around the genetic samples assigned to each cluster. We excluded four samples from Oregon, seven samples from Catalina and Coronado Islands, six samples from the eastern side of the Sierra Nevada mountains, and six *E. panamintina* samples

because there were large sampling gaps in between these and the remaining samples, which can make differentiating real spatial patterns from sampling artifacts difficult. Additionally, these excluded localities occur at the boundaries of the range, which often experience different or more extreme spatial processes and evolutionary forces (Sexton et al., 2009) and therefore may confound our attempts to identify more general processes and patterns. We selected presence and TGS background localities falling within these regional polygons to use in regional MAXENT models. We then projected these to generate regional current and historical suitability layers.

Assessing IBD and IBE

To assess the relative effects of IBD and IBE, we calculated five geographic distances (Euclidean geographic distance, least cost path distance for ENM_{CUR}, least cost path distance for ENM_{LGM}, optimized resistance distance for Bio13, optimized resistance distance based on slope), and three environmental distances (environmental dissimilarity, dissimilarity in current suitability, dissimilarity in LGM suitability).

Geographic distances: Landscape resistance surfaces can be used to generate a biologically-informed distance measure which is representative of the movement cost on a landscape (McRae, 2006; Peterman et al., 2019; Wang and Bradburd, 2014). Common approaches to generate resistance surfaces include utilizing suitability models (Wang et al., 2008) or relying on expert opinion. However, estimation of resistance values for particular landscape elements can be challenging when empirical data on movement are lacking (Peterman, 2018). We used two approaches to measure effective landscape distances, which we include under the umbrella of IBD hypotheses because they are representative of functional distances between localities rather

than differences in climatic or topographic qualities. First, we used ENM-defined habitat suitability to estimate pairwise least cost path (LCP) distances, which as been shown to reasonably estimate gene-flow in other systems (Wang et al., 2008). Pairwise LCP distances were calculated with *gdistance* v.1.2-2 (Etten, 2018) using Dijkstra's algorithm (Dijkstra, 1959) for range-ENM_{CUR} and range-ENM_{LGM}. We also calculated pairwise LCP distances within previously identified genetic clusters based on based on each regional ENM_{CUR} and regional ENM_{LGM} for within-region comparisons. Second, we used a maximum likelihood parameter estimation framework to estimate resistance values of individual landscape variables based genetic distances (*ResistanceGA* v4.0-14; (Peterman, 2018). We optimized resistance values based on commute distance for each of these remaining surfaces (N = 17) using the pairwise Euclidean genetic distance matrix as the response variable. We evaluated model fit in each of three replicates to select the two best performing optimized surfaces. Bio13 (precipitation in the wettest month) and slope were the best scoring surfaces in each replicate and we used *gdistance* to calculate pairwise commute distances between localities for each of these two surfaces. Finally, we calculated great-circle Euclidean geographic distance between localities in *fields* v9.7 (Nychka et al., 2017).

Environmental Dissimilarities: We characterized the environment at each locality by conducting a PCA of all variables using *prcomp* in R, and then calculated the environmental dissimilarity between points as the distance in all PCs. We calculated ENM_{CUR} and ENM_{LGM} dissimilarity as the difference in suitability between locality pairs. Dissimilarity in suitability is one measure of pairwise environmental distance between two localities, while LCP distance uses suitability to

calculate a weighted distance between localities and therefore is a measure of geographic distance (see above; Wang et al., 2013).

Statistical Approach: We used multiple matrix regression with randomization (MMRR) to quantify the effects of IBD and IBE on genetic distances using the R script 'MMRR' (Wang, 2013). Euclidean geographic distance, least cost path for ENM_{CUR}, and least cost path for ENM_{LGM}, exhibited high multicollinearity (Pearson's $r > 0.8$). Because each represents a plausible biological hypothesis explaining landscape genetic differentiation, we created three separate models which each included one of these three variables plus the remaining five uncorrelated distances. We constructed these three models for each of two spatial scales: 1) range-wide (excluding geographic outliers as outlined above) and 2) regional (the five primary regions defined by genetic clusters). The range-wide MMRRs included the LCP distances based on range-ENM_{CUR} and range-ENM_{LGM}, while the regional MMRRs include LCPs based on each region's ENM_{CUR} and ENM_{LGM}. All predictors were centered and scaled prior to analysis. We used 10,000 random permutations to construct a null distribution to assess model and parameter significance, and evaluated model fit (R^2) of the three different models at each spatial scale to determine the best model.

Niche identity and divergence

We used niched identity and background tests to examine niche divergence among the five primary genetic groups. Niche identity tests compare the empirically measured niches of each group to a null distribution of pseudo-niches, which are generated by randomly reshuffling the occurrence points of both groups and recalculating the niche space (Warren et al., 2008,

2010). Species inhabiting different geographic regions may exhibit divergent niches simply due to differences in available bioclimatic space resulting from spatial autocorrelation in the environment. We therefore also conducted background tests to further test whether two ENMs were more similar than expected by chance given their respective available environmental spaces. We used *ENMTools* v0.2 (Warren et al., 2010) in R to conduct both identity and background tests for pairwise comparison of all regions (five groups, $N = 10$ comparisons), and generated two similarity indices (Schoener's *D* and Warren's *I*) for each test, and generated a null distribution based on 99 pseudo-replicate MAXENT models using the set of 17 uncorrelated landscape variables to assess significance. To explore whether current ecological and genetic divergence are related to distributions at the LGM, we examined overlap of current and historic distributions of both the range-wide ENM and the regional ENMs. Relative to current distributions, greater overlap at the LGM would suggest recent post-glacial expansion and parapatric ecological divergence, whereas isolation at the LGM would support a role for past allopatry. To generate a conservative estimate of regional range overlap, we calculated the total number of regions at each point with suitability > 0.3. Finally, we visualized overall change in the suitable range of *E. multicarinata* since the LGM as the difference between the current and LGM range-wide suitability models.

RESULTS

Genetic data

The mean number of raw reads per sample was 3,311,251 (SE 1,988,018). After filtering, 43,964 SNP variants were retained across all species for use in RaxML, and 1,413 unlinked SNP variants were retained across *E. multicarinata* and *E. panamintina* and used for all other genetic

analyses. Tajima's *D* was positive for the South (0.067), coast (0.092), and Bay Area (0.002) clades, but negative for the north (-0.132) and west Sierra (-0.403) clades, indicative of recent demographic bottlenecks or range expansions in the latter two (Tajima, 1989). Within-clades genetic distances were significantly different among all clades, and followed a general south to north pattern of decreasing distances (Figure 2a and 2b).

Population structure and phylogenetic reconstruction

Maximum likelihood phylogenetic reconstruction identified four major clades with 100% bootstrap support (Figure 3.1b): 1) North (north of the San Francisco Bay through the north coast ranges, east through the Sacramento Valley, and northern Sierra Nevada Mountains), 2) Bay Area (coastal and valley habitats from San Francisco to Monterey Bays) , 3) Central Coast (coast ranges from Monterey Bay to south of Point Conception in Ventura County), 4) Western Sierra (west-slope of the central and southern Sierra Nevada Mountains). We also found strong support for an East Sierra group (eastern side of the Sierra Nevada Mountains in Inyo County) and *Elgaria panamintina* (the White and Inyo Mountains east of the Owen's Valley). The remaining southern California samples form a poorly support South group (Ventura county south through the Transverse Ranges and onto the Baja Peninsula). The rate of decrease in *tess3r* cross-validation scores slowed after $k = 7$, but reaches a minimum value at $k = 35$ before increasing (Figure 3.S1A). Examination of hierarchical *tess3r k*-values suggests that samples from the eastern Sierra Nevada mountains and *E. panamintina* are part of the South group at $k = 5$, comprise a single group at $k = 6$, and are two separate groups at $k = 7$ (Figure 3.S1B-C, Figure 3.1A and 3.1C). These two sample clusters are identified in RAXML as well supported monophyletic groups (consistent with $k = 7$), but nested within the south clade, rather than

reciprocally monophyletic with respect to it, as suggested by $k = 5$ and $k = 6$ (Figure 3.S1B and S.S1C) in *tess3r*. Internal RAXML node support within this south clade is generally low, so the precise arrangement of each of these groups remains unresolved. Pending this resolution, we focus on the five primary genetic clusters (North, Bay Area, Coast, West Sierra, and South) given that they are geographically cohesive and generally strongly supported.

Environmental niche modeling and niche divergence

Range-wide (Figures 3.3A and 3.3B) and regional ENMs (Figures 3.S2 and 3.S3) performed well (AUC > 0.64, Table 1). All niche identity tests among the five genetic clusters were significant ($P < 0.03$ for all comparisons), indicating that each cluster encompasses a unique environmental niche. The empirical values of Warren's *I* ranged from 0.18 to 0.57, and Schoener's *D* ranged from 0.05 to 0.33 (Table 3.2). Similarly, background tests demonstrated that habitats occupied by each of the two genetic clusters in each comparison were less similar than expected based on their available habitat ($P < 0.01$ for all comparisons). Empirical values of Warren's *I* ranged from 0.18 to 0.45, and Schoener's *D* ranged from 0.03 to 0.25 for background tests (Table 3.2). Comparisons of the geographic extent of ENM overlap in historical (Figures 3.3C, 3.S2 and 3.S3) and current (Figures 3.3D, 3.S2 and 3.S3) climates indicated that current potential distributions overlap more than overlap of historical distributions. Additionally, the total geographic extent of regional niches has generally increased since the LGM. Similarly, comparison of the range-wide current to historic ENMs demonstrates that the extent of suitable habitat has also increased since the LGM (Figure 3.4B and 3.S3).

IBD versus IBE

The range-wide model MMRR model explained 55.0% of total genetic variance while clade-specific models explained 29.3%-58.2% of genetic variance within each geographic region (Table 3.3, Figure 3.2B). The best fit range-wide model (R^2 = 0.55, P = 9.999e-05) included significant contributions from least cost path for ENM_{CUR}, (β_{IBD} = 4.80, P = 9.99e-05) and environmental dissimilarity (β_{IBE} = 0.97, P = 0.003). When clades were considered separately, all models were significant overall, and all best-fit models were primarily driven by IBD, though the strength of this correlation varied somewhat. LCP based on regional-ENM_{CUR} had the highest correlation with genetic distance in both the North (β_{IBD} = 1.22, P = 0.00009999) and West Sierra clades (β_{IBD} = 1.22, P = 0.00009999). The West Sierra model also included a significant effect of resistance distance based on slope (IBD, β_{IBD} = 0.58, P = 0.038), and environmental dissimilarity (IBE, β_{IBD} = 0.50, P = 0.033). Geographic distance was the only significant component in the model for the Bay Area group (IBD, β_{IBD} = 1.40, P = 0.00009999), while LCP based on regional-ENM_{LGM} was the only significant component in the model for the South (IBD, β_{IBD} = 1.89, P = 0.00019998). The Coast model was significant overall, but no individual predictor was individually significant (R^2 = 0.58, P = 0.02)

DISCUSSION

Range-wide population structure in *Elgaria multicarinata*

Our analyses of phylogeography and population structure identified five major genetic units across the range of *E. multicarinata,* which are strongly concordant with geography. Additionally, Bayesian clustering suggests that the eastern Sierras and *E. panamintina* may be most closely related to each other, and together related to the rest of southern California (Figure

3.S2B and 3.S2C). Maximum likelihood phylogenetic reconstruction strongly supports the North, Bay Area, Coast, and West Sierra groups as monophyletic, but returned relatively weak support for a monophyletic South group. The east Sierra and *E. panamintina* groups also received high bootstrap support individually, but may be nested within the Couth group (Figure 3.1B). This work, coupled with previous genetic and morphological analyses (Feldman and Spicer, 2006; Leavitt et al., 2017; Telemeco, 2014), firmly support the monophyly of *E. panamintina,* and the likely paraphyly of *E. multicarinata*. Previous phylogenetic work has also found strong support for northern and southern clades, but variably placed *E. panamintina* as sister to a northern *E. multicarinata* clade (Leavitt et al., 2017) or nested within a southern *E. multicarinata* clade (Feldman and Spicer, 2006; Leavitt et al., 2017). Also based on RAD data, we found that *E. panamintina* exhibits extremely low levels of genetic variation (Chapter 1 of this dissertation), which supports a hypothesis of peripatric speciation followed by incomplete lineage sorting as posited by Telemeco (2014). However, the position of *E. panamintina* relative to the rest of the southern clade remains unresolved and additional sampling of the Tehachapi Mountains, Transverse ranges and the southern extent of the eastern Sierra Nevada Mountains.

Within *E. multicarinata*, Feldman and Spicer (2006) identified four mtDNA lineages: Northern California, Southern Sierra Nevada, Coastal, and Southern California. With our large set of genome-wide genetic markers, we identified four well supported northern clades (Coast, Bay Area, West Sierra and North), as putatively sister to a southern group (Figure 3.1B). Phylogenetic and population genetic breaks occurring at the northern end of the Transverse Ranges, Monterey Bay, and the Sierra Nevada range echo similar patterns in many other western taxa and are largely concordant with the California Floristic Provinces (Calsbeek et al., 2003;

Feldman and Spicer, 2006; Lapointe and Rissler, 2005; McCartney-Melstad et al., 2018; Spinks and Shaffer, 2005).

IBD and IBE mediate patterns of divergence

Range-wide: We investigated how historical and contemporary landscape factors affect genetic differentiation across the range of *E. multicarinata*, and found evidence that range-wide genetic structure is linked to contemporary IBD and IBE. Range-wide, least cost path distance based on current suitability exhibited the strongest correlation with genetic distance ($\beta_{IBD} = 4.80$, $P = 0.0001$) and its regression coefficient was ~five times greater than that of environmental dissimilarity ($\beta_{IBE} = 0.97$, $P = 0.003$). Overall, this suggests that on large biogeographic scales, spatial isolation occurs as dispersal costs accrue due to low suitability in intervening habitat. Differences in dispersal has been linked to increased divergence, even on small spatial scales (Garant et al., 2005). One caveat is that LCP and Euclidean geographic distances are highly correlated, especially over long geographic distances. However, when we compared the Euclidean geographic distance based model to a LCP distance based model, the LCP model explained more variation than the simple model ($R^2 = 0.604$ versus 0.429), and in all models IBD was a stronger predictor than IBE.

While IBD appears to be the largest overall factor contributing to differentiation, the smaller signal of IBE was corroborated by niche background and identity tests. These tests, which quantify niche similarity, demonstrated that all five of the occupied regional niches are significantly less similar than expected by chance (Table 3.2). Moreover, across regional ENMs, permutation importance of landscape predictors varied widely (Table 3.2) with no clearly shared patterns. Taken together, these highly divergence niche spaces likely contribute to the pattern of

IBE on the range-wide scale. Moreover, the projected extents of current regional niches are much larger than the spatial extent of the underlying genetic clusters (Figure 3.S2), which suggests that niche divergence alone does not drive genetic divergence. If genetic isolation was due to niche divergence exclusively, we would expect that the geographic extent of the projected regional niche spaces should exhibit minimal overlap. Given the results of our correlational modeling, IBD appears to be the stronger driving force.

Comparison of the potential range at the Last Glacial Maximum (LGM) to the current range (Figure 3.3A and 3.3B) indicated that the total range of *E. multicarinata* has expanded since the LGM. Putative range overlap of the regional niches at the LGM was more limited than potential contemporary overlap (Figure 3.3C versus Figure 3.3D), and the intersections of regional niches correlate well with projected suitability based on range-wide data at point time points (Figure 3.3A compared to 3.3C, and Figure 3.3B compared to 3.3D). A smaller total number of sympatric LGM niches might suggest that contemporary divergence is a result of past allopatry through long term isolation. However, the niche currently occupied by the west Sierra cluster was virtually absent at the LGM (Figure 3.S2G). After adjusting the total number of potential overlaps at the LGM, it appears that potential sympatry across the LGM range is roughly similar to potential sympatry across the current range. Additionally, the limited of geographic overlap in current genetic groups is consistent with an parapatric or allopatric mode of divergence (Losos and Glor, 2003). Taken together, our range-wide correlational models and comparative niche modeling suggest that genetic divergence among genetic clusters is due primarily to IBD and that subsequent niche divergence occurred following postglacial expansion and geographic isolation.

Within regions: Within regions, IBD was again the primary correlate to divergence in most regions, though the type of IBD differed. The North and West Sierra clusters both show a strong signal of IBD based on LCP of current suitability. The model for the West Sierra clade also includes minor contributions from slope resistance (IBD) and environmental dissimilarity (IBE), and was the only region that exhibited any signal of IBE (Table 3.3). Based on both regional and range-wide models, the historic distribution for the West Sierran niche was extremely limited and overlapped minimally with current regional distribution (Figure 3.S3G-I) which suggests that *E. multicarinata* recently expanded into this niche space as the Sierra Nevada glaciations receded. Similarly, historic and current distributions for the northern clade, especially into the extreme north of the current distribution (dashed lines in Figure 3.S2A), exhibit minimal geographic overlap (Figure 3.S2A and S2B). Range expansions after the LGM northward into Oregon and Washington, and eastward into the Sierras are consistent with a genetic signature of bottlenecks associated with colonization (Tajima's *D*:-0.403 and -0.132 for West Sierra and North clades, respectively). Recent work on an amphibian species co-distributed with the northern clade also demonstrated negative Tajima's *D*, suggesting a shared history of range expansion (McCartney-Melstad et al., 2018). In the Bay Area cluster, IBD based on geographic distance was the only important correlate in the MMRR models (Table 3.3). Because suitability is relatively homogeneous across this small region (Figure 3.S3E), this signature of geographic distance likely resolves in the absence of a strong signal of suitability-based distance. For the South cluster, which exhibited the highest pairwise genetic distances, IBD based on LGM suitability was the only significant correlate with genetic distance (Table 3.3), although the alterative model including LCP of ENM_{CUR}, explained nearly as much variation (27% versus 29%, data not shown). Current suitability is marginally higher and more evenly distributed than

during the LGM (Figure 3.S3M-O), potentially leading to higher recent gene flow in the region. Historical isolation of sub-regional populations may have contributed to relatively higher pairwise genetic divergence in this region, although we did not find a strong concordant signal of reciprocal monophyly of putative sub clades in phylogenetic reconstruction (geographically proximate localities do tend to cluster together, but node support is low, Figure 3.1B). The current pattern of large pairwise distances but relative lack of strong local population structure may therefore be the result of historical subpopulation isolation or limited gene flow, followed by secondary contact and increased gene flow since the LGM.

While many studies examine landscape drivers of differentiation on small spatial scales, we combined landscape analyses on regional and range-wide scales to examine how the impact of isolating mechanisms across these spatial scales may change. Our results indicate that both geographic and environmental variation play important roles in contributing to genetic differentiation in species with wide and heterogeneous ranges. We found that patterns of differentiation in *E. multicarinata* are largely mediated by geographic distances based on habitat suitability, but that regional environmental niches are divergent. Future work examining whether there are further isolating mechanisms, especially at the contact zones between apparently ecologically and genetically divergent lineages, would be informative.

ACKNOWLEDGMENTS

We thank the following for sample contributions: the Department of Herpetology, California Academy of Sciences; the Museum of Vertebrate Zoology, University of California, Berkeley; the San Diego State University Museum of Biodiversity; the Natural History Museum of Los Angeles County; and many individuals, including: C.R. Feldman, R. Clark, R. Cooper, J.

Grismer, M. and L. La Kretz, J. Lee, K. Neal, S. Rice, Y. Shin, and the Spring 2018 UCLA Herpetology Course. This work was supported by funding from the National Science Foundation GRFP, the UCLA La Kretz Center for California Conservation Science, and the UCLA Dissertation Year Fellowship. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

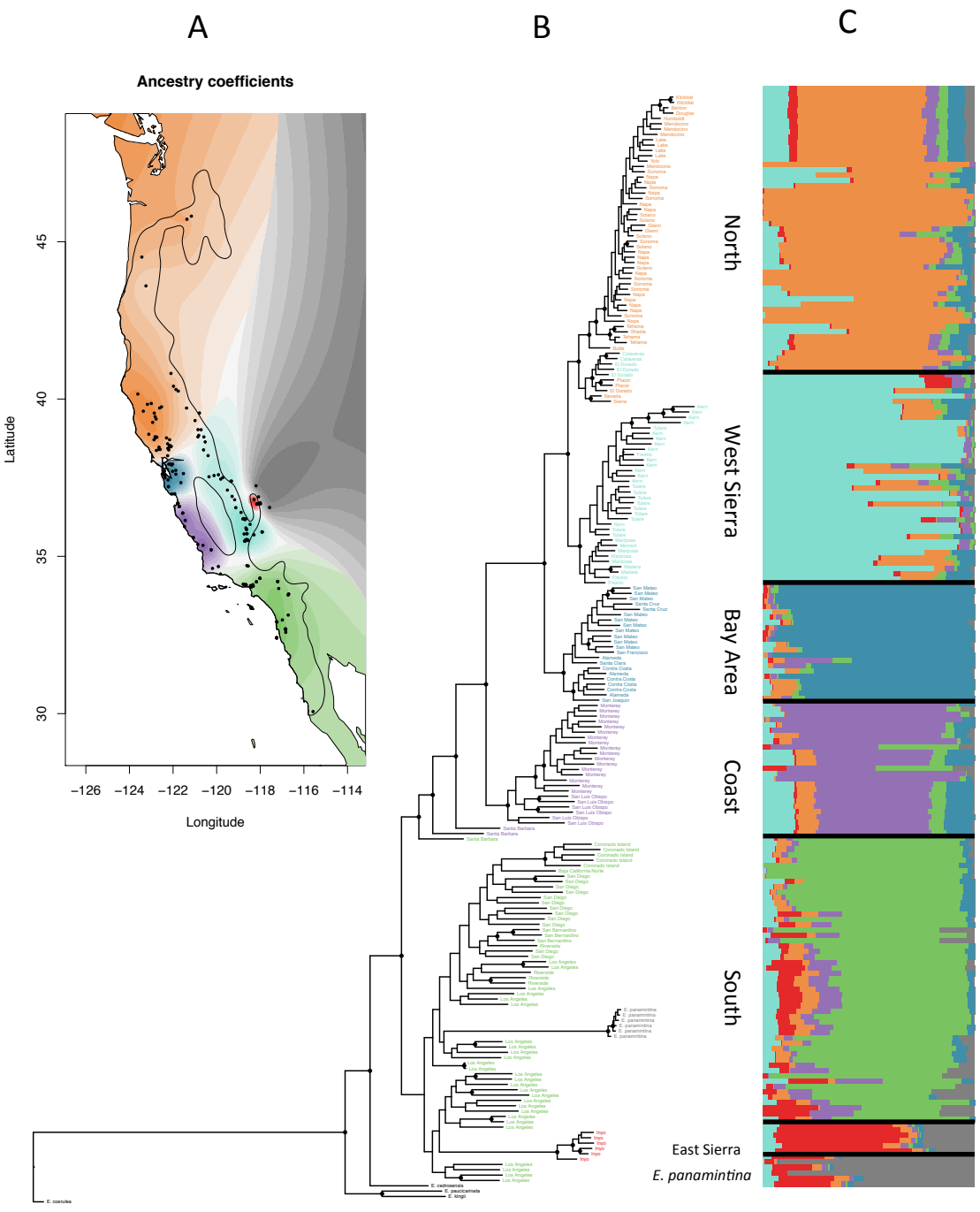

Figure 3.1: Range map of *E. multicarinata,* genetic sampling localities and spatial interpolation of ancestry coefficients for $k = 7$ from *tess3r* (**A**). The maximum likelihood phylogeny (**B**) recovers a split between a south clade and all localities north of the Transverse ranges (except eastern Sierra Nevada mountains *E. multicarinata* (red) and *E. panamintina* (grey), which cluster with the south clade (green)). Tips are labeled with county of origin and colored by their cluster assignment at $k = 7$ in *tess3r*. A barplot of spatially informed admixture assignment for $k = 7$ (**C**) demonstrates concordant patterns in range-wide genetic structure.

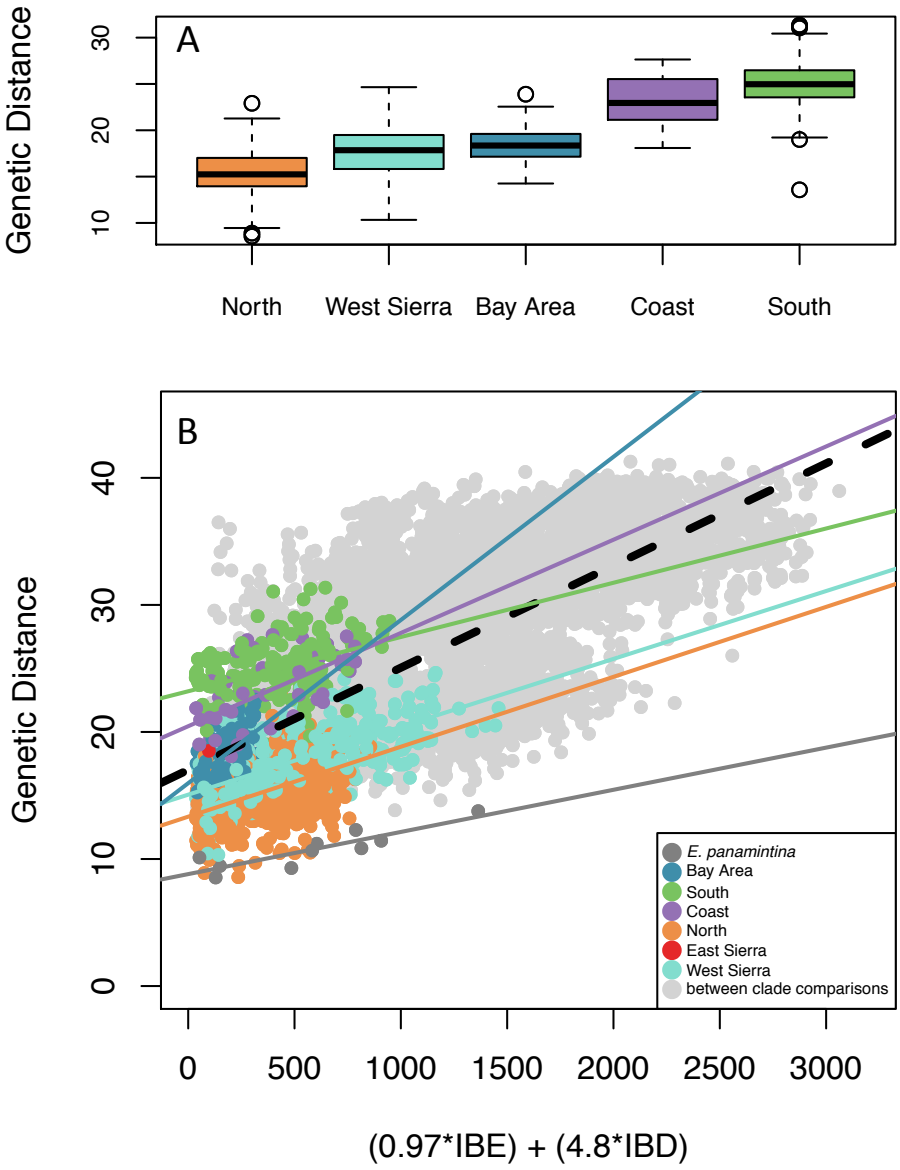

Figure 3.2: The distribution of within cluster genetic distances among major genetic clusters is significant for all comparisons between clusters and follows a general south to north trend of decreasing distances with increasing latitude (**A**). MMRR modeling demonstrates a strong correlation of IBD + IBE with pairwise genetic distances (**B**). In panel B, light grey points represent between clade comparisons, while colored points represent within-clade comparisons. The black dashed line represents the range-wide trend, while the thinner, colored lines represent models for each genetic cluster separately. We include localities from *E. panamintina* (dark grey) and *E. multicarinata* from the eastern Sierra Nevada Mountains (red) for comparison, but we did not build individual MMRR models for these groups due to limited sampling.

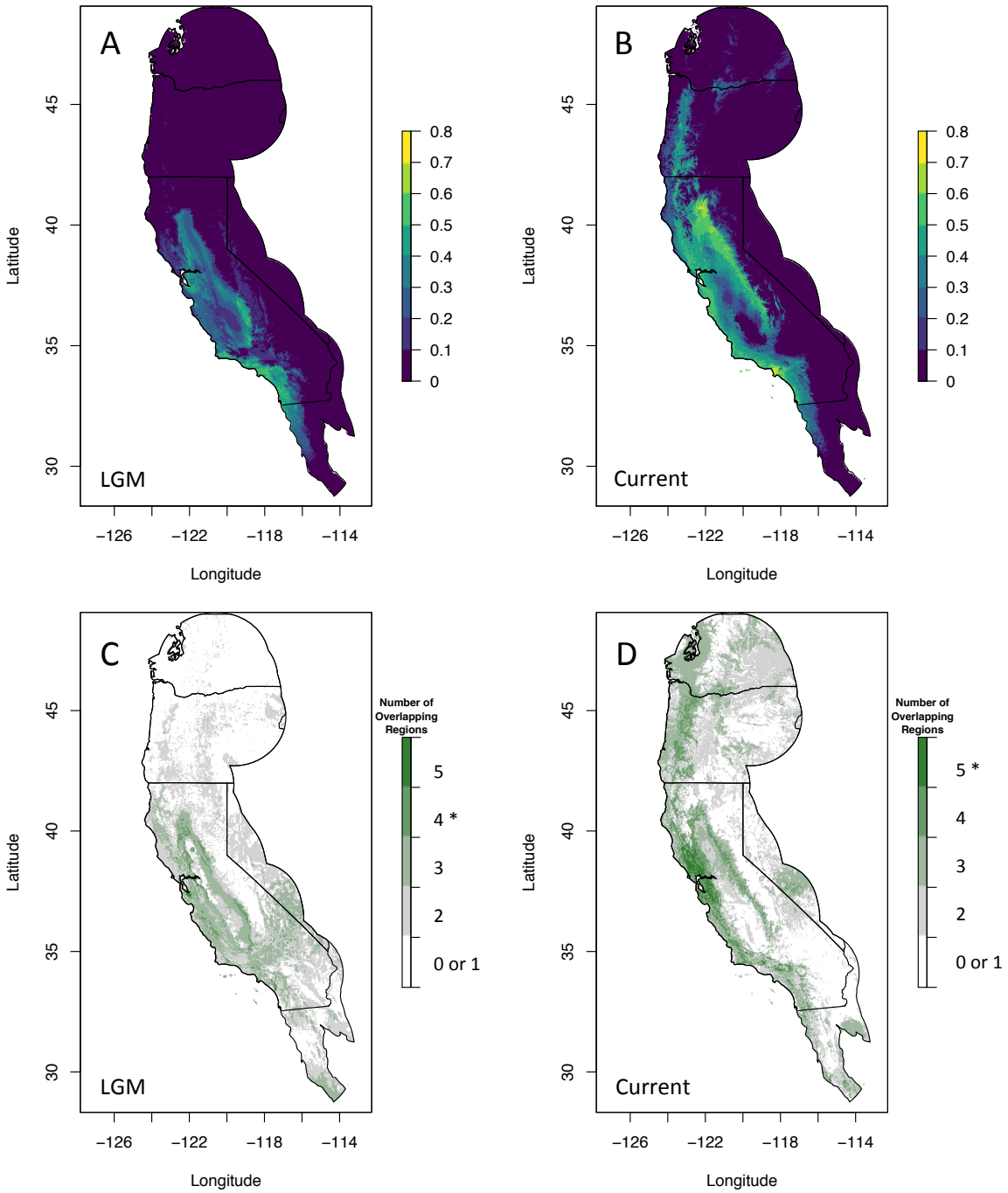

Figure 3.3: Range-wide ENM model projected to the LGM (**A**) and current (**B**) climates. Comparison of the ENMs suggests that the suitable range of *E. multicarinata* has increased since the LGM. See supplemental figures 3.S2 and 3.S3 for regional niche models, which differ slightly from the total range-wide model. Potential overlap of LGM (**C**) and contemporary (**D**) niches. (*) denotes the probable maximum potential number of overlaps during that time.

TABLES

Table 3.1: MAXENT performance and variable importance ranked by permutation importance for regional models. Values shaded with grey indicate top contributing variables to each model (Permutation importance ≥ 10).

	North	West Sierra	Bay Area	Coast	South
Model AUC	0.67	0.86	0.64	0.70	0.67
Landscape Variable – Permutation Importance					
Isothermality	23.17	0.94	14.33	19.49	14.90
Precipitation of Warmest Quarter	10.14	4.53	6.77	17.76	11.64
Mean Temperature of Driest Quarter	0.52	14.16	6.07	1.75	10.00
Mean Monthly PET of Driest Quarter	3.83	9.22	33.56	2.20	1.88
Mean Monthly PET of Wettest Quarter	1.23	11.15	2.92	1.30	26.20
Min Temperature of Warmest Quarter	18.68	2.89	0.00	7.23	1.98
Precipitation Seasonality	18.31	5.62	2.51	8.00	2.30
Aspect	3.72	9.20	3.36	1.28	0.90
Monthly Variability in Potential Evapotranspiration	2.55	10.01	0.00	5.66	0.37
Terrain Roughness Index	1.65	4.44	3.92	10.95	4.18
SAGA-GIS Topographic Wetness Index	0.70	5.07	3.53	5.27	1.73
Mean Temperature of Wettest Quarter	0.13	3.99	6.54	0.10	14.06
Topographic Position Index	7.49	7.89	4.28	1.11	2.63
Mean Diurnal Range	5.86	4.25	0.00	6.35	2.46
Slope	1.98	1.08	4.84	1.44	0.04
Precipitation of Wettest Month	0.05	5.55	4.50	5.65	3.06
Thornthwaite aridity index	0.00	0.00	2.88	4.46	1.67

Table 3.2: Results of background and identity tests for niche divergence. All tests were significant ($P < 0.01$) and indicated regional niches diverge more than expected under the null hypothesis of niche identity.

	Background				Identity			
	Schoener's		Warren's		Schoener's		Warren's	
Comparison	*D*	*P(D)*	*I*	*P(I)*	*D*	*P(D)*	*I*	*P(I)*
Coast – Bay Area	0.188	0.0100	0.447	0.0100	0.188	0.0300	0.447	0.0200
Coast - North	0.051	0.0100	0.182	0.0100	0.051	0.0100	0.182	0.0100
Coast - Sierra	0.172	0.0099	0.430	0.0099	0.172	0.0100	0.430	0.0100
Coast - South	0.035	0.0100	0.147	0.0100	0.035	0.0100	0.147	0.0100
North - Bay Area	0.230	0.0100	0.455	0.0100	0.230	0.0323	0.455	0.0323
North - Sierra	0.152	0.0100	0.377	0.0100	0.152	0.0100	0.377	0.0100
North - South	0.249	0.0100	0.471	0.0100	0.249	0.0100	0.471	0.0100
Sierra - Bay Area	0.174	0.0100	0.417	0.0100	0.174	0.0100	0.417	0.0100
Sierra - South	0.165	0.0100	0.418	0.0100	0.165	0.0100	0.418	0.0100
South - Bay Area	0.328	0.0100	0.555	0.0100	0.328	0.0200	0.569	0.0100

Table 3.3: Results of multiple matrix regression analyses for global and local models. Variables with significant contributions to the model are highlighted with bold text and grey shading. Bio13 is the precipitation in the wettest quarter.

	Model		Intercept		Isolation by Distance										Isolation by Environment					
					Geo. Distance		$LCP\ ENM_{CUR}$		$LCP\ ENM_{LGM}$		$Resistance_{Bio13}$		$Resistance_{Slope}$		Env. Dissimilarity		ENM_{CUR} Dissimilarity		ENM_{LGM} Dissimilarity	
	R^2	*P*	β	*P*	β	*P*	β	*P*	β	*P*	β	*P*	β	*P*	β	*P*	β	*P*	β	*P*
Global	0.604	0.0001	25.924	0.0001	-2.646	**0.001**	9.894	**0.0001**	-3.062	**0.0005**	0.710	0.142	-0.217	0.615	0.365	0.246	-0.078	0.783	0.309	0.071
North	0.272	0.0003	15.391	0.0003	-1.307	0.205	2.349	**0.035**	-0.159	0.727	0.355	0.285	-0.128	0.739	0.100	0.678	0.019	0.934	-0.209	0.163
Sierra	0.562	0.0001	17.554	0.0001	2.826	0.113	-0.366	0.828	-1.549	**0.011**	0.755	0.088	0.509	0.068	0.455	0.043	0.291	0.131	-0.081	0.633
Bay Area	0.475	0.0002	18.395	0.0007	1.102	**0.038**	0.511	0.537	-0.090	0.920	-0.568	0.133	-0.155	0.786	0.057	0.864	0.004	0.984	0.084	0.634
Coast	0.669	0.0238	22.929	0.0302	-9.866	0.250	28.251	**0.048**	-16.608	0.060	0.120	0.941	-0.490	0.679	0.052	0.923	-0.425	0.379	0.087	0.839
South	0.300	0.0008	24.984	0.0008	2.485	0.108	-4.379	**0.028**	3.789	0.104	-0.717	0.107	-0.951	0.113	0.767	0.105	-0.707	**0.023**	-0.002	0.995

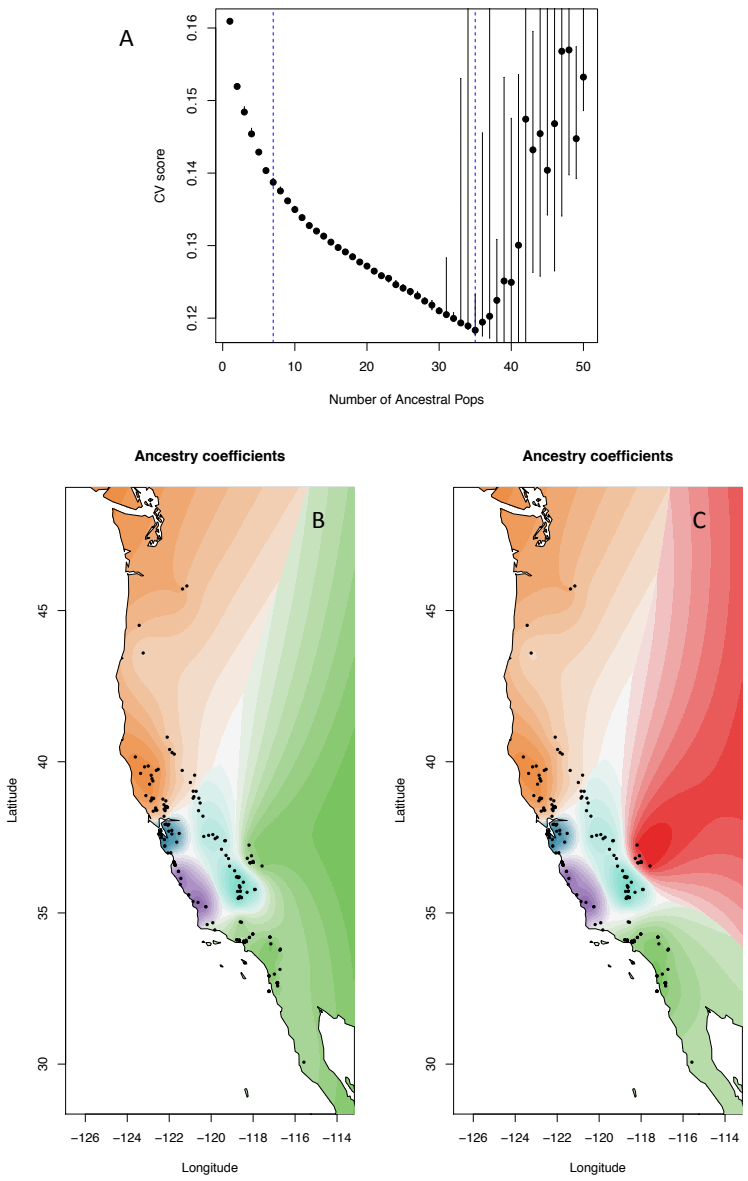

Figure 3.S1: Cross validation scores for $K = 1$ to $K = 50$ for *tess3r* (**A**). Spatial interpolation of admixture coefficients for $K = 5$ (**B**) and $K = 6$ (**C**).

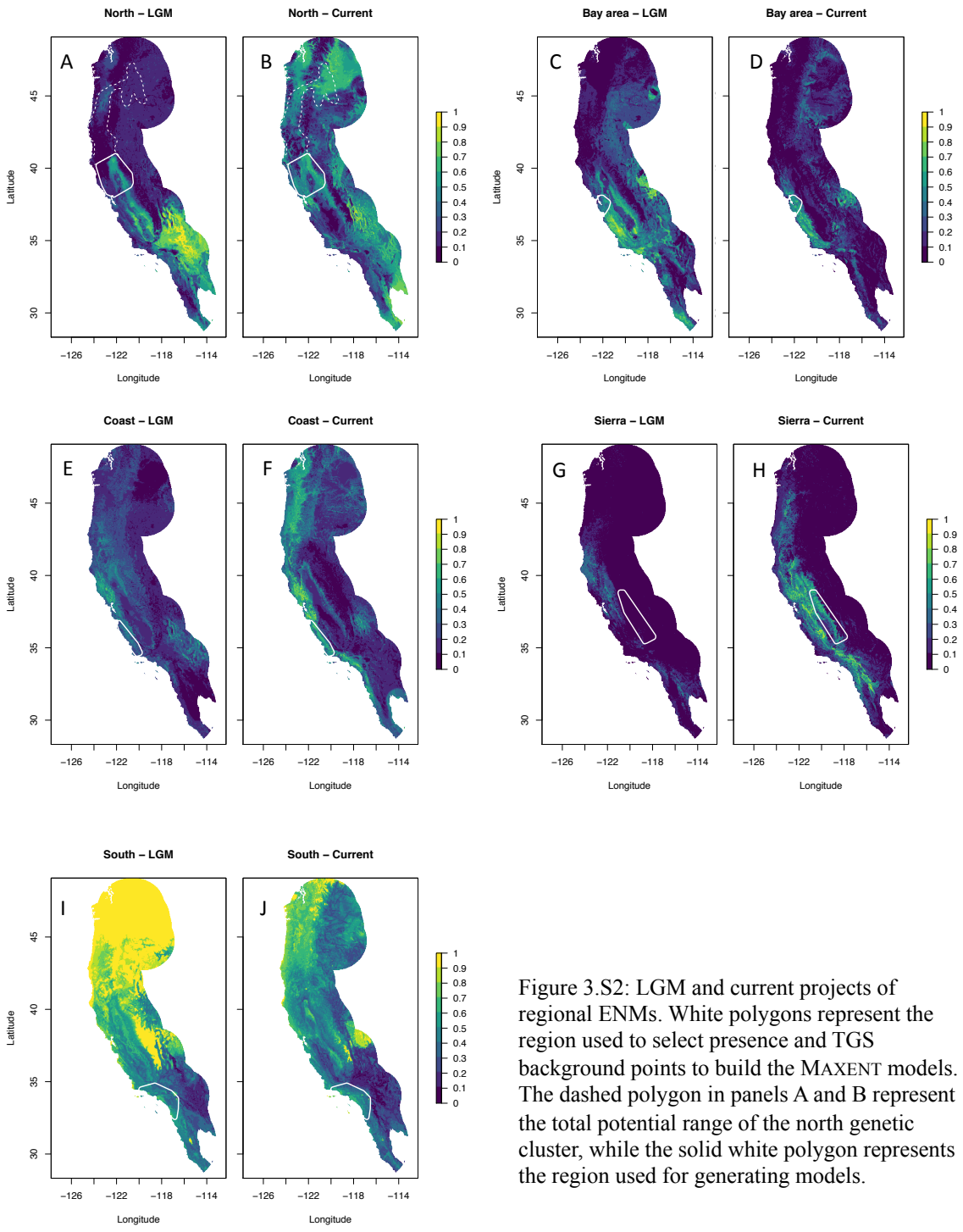

Figure 3.S2: LGM and current projects of regional ENMs. White polygons represent the region used to select presence and TGS background points to build the MAXENT models. The dashed polygon in panels A and B represent the total potential range of the north genetic cluster, while the solid white polygon represents the region used for generating models.

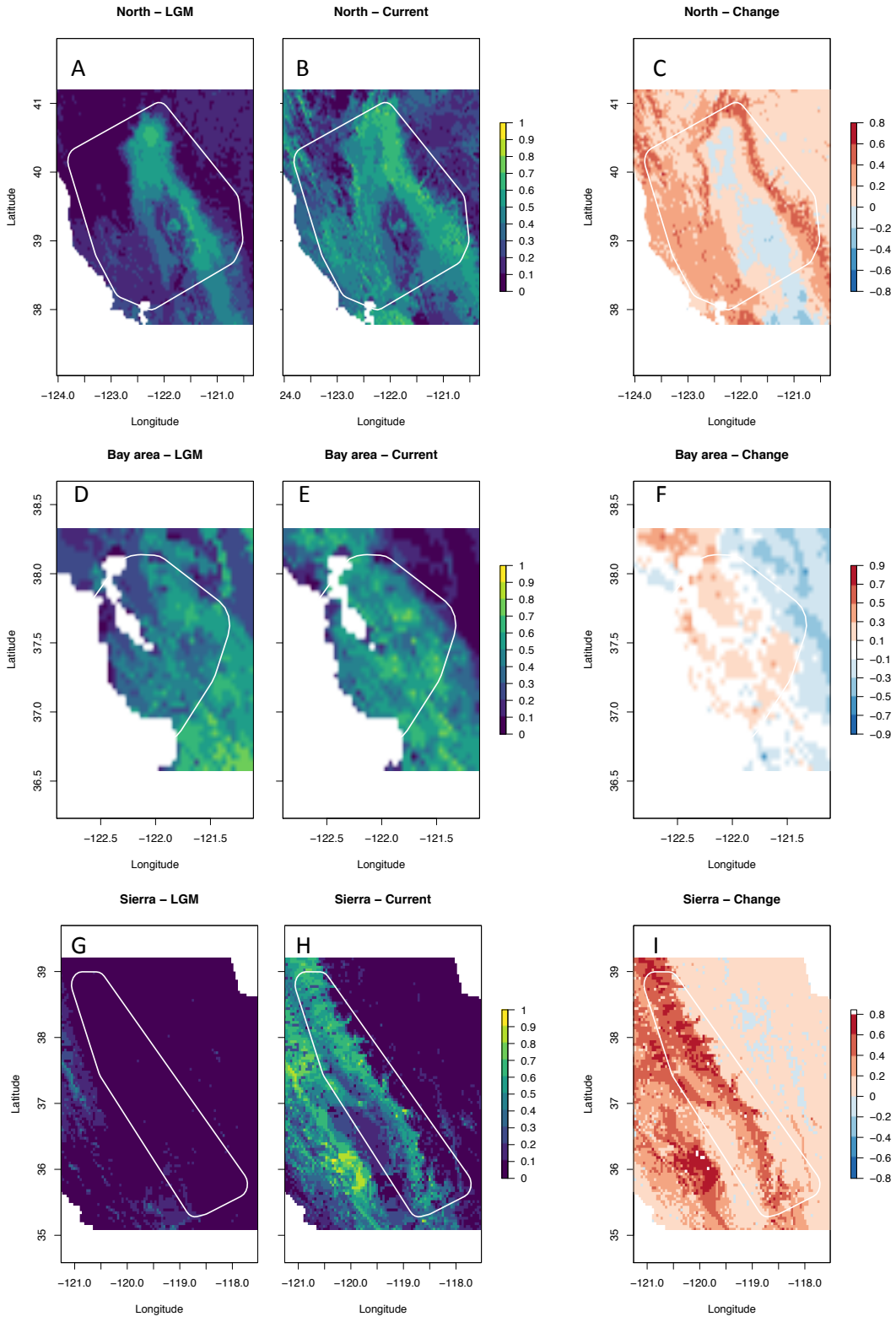

Figure 3.S3 (continued on next page)

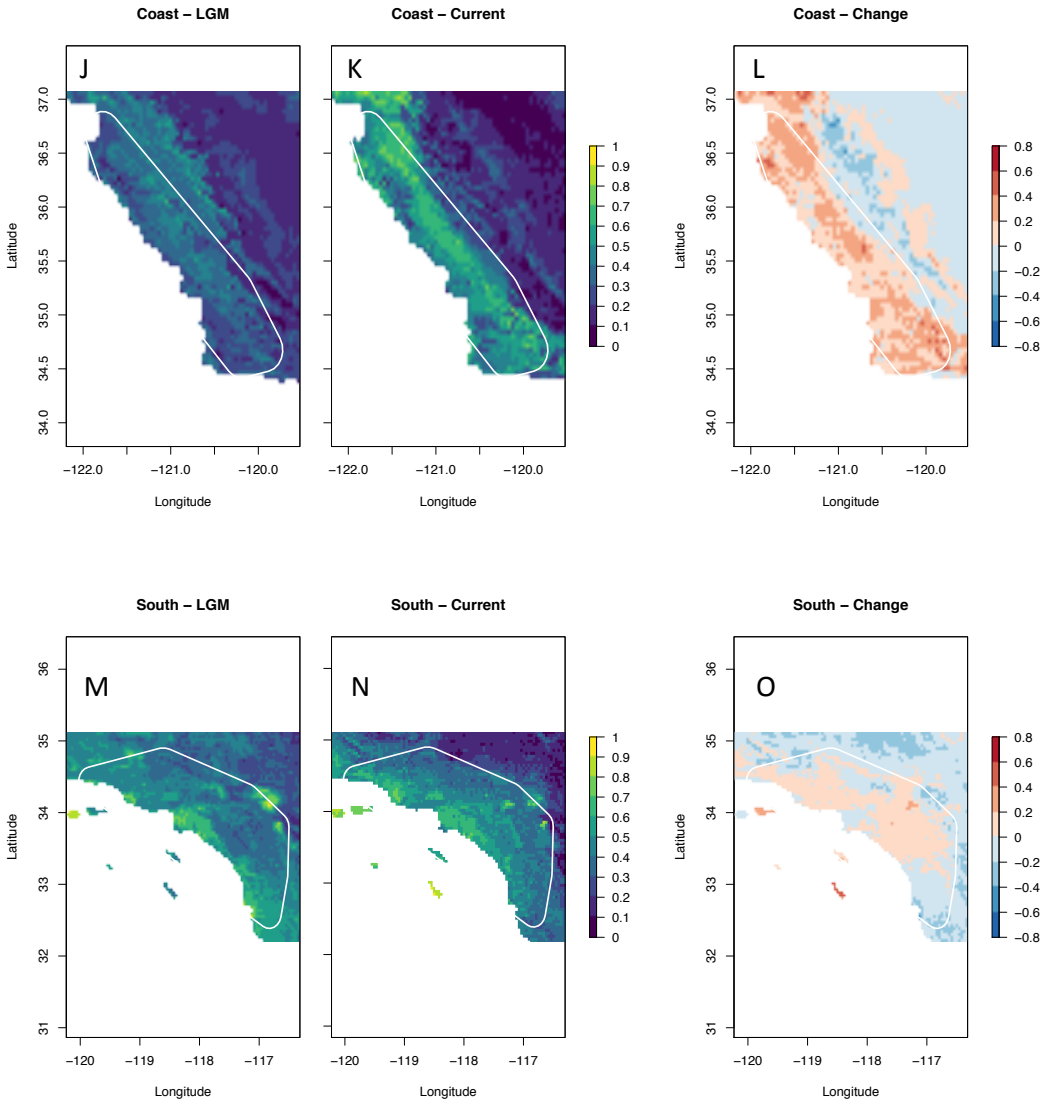

Figure 3.S3: Regional ENMs for historical (A, D, G, J, M) and current (B, E, H, K, N) climate. The difference between current and historical suitability in the regional ENMs (C, F, I, L, O)

Table 3.S1: Sample information for tissues used in this study. ELPA = *E. panamintina.*

Catalog Number	*Location*	*State*	*Country*	*Species*	*Genetic Cluster*
AGC 1279	Inyo	California	USA	*E. coerulea*	-
CAS 208956	Arizona	Arizona	USA	*E. kingii*	-
MVZ 182148	Baja California Norte	Baja California	Mexico	*E. cedrosensis*	-
MVZ 191079	Baja California Sur	Baja California Sur	Mexico	*E. paucicarinata*	-
CAS 205794	Monterey	California	USA	*E. multicarinata*	Coast
CAS 208948	Monterey	California	USA	*E. multicarinata*	Coast
CAS 226091	Monterey	California	USA	*E. multicarinata*	Coast
CAS 241848	Monterey	California	USA	*E. multicarinata*	Coast
EMT 075	Monterey	California	USA	*E. multicarinata*	Coast
EMT 076	Monterey	California	USA	*E. multicarinata*	Coast
EMT 077	Monterey	California	USA	*E. multicarinata*	Coast
RDC 1207	Monterey	California	USA	*E. multicarinata*	Coast
RDC 1208	Monterey	California	USA	*E. multicarinata*	Coast
RDC 1212	Monterey	California	USA	*E. multicarinata*	Coast
RDC 1218	Monterey	California	USA	*E. multicarinata*	Coast
RDC 1219	Monterey	California	USA	*E. multicarinata*	Coast
YS 012	Monterey	California	USA	*E. multicarinata*	Coast
YS 013	Monterey	California	USA	*E. multicarinata*	Coast
YS 014	Monterey	California	USA	*E. multicarinata*	Coast
YS 015	Monterey	California	USA	*E. multicarinata*	Coast
YS 016	Monterey	California	USA	*E. multicarinata*	Coast
CAS 208514	San Luis Obispo	California	USA	*E. multicarinata*	Coast
CAS 208515	San Luis Obispo	California	USA	*E. multicarinata*	Coast
CAS 241826	San Luis Obispo	California	USA	*E. multicarinata*	Coast
CAS 241827	San Luis Obispo	California	USA	*E. multicarinata*	Coast
MVZ 150173	San Luis Obispo	California	USA	*E. multicarinata*	Coast
MVZ 228813	San Luis Obispo	California	USA	*E. multicarinata*	Coast
CAS 241917	Santa Barbara	California	USA	*E. multicarinata*	Coast
MVZ 137539	Santa Barbara	California	USA	*E. multicarinata*	Coast
AGC 1155	Inyo	California	USA	*E. multicarinata*	EastSierra
AGC 1156	Inyo	California	USA	*E. multicarinata*	EastSierra
AGC 1157	Inyo	California	USA	*E. multicarinata*	EastSierra
AGC 1158	Inyo	California	USA	*E. multicarinata*	EastSierra
AGC 1159	Inyo	California	USA	*E. multicarinata*	EastSierra
MVZ 227733	Inyo	California	USA	*E. multicarinata*	EastSierra
AGC 1162	Inyo	California	USA	*E. panamintina*	ELPA

AGC 1224	Inyo	California	USA	*E. panamintina*	ELPA
AGC 1229	Inyo	California	USA	*E. panamintina*	ELPA
MVZ 191076	Inyo	California	USA	*E. panamintina*	ELPA
TC 1577	Inyo	California	USA	*E. panamintina*	ELPA
TC 2063	Inyo	California	USA	*E. panamintina*	ELPA
CAS 208683	Alameda	California	USA	*E. multicarinata*	Monterey
CAS 208964	Alameda	California	USA	*E. multicarinata*	Monterey
CAS 238961	Alameda	California	USA	*E. multicarinata*	Monterey
CAS 243754	Contra Costa	California	USA	*E. multicarinata*	Monterey
MVZ 162061	Contra Costa	California	USA	*E. multicarinata*	Monterey
MVZ 191101	Contra Costa	California	USA	*E. multicarinata*	Monterey
MVZ 191125	Contra Costa	California	USA	*E. multicarinata*	Monterey
CAS 247262	San Francisco	California	USA	*E. multicarinata*	Monterey
CRF 10	San Joaquin	California	USA	*E. multicarinata*	Monterey
CAS 201818	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 203541	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 204814	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 204817	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 208089	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 208947	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 218626	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 218719	San Mateo	California	USA	*E. multicarinata*	Monterey
CAS 223936	San Mateo	California	USA	*E. multicarinata*	Monterey
CRF 170	San Mateo	California	USA	*E. multicarinata*	Monterey
MVZ 230557	Santa Clara	California	USA	*E. multicarinata*	Monterey
CAS 252939	Santa Cruz	California	USA	*E. multicarinata*	Monterey
CAS 252942	Santa Cruz	California	USA	*E. multicarinata*	Monterey
MVZ 191127	Benton	Oregon	USA	*E. multicarinata*	North
CAS 205830	Butte	California	USA	*E. multicarinata*	North
CAS 228313	Douglas	Oregon	USA	*E. multicarinata*	North
MVZ 150172	El Dorado	California	USA	*E. multicarinata*	North
CAS 223707	Glenn	California	USA	*E. multicarinata*	North
MVZ 191137	Glenn	California	USA	*E. multicarinata*	North
MVZ 162062	Humboldt	California	USA	*E. multicarinata*	North
REW 301	Klickitat	Oregon	USA	*E. multicarinata*	North
REW 302	Klickitat	Oregon	USA	*E. multicarinata*	North
CAS 209176	Lake	California	USA	*E. multicarinata*	North
CAS 219450	Lake	California	USA	*E. multicarinata*	North
CAS 219530	Lake	California	USA	*E. multicarinata*	North
CAS 220804	Lake	California	USA	*E. multicarinata*	North
CAS 201252	Mendocino	California	USA	*E. multicarinata*	North

CAS 208852	Mendocino	California	USA	*E. multicarinata*	North
CAS 220770	Mendocino	California	USA	*E. multicarinata*	North
MVZ 162059	Mendocino	California	USA	*E. multicarinata*	North
CAS 241784	Napa	California	USA	*E. multicarinata*	North
CAS 251756	Napa	California	USA	*E. multicarinata*	North
HBS 122003	Napa	California	USA	*E. multicarinata*	North
HBS 122146	Napa	California	USA	*E. multicarinata*	North
HBS 122147	Napa	California	USA	*E. multicarinata*	North
HBS 122149	Napa	California	USA	*E. multicarinata*	North
HBS 122150	Napa	California	USA	*E. multicarinata*	North
HBS 122218A	Napa	California	USA	*E. multicarinata*	North
HBS 123537	Napa	California	USA	*E. multicarinata*	North
HBS 123538	Napa	California	USA	*E. multicarinata*	North
HBS 123540	Napa	California	USA	*E. multicarinata*	North
MVZ 191108	Napa	California	USA	*E. multicarinata*	North
MVZ 191110	Napa	California	USA	*E. multicarinata*	North
MVZ 191113	Napa	California	USA	*E. multicarinata*	North
MVZ 175428	Nevada	California	USA	*E. multicarinata*	North
CAS 206370	Placer	California	USA	*E. multicarinata*	North
MVZ 175432	Placer	California	USA	*E. multicarinata*	North
MVZ 162053	Shasta	California	USA	*E. multicarinata*	North
CAS 202917	Sierra	California	USA	*E. multicarinata*	North
HBS 122139	Solano	California	USA	*E. multicarinata*	North
HBS 122141	Solano	California	USA	*E. multicarinata*	North
HBS 122142	Solano	California	USA	*E. multicarinata*	North
HBS 122206A	Solano	California	USA	*E. multicarinata*	North
HBS 122611	Solano	California	USA	*E. multicarinata*	North
CAS 236169	Sonoma	California	USA	*E. multicarinata*	North
CAS 236534	Sonoma	California	USA	*E. multicarinata*	North
CAS 236558	Sonoma	California	USA	*E. multicarinata*	North
CAS 238471	Sonoma	California	USA	*E. multicarinata*	North
CAS 250301	Sonoma	California	USA	*E. multicarinata*	North
CAS 251453	Sonoma	California	USA	*E. multicarinata*	North
CAS 251457	Sonoma	California	USA	*E. multicarinata*	North
MVZ 162054	Sonoma	California	USA	*E. multicarinata*	North
CAS 206459	Tehama	California	USA	*E. multicarinata*	North
CAS 227441	Tehama	California	USA	*E. multicarinata*	North
CAS 235901	Tehama	California	USA	*E. multicarinata*	North
CAS 252952	Yolo	California	USA	*E. multicarinata*	North
MVZ 161393	Baja California Norte	Baja California	Mexico	*E. multicarinata*	South
HBS 130992	Isla Coronado	Baja California	Mexico	*E. multicarinata*	South

HBS 130993	Isla Coronado	Baja California	Mexico	*E. multicarinata*	South
HBS 130994	Isla Coronado	Baja California	Mexico	*E. multicarinata*	South
HBS 130995	Isla Coronado	Baja California	Mexico	*E. multicarinata*	South
HBS 130996	Isla Coronado	Baja California	Mexico	*E. multicarinata*	South
BG	Los Angeles	California	USA	*E. multicarinata*	South
CAS 227987	Los Angeles	California	USA	*E. multicarinata*	South
CHAR03	Los Angeles	California	USA	*E. multicarinata*	South
CHAR04	Los Angeles	California	USA	*E. multicarinata*	South
CHAR05	Los Angeles	California	USA	*E. multicarinata*	South
CRF 146	Los Angeles	California	USA	*E. multicarinata*	South
CRF 149	Los Angeles	California	USA	*E. multicarinata*	South
EMT 059	Los Angeles	California	USA	*E. multicarinata*	South
EMT 060	Los Angeles	California	USA	*E. multicarinata*	South
EMT 061	Los Angeles	California	USA	*E. multicarinata*	South
EMT 078	Los Angeles	California	USA	*E. multicarinata*	South
EMT 079	Los Angeles	California	USA	*E. multicarinata*	South
EMT 080	Los Angeles	California	USA	*E. multicarinata*	South
EMT 081	Los Angeles	California	USA	*E. multicarinata*	South
EMT 082	Los Angeles	California	USA	*E. multicarinata*	South
HAHA01	Los Angeles	California	USA	*E. multicarinata*	South
HAHA02	Los Angeles	California	USA	*E. multicarinata*	South
HBS 130997	Los Angeles	California	USA	*E. multicarinata*	South
JL1113	Los Angeles	California	USA	*E. multicarinata*	South
JL2015	Los Angeles	California	USA	*E. multicarinata*	South
LACM 145480	Los Angeles	California	USA	*E. multicarinata*	South
SDSU 4071	Los Angeles	California	USA	*E. multicarinata*	South
SDSU 4072	Los Angeles	California	USA	*E. multicarinata*	South
ST	Los Angeles	California	USA	*E. multicarinata*	South
TB SAL1	Los Angeles	California	USA	*E. multicarinata*	South
TB SAL2	Los Angeles	California	USA	*E. multicarinata*	South
TP	Los Angeles	California	USA	*E. multicarinata*	South
CAS 208712	Riverside	California	USA	*E. multicarinata*	South
CAS 228319	Riverside	California	USA	*E. multicarinata*	South
JLG701	Riverside	California	USA	*E. multicarinata*	South
JLG702	Riverside	California	USA	*E. multicarinata*	South
JLG703	San Bernardino	California	USA	*E. multicarinata*	South
JLG704	San Bernardino	California	USA	*E. multicarinata*	South
JLG705	San Bernardino	California	USA	*E. multicarinata*	South
HBS 130982	San Diego	California	USA	*E. multicarinata*	South
HBS 130983	San Diego	California	USA	*E. multicarinata*	South
HBS 130984	San Diego	California	USA	*E. multicarinata*	South

HBS 130985	San Diego	California	USA	*E. multicarinata*	South
HBS 130986	San Diego	California	USA	*E. multicarinata*	South
HBS 130987	San Diego	California	USA	*E. multicarinata*	South
HBS 130988	San Diego	California	USA	*E. multicarinata*	South
HBS 130989	San Diego	California	USA	*E. multicarinata*	South
HBS 130990	San Diego	California	USA	*E. multicarinata*	South
HBS 130991	San Diego	California	USA	*E. multicarinata*	South
MVZ 162058	San Diego	California	USA	*E. multicarinata*	South
MVZ 230919	San Diego	California	USA	*E. multicarinata*	South
MVZ 162381	Santa Barbara	California	USA	*E. multicarinata*	South
MVZ 162066	Calaveras	California	USA	*E. multicarinata*	WestSierra
MVZ 175287	Calaveras	California	USA	*E. multicarinata*	WestSierra
CAS 234675	El Dorado	California	USA	*E. multicarinata*	WestSierra
CAS 235942	El Dorado	California	USA	*E. multicarinata*	WestSierra
CAS 236556	El Dorado	California	USA	*E. multicarinata*	WestSierra
CAS 208825	Fresno	California	USA	*E. multicarinata*	WestSierra
CAS 224866	Fresno	California	USA	*E. multicarinata*	WestSierra
CAS 235937	Fresno	California	USA	*E. multicarinata*	WestSierra
AGC 1151	Kern	California	USA	*E. multicarinata*	WestSierra
AGC 1152	Kern	California	USA	*E. multicarinata*	WestSierra
AGC 1153	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 206440	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 220919	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 220920	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 236212	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 247436	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 247437	Kern	California	USA	*E. multicarinata*	WestSierra
KLW 69	Kern	California	USA	*E. multicarinata*	WestSierra
MVZ 137822	Kern	California	USA	*E. multicarinata*	WestSierra
MVZ 137823	Kern	California	USA	*E. multicarinata*	WestSierra
MVZ 137826	Kern	California	USA	*E. multicarinata*	WestSierra
MVZ 137827	Kern	California	USA	*E. multicarinata*	WestSierra
CAS 212976	Madera	California	USA	*E. multicarinata*	WestSierra
CAS 212986	Madera	California	USA	*E. multicarinata*	WestSierra
CAS 205780	Mariposa	California	USA	*E. multicarinata*	WestSierra
CAS 209201	Mariposa	California	USA	*E. multicarinata*	WestSierra
MVZ 137825	Mariposa	California	USA	*E. multicarinata*	WestSierra
MVZ 243330	Mariposa	California	USA	*E. multicarinata*	WestSierra
MVZ 243333	Merced	California	USA	*E. multicarinata*	WestSierra
CAS 219611	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 220864	Tulare	California	USA	*E. multicarinata*	WestSierra

CAS 220884	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 220908	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 223541	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 223578	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 235936	Tulare	California	USA	*E. multicarinata*	WestSierra
CAS 236172	Tulare	California	USA	*E. multicarinata*	WestSierra
MVZ 137828	Tulare	California	USA	*E. multicarinata*	WestSierra
MVZ 230094	Tulare	California	USA	*E. multicarinata*	WestSierra

REFERENCES

Avise, J.C. (2000). Phylogeography: the history and formation of species (Harvard University Press).

Calsbeek, R., Thompson, J.N., and Richardson, J.E. (2003). Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Mol. Ecol. *12*, 1021–1029.

Caye, K., Deist, T.M., Martins, H., Michel, O., and François, O. (2016). TESS3: fast inference of spatial population structure and genome scans for selection. Mol. Ecol. Resour. *16*, 540–548.

Coyne, J.A., and Orr, H.A. (1998). The evolutionary genetics of speciation. Philos. Trans. R. Soc. B Biol. Sci. *353*, 287–305.

Crispo, E., Bentzen, P., Reznick, D.N., Kinnison, M.T., and Hendry, A.P. (2006). The relative influence of natural selection and geography on gene flow in guppies. Mol. Ecol. *15*, 49–62.

Cushman, S.A., and McGarigal, K. (2002). Hierarchical, Multi-scale decomposition of species-environment relationships. Landsc. Ecol. *17*, 637–646.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. Numer. Math. *1*, 269–271.

Duforet-Frebourg, N., and Slatkin, M. (2016). Isolation–By–Distance–and–Time in a stepping–stone model. Theor. Popul. Biol. *108*, 24–35.

Eaton, D.A.R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics *30*, 1844–1849.

Etten, J. van (2018). gdistance: Distances and Routes on Geographical Grids.

Feldman, C.R., and Spicer, G.S. (2006). Comparative phylogeography of woodland reptiles in California: repeated patterns of cladogenesis and population expansion. Mol. Ecol. *15*, 2201–2222.

Fitch, H.S. (1934). New alligator lizards from the Pacific coast. Copeia *1934*, 6–7.

Fitch, H.S. (1938). A systematic account of the alligator lizards (*Gerrhonotus*) in the western United States and lower California (University Press).

Garant, D., Kruuk, L.E.B., Wilkin, T.A., McCleery, R.H., and Sheldon, B.C. (2005). Evolution driven by differential dispersal within a wild bird population. Nature *433*, 60–65.

Glenn, T.C., Bayona-Vasquez, N.J., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A., Bentley, K.E., Finger, J.W., Watson, P.R., Louha, S., et al. (2017). Adapterama III: Quadruple-indexed, triple-enzyme RADseq libraries for about $1USD per Sample (3RAD).

Hijmans, R.J., Etten, J. van, Sumner, M., Cheng, J., Bevan, A., Bivand, R., Busetto, L., Canty, M., Forrest, D., Ghosh, A., et al. (2019). raster: Geographic Data Analysis and Modeling.

Holderegger, R., and Wagner, H.H. (2008). Landscape Genetics. BioScience *58*, 199–207.

Jenkins, D.G., Carey, M., Czerniewska, J., Fletcher, J., Hether, T., Jones, A., Knight, S., Knox, J., Long, T., Mannino, M., et al. (2010). A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? Ecography *33*, 315–320.

Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics *27*, 3070–3071.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P., and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. Sci. Data *4*, 170122.

Lapointe, F., and Rissler, L.J. (2005). Congruence, Consensus, and the Comparative Phylogeography of Codistributed Species in California. Am. Nat. *166*, 290–299.

Leaché, A.D., Banbury, B.L., Felsenstein, J., Oca, A.N.-M. de, and Stamatakis, A. (2015). Short Tree, Long Tree, Right Tree, Wrong Tree: New Acquisition Bias Corrections for Inferring SNP Phylogenies. Syst. Biol. syv053.

Leavitt, D.H., Marion, A.B., Hollingsworth, B.D., and Reeder, T.W. (2017). Multilocus phylogeny of alligator lizards (*Elgaria*, Anguidae): Testing mtDNA introgression as the source of discordant molecular phylogenetic hypotheses. Mol. Phylogenet. Evol. *110*, 104–121.

Lee, C.-R., and Mitchell-Olds, T. (2011). Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. Mol. Ecol. *20*, 4631–4642.

Losos, J.B., and Glor, R.E. (2003). Phylogenetic comparative methods and the geography of speciation. Trends Ecol. Evol. *18*, 220–227.

Manel, S., Schwartz, M.K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. *18*, 189–197.

Maniatas, T., Fristsch, E.F., and Sambrook, J. (1982). Molecular cloning: a laboratory manual (Cold Spring Harbour, New York: Cold Spring Harbor Laboratory).

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *17*, 10–12.

Mayr, E. (1963). Animal species and evolution Belknap Press of Harvard University Press. Camb. MA.

McCartney-Melstad, E., Gidiş, M., and Shaffer, H.B. (2018). Population genomic data reveal extreme geographic subdivision and novel conservation actions for the declining foothill yellow-legged frog. Heredity *121*, 112–125.

McRae, B.H. (2006). Isolation by Resistance. Evolution *60*, 1551–1561.

Nosil, P., Vines, T.H., and Funk, D.J. (2005). Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution *59*, 705–719.

Nychka, D., Furrer, R., Paige, J., and Sain, S. (2017). fields: Tools for spatial data.

Ortego, J., Riordan, E.C., Gugger, P.F., and Sork, V.L. (2012). Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. Mol. Ecol. *21*, 3210–3223.

Peterman, W.E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. Methods Ecol. Evol. *9*, 1638–1647.

Peterman, W.E., Winiarski, K.J., Moore, C.E., Carvalho, C. da S., Gilbert, A.L., and Spear, S.F. (2019). A comparison of popular approaches to optimize landscape resistance surfaces. Landsc. Ecol.

Phillips, S.J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography *31*, 161–175.

Rousset, F. (1997). Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. Genetics *145*, 1219–1228.

Searcy, C.A., and Shaffer, H.B. (2014). Field validation supports novel niche modeling strategies in a cryptic endangered amphibian. Ecography *37*, 983–992.

Searcy, C.A., and Shaffer, H.B. (2016). Do Ecological Niche Models Accurately Identify Climatic Determinants of Species Ranges? Am. Nat. *187*, 423–435.

Sexton, J.P., McIntyre, P.J., Angert, A.L., and Rice, K.J. (2009). Evolution and ecology of species range limits. Annu. Rev. Ecol. Evol. Syst. *40*, 415–436.

Sexton, J.P., Hangartner, S.B., and Hoffmann, A.A. (2014). Genetic Isolation by Environment or Distance: Which Pattern of Gene Flow Is Most Common? Evolution *68*, 1–15.

Slatkin, M. (1985). Gene flow in natural populations. Annu. Rev. Ecol. Syst. *16*, 393–430.

Sork, V., and Waits, L. (2010). Contributions of landscape genetics–approaches, insights, and future potential. Mol. Ecol. *19*, 3489–3495.

Spinks, P.Q., and Shaffer, H.B. (2005). Range-wide molecular analysis of the western pond turtle (*Emys marmorata*): cryptic variation, isolation by distance, and their conservation implications. Mol. Ecol. *14*, 2047–2064.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Stebbins, R.C. (2003). A Field Guide to Western Reptiles and Amphibians (Houghton Mifflin Harcourt).

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics *123*, 585–595.

Telemeco, R.S. (2014). Here be dragons: Functional analyses of thermal adaptation and biogeography of reptiles in a changing world.

Thorpe, R.S., Surget-Groba, Y., and Johansson, H. (2008). The relative importance of ecology and geographic isolation for speciation in anoles. Philos. Trans. R. Soc. B Biol. Sci. *363*, 3071–3081.

Title, P.O., and Bemmels, J.B. (2018). ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography *41*, 291–307.

Trumbo, D.R., Spear, S.F., Baumsteiger, J., and Storfer, A. (2013). Rangewide landscape genetics of an endemic Pacific northwestern salamander. Mol. Ecol. *22*, 1250–1266.

US Geological Survey (2013). USGS NED n36w121 1/3 arc-second 2013 1 x 1 degree IMG (U.S. Geological Survey).

Wang, I.J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution *67*, 3403–3411.

Wang, I.J., and Bradburd, G.S. (2014). Isolation by environment. Mol. Ecol. *23*, 5649–5662.

Wang, I.J., and Summers, K. (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. Mol. Ecol. *19*, 447–458.

Wang, I.J., Glor, R.E., and Losos, J.B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol. Lett. *16*, 175–182.

Wang, Y.-H., Yang, K.-C., Bridgman, C.L., and Lin, L.-K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. Landsc. Ecol. *23*, 989–1000.

Warren, D.L., Glor, R.E., and Turelli, M. (2008). Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. Evolution *62*, 2868–2883.

Warren, D.L., Glor, R.E., and Turelli, M. (2010). ENMTools: a toolbox for comparative studies of environmental niche models. Ecography.

Wiens, J.A. (1989). Spatial Scaling in Ecology. Funct. Ecol. *3*, 385–397.

Wright, S. (1943). Isolation by distance. Genetics *28*, 114.