American Genetic Association

Journal of Heredity, 2020, 613–627
doi:10.1093/jhered/esaa049
Original Article
Advance Access publication November 27, 2020

OXFORD

Original Article

Conservation Genomics of the Threatened Western Spadefoot, *Spea hammondii*, in Urbanized Southern California

Kevin M. Neal, Robert N. Fisher, Milan J. Mitrovich, and H. Bradley Shaffer

From the Department of Ecology and Evolutionary Biology, and La Kretz Center for California Conservation Science, University of California Los Angeles, Los Angeles, CA 90095 (Neal and Shaffer); Western Ecological Research Center, U.S. Geological Survey, San Diego, CA 92101 (Fisher); and Natural Communities Coalition, Irvine, CA 92602 (Mitrovich).

Address correspondence to H. B. Shaffer at the address above, or e-mail: brad.shaffer@ucla.edu.

Received December 9, 2019; First decision February 17, 2020; Accepted November 19, 2020.

Corresponding Editor: Ryan Garrick

Abstract

Populations of the western spadefoot (*Spea hammondii*) in southern California occur in one of the most urbanized and fragmented landscapes on the planet and have lost up to 80% of their native habitat. Orange County is one of the last strongholds for this pond-breeding amphibian in the region, and ongoing restoration efforts targeting *S. hammondii* have involved habitat protection and the construction of artificial breeding ponds. These efforts have successfully increased breeding activity, but genetic characterization of the populations, including estimates of effective population size and admixture between the gene pools of constructed artificial and natural ponds, has never been undertaken. Using thousands of genome-wide single-nucleotide polymorphisms, we characterized the population structure, genetic diversity, and genetic connectivity of spadefoots in Orange County to guide ongoing and future management efforts. We identified at least 2, and possibly 3 major genetic clusters, with additional substructure within clusters indicating that individual ponds are often genetically distinct. Estimates of landscape resistance suggest that ponds on either side of the Los Angeles Basin were likely interconnected historically, but intense urban development has rendered them essentially isolated, and the resulting risk of interruption to natural metapopulation dynamics appears to be high. Resistance surfaces show that the existing artificial ponds were well-placed and connected to natural populations by low-resistance corridors. Toad samples from all ponds (natural and artificial) returned extremely low estimates of effective population size, possibly due to a bottleneck caused by a recent multi-year drought. Management efforts should focus on maintaining gene flow among natural and artificial ponds by both assisted migration and construction of new ponds to bolster the existing pond network in the region.

Subject area: Conservation genomics and biodiversity
Key words: amphibian conservation, landscape genetics, RAD-seq, *Spea*, landscape resistance

The decline of amphibians is a global conservation crisis, and habitat loss and fragmentation are major contributing factors (Cushman 2006; Hamer and McDonnell 2008). However, the degree of fragmentation and its effects on populations vary substantially at regional and local scales (Marsh and Trenham 2001; Grant et al. 2016). This variability limits generalizability for conservation and management of amphibians at the finer scales of detection and emphasizes the importance of determining appropriate management units across scales. Landscape genetics provides a useful framework for such fine-scale conservation efforts by quantifying historical isolation and functional connectivity among populations and the role of landscape features in shaping those patterns (Manel et al. 2003; Storfer et al. 2007). The role of environmental variables in the isolation and connectivity of populations can be inferred by comparing landscape resistance of environmental variables using spatially explicit models and data layers.

Pond-breeding amphibians are particularly attractive targets for landscape genetic analysis. Breeding ponds function as discrete natural subunits for analysis using spatially explicit genetic models, and larvae usually represent a single breeding season cohort of same-age individuals. Although adults are often exclusively nocturnal and only active during heavy rain events, tadpoles and larvae are often concentrated in large numbers at breeding ponds and are straightforward to sample for genetics with minimal harm to individuals (Beebee 2005; Polich et al. 2013). For the same reasons, pond-breeding amphibians are uniquely vulnerable to local habitat disruption: their reliance on ephemeral aquatic breeding sites and generally limited dispersal abilities means that the destruction or isolation of a few key ponds can reduce functional connectivity across a largely unsuitable habitat matrix by eliminating habitat stepping stones (Willson and Hopkins 2013; Unglaub et al. 2015; Zamberletti et al. 2018), potentially leading to the collapse of preexisting metapopulation dynamics. Identifying, maintaining, and restoring historical levels of connectivity is a key way in which landscape genetics can and should inform amphibian conservation biology, particularly at the fine-scale level where most conservation actions occur (McCartney-Melstad et al. 2018b).

Southern California is host to a number of sensitive, threatened, and endangered amphibians (Thomson et al. 2016). Among them, the western spadefoot, *Spea hammondii,* stands out as a key species for management. Vernal pools utilized by the species occur in grasslands, coastal sage scrub, oak woodlands, and chaparral. These areas have been decimated by urban and agricultural development, and as a consequence, the western spadefoot is now extirpated across much of its range in southern California (Thomson et al. 2016). Noting this dramatic decline, the California Department of Fish and Wildlife identified the western spadefoot as a Species of Special Concern (SSC) in both its original 1994 (Jennings and Hayes 1994) and most recent (Thomson et al. 2016) assessments, and the species is currently under review by the U.S. Fish and Wildlife Service for listing under the federal Endangered Species Act (U.S. Fish and Wildlife Service 2015). The western spadefoot is also a Species of Interest in the Central Subarea and a fully covered species in the Coastal Subarea under the Natural Community Conservation Plan (NCCP) and Habitat Conservation Plan (HCP) for the County of Orange Central and Coastal Subregion (County of Orange 1996), where it is managed as if it were a listed species under the federal and California Endangered Species Acts. It is in this critical region for the southern *S. hammondii* clade (Neal et al. 2019) that our study took place. The core of the study area is a roughly 15 400-hectare open space managed for both wildlife protection and human recreation. The scale of the study area presents a unique opportunity to focus landscape and population genomic inference at a spatial scale appropriate for many amphibian species: within and among vernal pools separated by hundreds to a few thousand meters (Baumberger et al. 2019). In addition, a significant *S. hammondii* restoration effort within Orange County has involved the construction of seasonal ponds that were stocked with spadefoot tadpoles in 2005 and 2006. Although these ponds have existed for several *S. hammondii* generations, follow-up genetic monitoring of the constructed ponds and surrounding natural breeding sites has never been undertaken (Baumberger et al. 2020).

Management of *S. hammondii* is complicated both by the elusive life history of the species and the repeated droughts that increasingly occur in southern California. As with many other pond-breeding amphibians, direct observations of adult *S. hammondii* are rare. Like all members of the genus, *S. hammondii* dig burrows up to 1 meter deep using their keratinous hindfoot spade, and they estivate in these retreats until seasonal rains trigger their emergence (Morey and Reznick 2004). The few existing observations of *S. hammondii* movements indicate very strong philopatry of adults to their breeding ponds, but these observations are extremely limited and may not represent the normal range of species movement dynamics (Baumberger et al. 2019). Genetic sampling of tadpoles has been suggested as the best way to determine patterns of gene flow and migration among ponds, and genetic assessment of the degree of philopatry and subsequent population structure has been highlighted as an important conservation need (Thomson et al. 2016).

Our goal in this study was to provide a detailed landscape genomic analysis of *S. hammondii* that will simultaneously inform scientists about the basic landscape ecology of the species and help guide its management in a region with active and ongoing conservation efforts. Although studies like ours are by definition inferential and somewhat exploratory, they can fill in critical gaps in our knowledge that are essential for effective management. We used restriction site-associated DNA sequencing (RAD-seq) to generate a large genome-wide dataset for nearly all documented *S. hammondii* breeding sites in the largely protected lands of Orange County. Given the fine spatial scale of this study and the inability of a previous effort with limited multi-locus genetic sampling (5 nuclear loci) to resolve population-level relationships in *S. hammondii* (Neal et al. 2018), we sequenced thousands of markers to increase the precision and resolution of our genetic assessment (Andrews et al. 2016; McCartney-Melstad et al. 2018a). We analyzed our RAD-seq data to address several interrelated goals. First, we sought to identify whether spadefoots at this spatial scale segregate into distinct genetic clusters that could be considered potential conservation units. Second, we determined the relative genetic health, as quantified by levels of genetic diversity and estimates of effective population sizes of gene pools within ponds and the correlation of genetic diversity measures with pond area to help guide the construction of new ponds. Third, we aimed to quantify the association of geographic and environmental distance with genetic distance and to identify potential corridors for conservation prioritization. Finally, we assessed the genetic health of a set of constructed artificial ponds, their diversity levels relative to nearby natural ponds, and the extent of admixture between the 2 that would indicate whether natural gene flow has been reestablished. We discuss how these findings can inform the management of the western spadefoot across this critical conservation landscape.

Methods

Study Area and Sampling

The study area comprises a roughly 15 400-hectare open space in central Orange County, California, with a few sample areas further south in Orange County. The study region is bounded by the coastal San Joaquin Hills to the southwest and the inland Santa Ana Mountains to the east/northeast, with the southeastern extent of the Los Angeles Basin (i.e., the Orange County Coastal Plain) in between (Figure 1). Spadefoots were historically observed within the Los Angeles Basin as recently as the 1960s (GBIF.org 2017), but the area is now urbanized to the extent that spadefoot localities occur almost exclusively on isolated uplands and ridgetops adjacent to the developed lowlands (Thomson et al. 2016).

We sampled up to 30 tadpoles from 26 sites within Orange County (Figure 1) between 2015 and 2017, with the majority in 2017. Except for the constructed ponds at Irvine Mesa (IM) and Shoestring Canyon (SHOE), sampled ponds are assumed to have been colonized naturally, even though some are human-modified. We refer to IM and SHOE as "artificial" and naturally colonized sites as "natural." Most localities sampled for this study are situated on protected lands collectively known as the Nature Reserve of Orange County and are managed by a consortium of public and private landowners. Our field work targeted known localities plus potential new sites based on Google Earth satellite imagery. Tadpoles were sampled using dipnets at different locations along the perimeter of each pond. For large tadpoles, roughly 1 cm of the tail tip was removed and the tadpole released at point of capture; for smaller tadpoles, whole individuals were sometimes sacrificed. Tissues were stored in 95% ethanol in the field and then in −20 °C freezers until DNA extraction. We also utilized a much smaller set of tissues from adults and tadpoles collected serendipitously, dating back to 1991. Tissues collected prior to 2015 are from 3 sites only: TENA, SANCAN, and TORO. All collections were conducted via IACUC-approved protocols.

DNA Extraction and Sequencing

We extracted genomic DNA from up to 10 randomly subsampled individuals from each site using either a standard salt extraction or the Kingfisher extraction protocol (Thermoscientific), followed by the 3RAD RAD-seq protocol (Bayona-Vásquez et al. 2019) to isolate genome-wide fragments of DNA for sequencing. 3RAD adds a third enzyme to ddRAD (Peterson et al. 2012) that cuts adapter dimers. We used Illumina iTru adapters as well as custom internal adapters (Integrated DNA Technologies) for multiplexing, and PstI-HF (dimer cutter), NsiI-HF (rare cutter), and MspI (common cutter) (New England BioLabs) for the 3 restriction enzymes. Libraries were pooled with 200–300 ng of DNA per individual and we used BluePippin (Sage Science) to size select 350–450 base pair fragments. Pooled

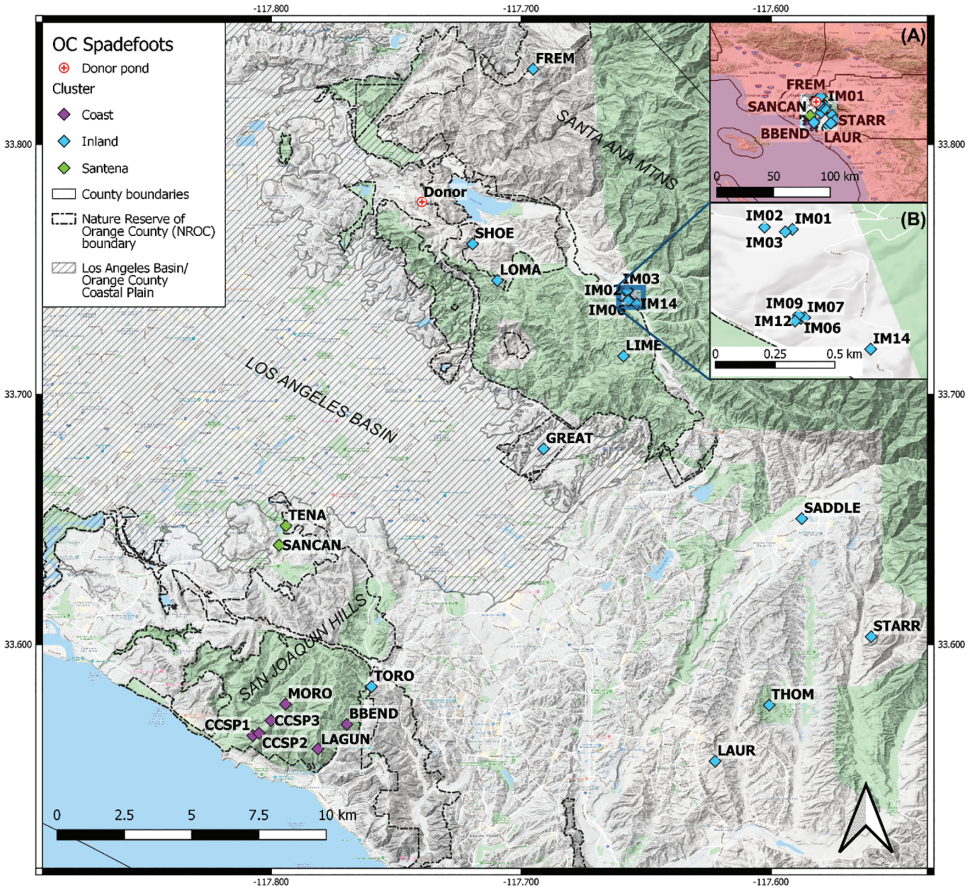

Figure 1. Map of study area with sample sites colored by FastStructure genetic cluster membership for 3 hierarchically determined clusters. Dot-dash outlines mark the boundary of the Nature Reserves of Orange County. IM and SHOE ponds are artificial and were stocked with individuals from donor ponds at the marked site. Inset **A** shows the broader context of the study area in southern California. Inset **B** zooms in on the blue box to show the artificial ponds at Irvine Mesa (IM). Green-shaded areas are public lands on Google Terrain Hybrid basemap.

libraries were sent to the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley, for sequencing on an Illumina HiSeq 4000 with either PE150, SR100, or SR150, although we used only the higher-quality first reads (R1's) from the paired end (PE150) data for downstream protocols.

DNA Post-Processing

We first trimmed the internal adapters from the demultiplexed sequences using cutadapt (Martin 2011). As an additional filtering step, we used fastp (Chen et al. 2018) to filter reads below 50% complexity (percentage of bases that are different from the next base in a sequence) and with a mean Phred quality score below Q30 (base call accuracy of 99.9%). To capture the highest-quality portions at the start of each read and because of different read lengths among SR100 and SR150/PE150 sequencing runs, every read was trimmed to 60 base pairs, and reads less than 30 base pairs in length were removed. We used ipyrad 0.7.29 (Eaton 2015, 2019; Eaton 2014) to assemble the RAD loci using a reference draft genome of a *S. hammondii* metamorph from Santa Barbara County, California; this individual is a part of the northern clade of *S. hammondii*, the sister clade of the southern *S. hammondii* that is the focus of the current study (Neal et al. 2018). The genome was sequenced by 10X Genomics, assembled using Supernova (10X Genomics) and analyzed using BUSCO (Simão et al. 2015) to evaluate the quality of the assembly (Supplementary Table S1).

From the ipyrad VCF file output, we used vcftools 1.1.5 (Danecek et al. 2011) to produce 2 genetic datasets for Orange County spadefoots based on different subsets of individuals. OC_all_208 is the complete dataset, including a maximum of 10 individuals from all sampled ponds. OC_reduced_148 reduced the sampling of the 9 artificial ponds from 79 total individuals to 19 by including 2–3 individuals from each artificial pond rather than 10; given that the spadefoots in all 9 artificial ponds were derived from salvaged tadpoles from a single group of closely adjacent donor ponds (Glenn Lukos Associates, Inc. 2005), this subsampling was an attempt to reduce the overrepresentation of that set of donor individuals and roughly equalize sampling between natural and artificial ponds. For each dataset, we removed singleton SNPs to filter noise from the dataset (Linck and Battey 2019), included only biallelic SNPs, and retained only the first SNP in each RAD locus to reduce the effects of physical linkage among markers. We explored missing data thresholds from 20% to 75% in clustering analyses. Because we found no major differences in optimal cluster composition, we allowed up to 75% missing data to maximize the number of independent SNPs for pairwise and intrapopulation statistics. Additional parameters used in ipyrad and vcftools to produce each final VCF dataset can be found in scripts in the included figshare repository.

Data Analysis: Population Structure

We evaluated population genetic structure of the OC_reduced_148 dataset using the Bayesian clustering method implemented in FastStructure 1.0 (Raj, Stephens, and Pritchard 2014). The VCF file was first converted to a STRUCTURE file using PGDSpider (Lischer and Excoffier 2012). We ran FastStructure using the simple prior with K values from 1 to 20, repeating each K value 10 times with different random seeds. We used the included chooseK.py or StructureSelector (Li and Liu 2018) and selected the optimal K value based on the value that maximized the marginal likelihood. To identify hierarchical genetic structure, we split the dataset into its component clusters based on the optimal K, and reran FastStructure on each cluster using K from 1 to 10 and 10 random seeds, stopping when K = 1 had the highest marginal likelihood or when K equaled the number of sampling localities (ponds) in the subset. We also ran FastStructure with 10 cross-validation iterations with K from 1 to 20 to examine clustering at K's above the optimal value.

To compare the model-based FastStructure to a model-free ordination approach, we ran a principal components analysis (PCA) in R 3.5.1 (R Core Team 2018) using adegenet::glPca (Jombart 2008; Jombart and Ahmed 2011; Jombart et al. 2018) to visualize genetic structure along the first 8 principal components. We additionally ran fineRADstructure (Malinsky et al. 2018), an extension of finestructure (Lawson et al. 2012) for RAD-seq data, as an alternative examination of hierarchical genetic structure. FineRADstructure utilizes full haplotype sequence data in the alleles.loci file generated by ipyrad to create a matrix of individual pairwise coancestry. We ran FineRADstructure to assign individuals to populations with 1 000 000 MCMC generations after 1 000 000 generations of burn-in and constructed a population tree with 100 000 tree-building iterations, with the tree demonstrating hierarchical clustering based on coancestry. We calculated global F-statistics as a summary measure of population structure with hierfstat::fstat (Goudet and Jombart 2015).

Data Analysis: Genetic Diversity, Effective Population Size, and Their Relationship to Pond Size

Pond size has been implicated as an important determinant of effective population size in pond-breeding salamanders (Wang et al. 2011; McCartney-Melstad et al. 2018b), and we tested for this same potential effect at spadefoot breeding sites. We ran the OC_all_208 VCF in the "populations" module of Stacks 2.3d (Catchen et al. 2013) to calculate pond-level and genetic cluster-level observed heterozygosity (H_O), expected heterozygosity (H_E), nucleotide diversity (*pi*), and inbreeding coefficient (F_{IS}), including invariant sites in the calculations. We required that loci be present in at least 2 populations, be present in at least 50% of individuals within a population and in 25% of all individuals, and that minor allele counts be ≥2 to remove singletons. We calculated effective population size (N_e) for individual ponds using the linkage disequilibrium method (LD N_e) without singletons, and effective number of breeders (N_{eb}) using the molecular coancestry method (Nomura 2008) in NeEstimator 2.1 (Do et al. 2014; Waples et al. 2016); the R package radiator (Gosselin 2017, 2019) was used to convert the VCF to GENEPOP format as input for NeEstimator. We estimated maximum pond surface area from historical satellite imagery in Google Earth based on our visual assessment of the water perimeter in March 2011, when ponds are normally at maximum capacity near the end of California's rainy season. We then calculated correlations as Pearson's r between measures of genetic diversity and the logarithm of pond area using scipy 1.18 for all ponds with 3 or more DNA samples. We excluded SANCAN and TENA from these correlation measures because the 2 water bodies coalesced into a single pond in the March 2011 satellite imagery. SANCAN is also an artificial reservoir, and our samples for both SANCAN and TENA are much older than for the other sites (1995 vs. 2011–2017).

Data Analysis: Genetic Distance and Landscape Resistance

We used hierfstat::genet.dist to calculate pairwise Nei's D_a (Nei et al. 1983; Takezaki and Nei 1996) between individuals, sites, and clusters using the OC_all_208 data set; we also used hierfstat::pairwise.

WCfst to calculate pairwise Weir and Cockerham's FST (Weir and Cockerham 1984) between sites and clusters, and mmod::pairwise_D (Winter 2012; Winter et al. 2017) to calculate pairwise Jost's D (Jost 2008; Jost et al. 2018). StAMPP::stamppPhylip (Pembleton et al. 2013; Pembleton 2017) was used to convert the FST, Jost's D, and Nei's D_a distance matrices to phylip distance files, and these distances were visualized as NeighborNet networks (Bryant and Moulton 2004) in SplitsTree 4.14.8 (Huson and Bryant 2006). Pairwise FST, Jost's D, and Nei's D_a are available as Supplementary Tables S2–S4, respectively.

An important goal of this project was to quantify the effects of local environmental variation in hindering or facilitating gene flow. Given that spadefoots are a rain- and pond-dependent amphibian that burrows in soil, we chose common bioclimatic variables that quantify seasonal aspects of precipitation and temperature, as well as several soil and landcover layers, particularly those associated with the thickness of the topsoil (depth to bedrock). Although we lacked the ecological data to articulate specific hypotheses, we predicted that rainfall, slope, and depth to bedrock would be important predictors of low resistance to gene flow. We chose ResistanceGA (Peterman 2018) over other landscape resistance methods as our analysis tool of choice because of its demonstrated accuracy in optimizing the parameterization of resistance surfaces compared to other methods (Peterman et al. 2019). ResistanceGA takes a pairwise genetic distance matrix and a series of environmental rasters or surfaces as input and uses a genetic algorithm to transform the surfaces into resistance surfaces, using linear mixed effect models to iteratively optimize the fit of resistance distance to genetic distance; the program returns a ranking of the optimized resistance surfaces. Using ResistanceGA removes the subjectivity inherent in expert opinion to determine resistance values of an environmental variable and allows for exploration of a wide range of parameter space (Peterman et al. 2014; Peterman 2018). We limited our analysis to natural sites with 3 or more individuals, removing artificial sites from the input files since they are relatively new and may not be in migration-drift equilibrium. We also removed CCSP2 because it is in the same raster cell as CCSP1. By using only natural sites, we sought to estimate landscape resistance as it occurred before artificial sites were constructed. This assumes that little to no gene flow has occurred between natural and artificial sites since the artificial sites were constructed, and therefore that the historical relationship between landscape features and movement is still discernable among natural ponds. Although we were unable to empirically test this assumption due to the extremely limited genetic samples available from before the construction of the artificial sites, we consider this assumption reasonable given the hilly terrain and distance between artificial sites and known natural sites. We calculated resistance distance between natural sites using gdistance::commuteDistance, which determines distance as an average of a number of random walks between points on the environmental surface. Pairwise Nei's D_a, calculated from the OC_reduced_148 data matrix but restricted to distances between the natural sites was used for the genetic distance input. We opted for Nei's D_a given the demonstrated superiority of nucleotide distance measures over pairwise F-statistics for identifying relationships with geographic distance (Séré et al. 2017), although all of our measures of pairwise genetic divergence were highly correlated. We ran ResistanceGA using 2 sets of environmental variables of differing spatial resolutions, and used a broad, inclusive set of variables based on our limited a priori assumptions of variable importance. The first set of layers included roughly 1 km (0.00833 decimal degrees) resolution bioclimatic layers from CHELSA 1.2 (Karger et al. 2017) and additional climatic and topographic layers from ENVIREM (Title and Bemmels 2018). The second set included several soil and landcover layers, including percent impervious surface, available at sub-kilometer resolution to facilitate a finer-scale analysis within Orange County. These layers were downloaded from the 2011 National Landcover Database (NLCD) (Homer et al. 2012) and from SoilGrids (Hengl et al. 2017). We downscaled the very high-resolution NLCD layers to 250 m/0.0020833 decimal degrees to match the maximum resolution available for the SoilGrids layers. To enable comparison among rasters of both resolutions, we favored using marginal R^2 to rank surfaces as it provides a measure of model fit comparable across different datasets (Nakagawa and Schielzeth 2013), and also because we were more interested in the surface that explained the most variance (indicated by R^2) rather than model fit per se (although we also report AIC and log-likelihood). We used Circuitscape 4.0.5 (Shah and McRae 2008) with the optimized resistance surface with the highest marginal R^2 to visualize potential high-current dispersal corridors on the landscape. For Circuitscape, we used default parameters for the pairwise mode, with sites used in ResistanceGA as focal nodes and the output raster as the resistance surface. Environmental layers are described in Supplementary Table S5. The subset distance matrix (Nei's Da) and the population coordinates used as inputs for ResistanceGA are provided in Supplementary Tables S6 and S7, respectively.

Results

RAD-Seq Data Assembly

The output of the OC_all_208 VCF file contained 6660 SNPs with up to 75% missing data per site. Per individual, the mean percent missingness was 34.28% (median 35.30%, range 16.30% to 50.86%). Mean sequencing depth per individual was 23.27 (median: 21.11; range: 9.182 to 59.93). The reduced dataset, OC_reduced_148, contained 6864 SNPs with up to 75% missing data per site, had mean missingness per individual of 35.24% (median of 35.50%, range of 16.11% to 50.68%), and mean sequencing depth per individual of 21.06 (median of 19.41, range of 9.659 to 55.63). The slight increase in the number of SNPs in OC_reduced_148 compared to OC_all_208 likely reflects our strategy of subsampling individuals from artificial sites with the lowest individual missingness.

Genetic Structure Across Orange County

Using OC_reduced_148, all 10 replicates of FastStructure returned a maximum marginal likelihood at $K = 2$, splitting coastal Orange County ("Coast") from more interior ("Inland") sites (Figure 2). FastStructure also reported the number of model components used to explain the structure in the data as 2 (7 of 10 replicates), or 3 (3 of 10). There was little inferred admixture at $K = 2$ at most sites except for Sand Canyon Reservoir (SANCAN), where many individuals had a roughly even mix of Inland and Coast genetic cluster ancestry. Tenaja Pond (TENA), a breeding pond directly adjacent to the Sand Canyon Reservoir (Figure 1), grouped with Coast in these initial FastStructure analyses. Both SANCAN and TENA sites are located adjacent to the Los Angeles Basin on the northern edge of the San Joaquin Hills, roughly 10 km from the other Coast sites (Figure 1). At $K = 3$ (non-hierarchical), TENA was split from the other Coast sites in all replicates. Running FastStructure separately on the Coast+SANCAN+TENA and Inland clusters, $K = 1$ had the highest marginal likelihood for all 10 replicates for the Inland cluster. $K = 2$ had the highest for

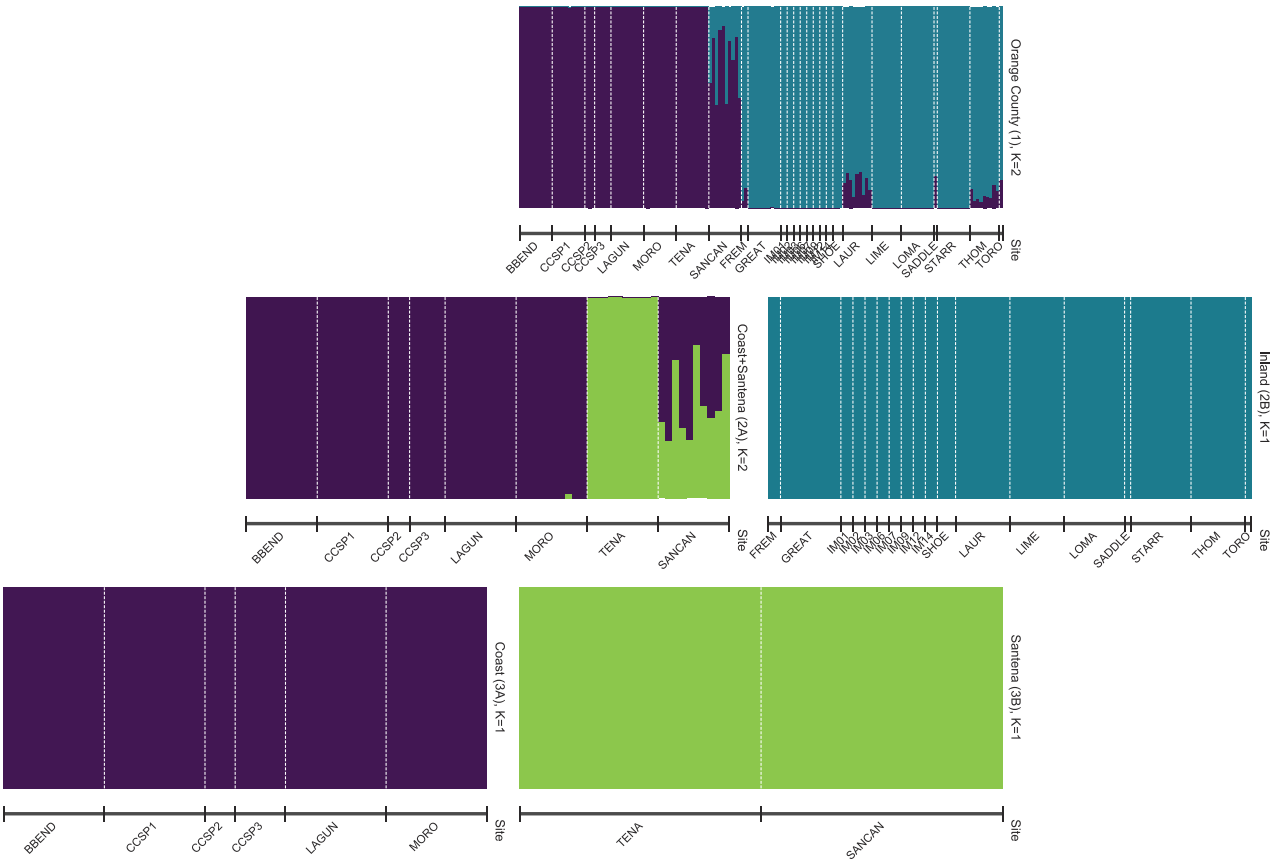

Figure 2. Hierarchical FastStructure barplots using 6864 unlinked SNPs. The upper plot shows results for $K = 2$ populations for 148 Orange County samples. The middle plot shows the optimal results for the Inland cluster (blue, $K = 1$) and the Coast (purple) + Santena (green) samples. The bottom row shows the single cluster that results when each of the Coast and Santena clusters are analyzed separately.

all 10 replicates of the Coast+SANCAN+TENA cluster, splitting off TENA, with SANCAN admixed between TENA and Coast. Given their close proximity (satellite imagery showed the ponds connected during high rain years), we considered SANCAN and TENA together as their own subcluster ("Santena") with no further subdivision. In the remaining Coast cluster, all 10 replicates reported $K = 1$ as the highest marginal likelihood. In summary, both hierarchical and non-hierarchical FastStructure identified 3 major clusters: Inland, Santena, and Coast (Figure 2), although the non-hierarchical analysis returned a maximum marginal likelihood for $K = 2$ (combining Santena + Coast). The artificial sites IM 01-14 (8 total ponds) and SHOE all clustered with Inland, consistent with their geographical proximity and the location of their donor pond (Figure 1) (unfortunately, no samples exist from the donor pond). Although FastStructure indicated that higher values of K explain the structure in the data in the Coast and Inland subsets (between 3 and 5 for all replicates), marginal likelihoods dropped steeply, and the modes of these replicates varied substantially at the higher K values, with sites moving in and out of different clusters and often fewer than K clusters appearing in a given replicate (Supplementary Figure S1).

PCA (Figure 3) showed a similar pattern, but also revealed additional population structure. PC1 (6.42% of variance explained) separated Coast from Inland. Santena occupied an intermediate position between the 2, although TENA and parts of SANCAN were closer on PC1 to the Inland sites. Within Santena, the admixed SANCAN spanned the Coast and Inland clusters. PC2 (4.36% of variance) separated STARR from the remaining Inland sites, PC3 (3.65% of variance) separated TENA, and PC4 (3.36%) further characterized genetic variation within the Inland cluster, distinguishing LOMA and LIME from each other and other sites. PCs 5–8 captured additional genetic variation within the Inland and Coast clusters. The artificial ponds at IM & SHOE clustered with the northern Inland FREM site (PC5, 2.88%), corresponding to the geographic proximity of the IMSHOE donor ponds (Figure 1). PC6 (2.51%) separated the Coastal BBEND and CCSP1, and PC7/8 separated most of the remaining Inland ponds.

FastStructure (run non-hierarchically, Supplementary Figure S1) and fineRADstructure (Supplementary Figure S2), like PCA, indicated substructure beyond the optimal FastStructure results using marginal likelihoods. While FineRADstructure sometimes oversplit individuals within the same pond into their own clusters, its hierarchical splitting did show agreement with FastStructure and PCA in that Inland and Coast are separate clusters. Similar to the later PC's, FineRADstructure identified most sites within Inland as unique, cohesive clusters. IMSHOE + FREM (the artificial sites plus the most northerly of the natural sites) and Coast each formed independent clusters, but ponds within both were often not readily distinguishable from one another (Supplementary Figure S2). FineRADstructure also showed SANCAN having high levels of shared coancestry with both Inland and Coast sites, again corroborating FastStructure and PCA results.

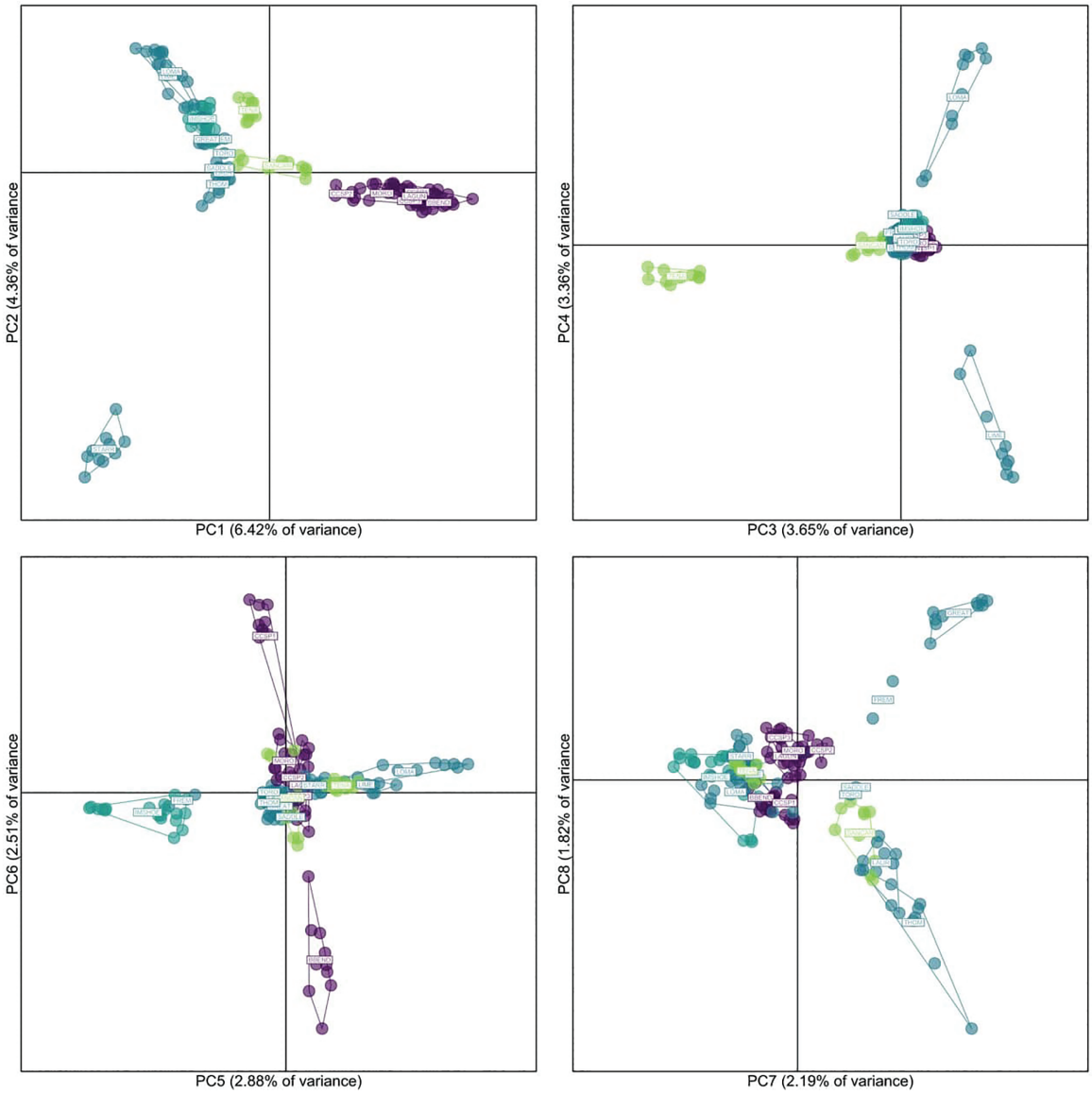

Figure 3. Principal components analysis of 148 natural site Orange County samples with 6864 unlinked SNPs. Colors are as in Figure 2: Purple points are individuals of the Coast FastStructure cluster; blue, Inland; green, Santena.

Genetic Distance

Considering the 3 FastStructure clusters (Inland, Coastal, Santena) as genetic units with the OC_reduced_148 dataset, hierfstat returned a global (among those 3 groups) F_{ST} = 0.111, F_{IT} = 0.278, and F_{IS} = 0.188. The OC_all_208 dataset, when individual pond-breeding sites (including all artificial ponds) were treated as separate populations, had a much higher F_{ST} and lower F_{IS}: global F_{ST} was 0.247, F_{IT} = 0.256, and F_{IS} = 0.012, consistent with the subtle, but consistent differentiation among ponds in PC 2–8 and in the fineRADstructure results. In pairwise comparisons (Supplementary Figure S3), we found lower F_{ST} values among the 3 clusters (0.079–0.142, mean: 0.112; median: 0.116) than among individual breeding sites within the Coast (0.049–0.244, mean: 0.142; median: 0.163) or Inland (0.119–0.370, mean: 0.228; median: 0.218) clusters when calculated for all sites with more than a single sample. Pairwise F_{ST} values among the artificial ponds in IMSHOE range from 0.005 between IM06 and IM07, to 0.297 between IM06 and IM12.

NeighborNet networks of pairwise F_{ST}, Jost's D, and Nei's D_a all revealed similar patterns of genetic distance (Figure 4, Supplementary Figures S5 and S6), showing clear separation between Inland and Coast with SANCAN and TENA (Santena) falling in between, but closer to the Inland group. Reticulations in the networks were common among sites within genetic clusters. Notably, the pairwise Jost's D and Nei's D_a networks indicated that the artificial site SHOE appeared to share a high number of edges with natural Inland and the artificial IM populations (Supplementary Figures S4 and S5), and SHOE is also the artificial site in the closest geographic proximity to known natural sites.

Genetic Diversity and Effective Population Size

Among the 3 FastStructure genetic clusters, Inland (excluding artificial ponds) had the highest genetic diversity across measures, but also the highest F_{IS} value (Table 1). Coast had the lowest measures of genetic diversity—even lower than the artificial ponds (IMSHOE)

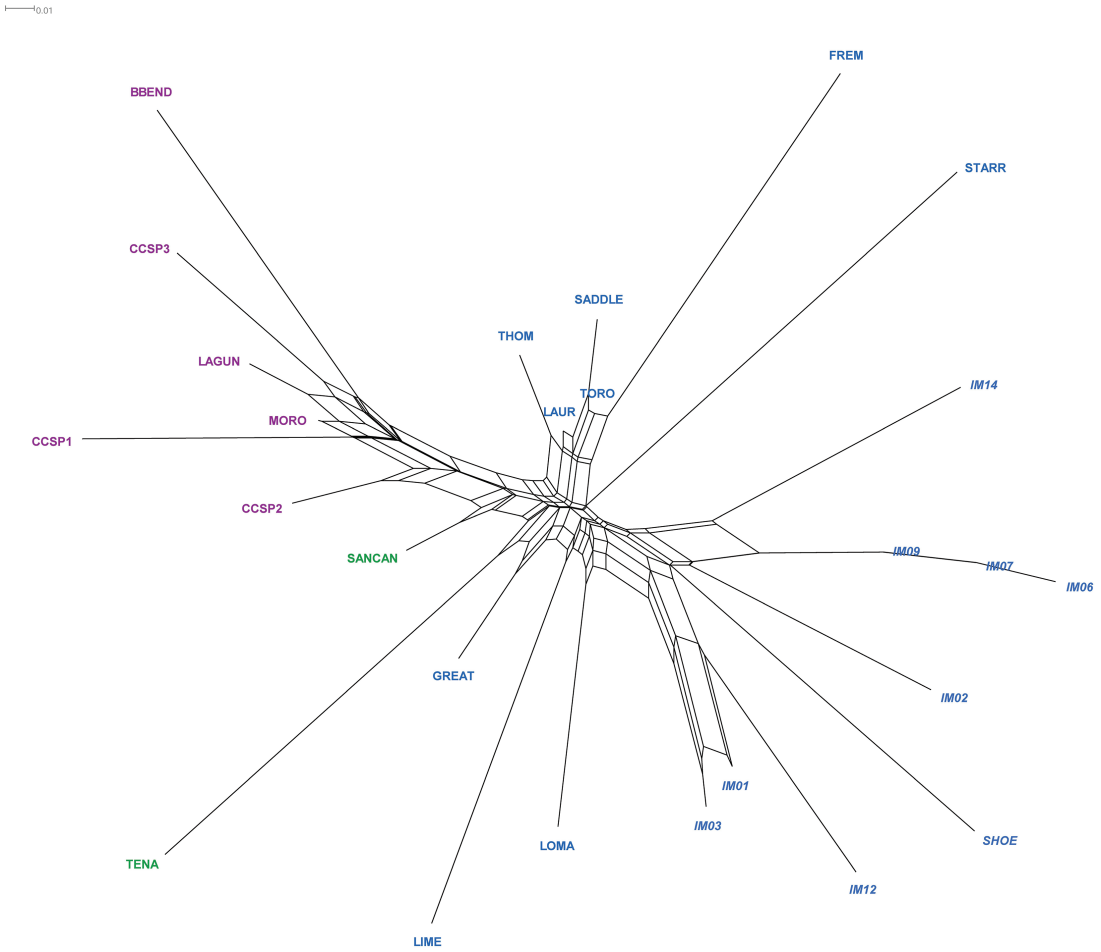

Figure 4. NeighborNet distance network of Weir and Cockerham's pairwise FST among 26 sites for OC_all_208 with up to 75% missing data per locus. Colors are as in Figure 2: Purple are sites that fall into the Coast FastStructure cluster; blue, Inland; green, Santena.

when considered as a group—but had a slightly lower F_{IS} than IMSHOE (0.0213 vs. 0.0257). At the site/pond level, natural sites within the Inland cluster tended to have higher values of genetic diversity (*pi* = 5.40E-4) than Coast (4.06E-4), Santena (4.36E-4), and artificial sites (4.20E-4). Eight out of 17 natural sites with 2 or more available samples had slightly negative F_{IS} values, indicating that mild outbreeding (or at least no discernable inbreeding) may be taking place.

Effective population size and effective number of breeders at sites were also generally low. Estimates of the effective number of breeders (N_{eb}) at breeding ponds based on the molecular coancestry method ranged from 1.2 to 12.2, while effective population size estimates based on the linkage disequilibrium method (LDN_e) ranged from 1.4 to 19.8, with 3 outliers: 62 (SHOE, an artificial pond), 64 (THOM), and 167 (LOMA). Given that these populations did not have high numbers of effective breeders (1.4, 1.2, 3.5), we view these outliers as likely statistical artifacts rather than historically large populations.

Considering all sites with 3 or more samples (except SANCAN and TENA; see methods), we found statistically significant correlations between the log pond area and both HO ($r = 0.839$, $P < 0.001$) and HE ($r = 0.828$, $P < 0.001$) and the logarithm of pond area, but no significant correlation with N_{eb} or FIS (Supplementary Figure S6). We also found significant correlations between HO ($r = 0.871$, $P < 0.001$) and HE ($r = 0.825$, $P = 0.002$) with log pond area when excluding artificial ponds. When broken down by Coast, Inland artificial, and Inland natural, we found a significant correlation only for the natural Inland ponds between HE and log pond area ($r = 0.836$, $P = 0.034$) (Supplementary Figure S6).

Landscape Resistance

ResistanceGA results provided clear evidence of isolation by environment and distance among natural ponds. Using marginal R^2 (R^2_m) as the rank measure, ResistanceGA identified a transformation of depth to bedrock combined with slope (Figure 5a) as the highest ranked surface among the environmental surfaces, explaining 77.0% ($R^2_m = 0.770$) of the variance in pairwise genetic distance. Depth to bedrock alone ($R^2_m = 0.761$) and landcover alone ($R^2_m = 0.742$) ranked second and third, respectively (Supplementary Table S8), in their ability to explain observed genetic distance. For environmental surfaces available at both high-resolution and low-resolution, the high-resolution version always outperformed the low, although the signal from depth to bedrock was strong enough that the low-resolution version ranked just behind the high-resolution version with a R^2_m of 0.741. Geographic distance alone ranked near the middle with a marginal R^2 of 0.562 (0.493 at low resolution), suggesting that isolation by distance is an important component of *Spea* landscape resistance in Orange County, but not the most important.

Table 1. Genetic diversity, genetic cluster membership, and site information of sampled ponds

Pond	Cluster_K = 3	Pond_natural	n	Sample_years	HO	HE	FIS	Neb	LDN_e	Pond area (m^2)	Longitude	Latitude
BBEND	Coast	Natural	10	2011, 2017	0.07391	0.06515	−0.00286	3.2	11	11	−117.769774	33.568042
CCSP1	Coast	Natural	10	2017	0.08244	0.06979	−0.01037	2.6	2.2	21.8	−117.807412	33.563498
CCSP2	Coast	Natural	3	2017	0.08748	0.06318	−0.01516	NA	NA	NA	−117.804929	33.564458
CCSP3	Coast	Natural	5	2017	0.07735	0.05922	−0.0184	12.2	NA	32.065	−117.800118	33.56952
FREM	Inland	Natural	2	2004, 2011	0.07864	0.0555	0.00783	NA	NA	NA	−117.69521	33.83007
GREAT	Inland	Natural	10	2015	0.1098	0.09804	−0.01399	3.7	2	2326.98	−117.691022	33.678224
IM01	Inland	Artificial	10	2016	0.07826	0.07717	0.0073	2	8.2	39.295	−117.65726	33.74162
IM02	Inland	Artificial	8	2016	0.07967	0.06579	−0.01564	3.1	1.4	26.67	−117.65852	33.74172
IM03	Inland	Artificial	7	2016	0.06927	0.05873	−0.011	1.7	19.8	35.12	−117.65758	33.74151
IM06	Inland	Artificial	10	2016	0.07798	0.06383	−0.02134	2.2	6.1	24.525	−117.65688	33.7376
IM07	Inland	Artificial	10	2016	0.07617	0.0656	−0.01256	1.7	2.9	24.96	−117.6567	33.73763
IM09	Inland	Artificial	10	2016, 2017	0.07297	0.07521	0.02117	1.9	3	34.775	−117.65698	33.73775
IM12	Inland	Artificial	6	2016	0.06051	0.05421	0.00152	4.9	3	22.38	−117.65714	33.73748
IM14	Inland	Artificial	8	2016	0.07426	0.06087	−0.01898	2.1	8.7	65.82	−117.65372	33.73621
LAGUN	Coast	Natural	10	2011, 2017	0.08403	0.0897	0.03296	3	2.1	22.6	−117.78134	33.558171
LAUR	Inland	Natural	9	2017	0.11	0.11396	0.03452	1.4	9.1	2249.185	−117.622406	33.553285
LIME	Inland	Natural	9	2017	0.09991	0.087	−0.00832	2.1	11.4	168.96	−117.659055	33.715315
LOMA	Inland	Natural	10	2017	0.09861	0.09243	0.00565	3.5	NA	768.305	−117.709481	33.745564
MORO	Coast	Natural	10	2011, 2017	0.08974	0.09118	0.02417	1.8	2	44.135	−117.794304	33.576053
SADDLE	Inland	Natural	1	2003	0.10761	0.0538	NA	NA	NA	NA	−117.58774	33.65029
SANCAN	Santena	Natural	10	1995	0.09819	0.10168	0.02698	4.2	2.6	NA	−117.796878	33.639673
SHOE	Inland	Artificial	10	2017	0.09114	0.07622	−0.01889	1.4	NA	174.385	−117.719416	33.760152
STARR	Inland	Natural	10	2017	0.10897	0.08915	−0.02632	4.4	2.2	138.965	−117.559973	33.603068
TENA	Santena	Natural	10	1995	0.07938	0.06709	−0.01351	9	NA	1032.655	−117.794122	33.647421
THOM	Inland	Natural	9	2017	0.10429	0.10682	0.02759	1.2	63.8	2542.82	−117.600762	33.575703
TORO	Inland	Natural	1	1991	0.12134	0.06067	NA	NA	NA	NA	−117.759902	33.583093

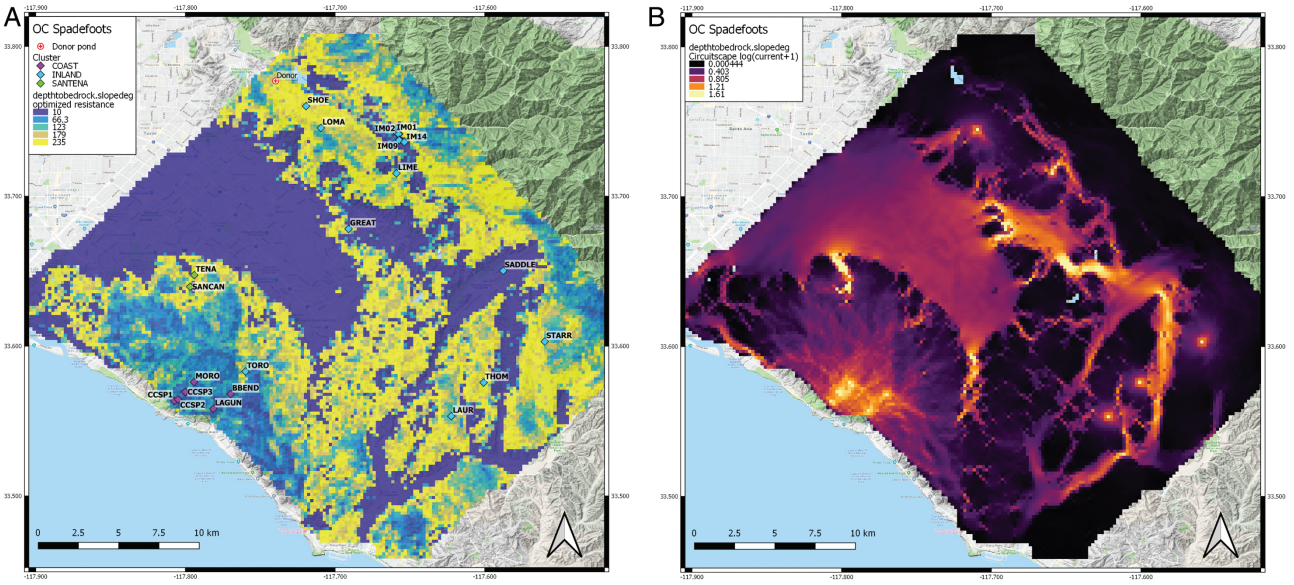

Figure 5. (**A**) ResistanceGA-optimized resistance parameterization of depth to bedrock + slope raster. Purple/Blue is lower resistance, green/yellow is higher. Diamond points show sampled localities, with color indicating genetic cluster membership as in Figure 1. (**B**) Circuitscape current run through optimized (depth to bedrock + slope) raster, using natural localities as focal nodes. Current is displayed as the logarithmic transform of the value. Highest current is in yellow/orange, lowest current in purple/black. High-current areas are inferred as corridors with the highest level of potential movement.

The optimized depth to bedrock + slope surface tended to show the lowest resistance in the highly human-modified lowlands of Orange County (Los Angeles Basin), where spadefoots are now largely extirpated (coinciding with the lowest untransformed values of slope and highest untransformed values of depth to bedrock in the map area). Running current through the optimized depth to bedrock + slope resistance surface in Circuitscape confirmed low-lying valleys as generally the lowest resistance potential corridors between sample sites (Figure 5b).

Discussion

Although widely viewed as a critical element in the conservation biology toolkit, there is often a mismatch between the application of landscape genetic studies and the inferences that they inform. The reason for this lies in the scale of most studies—evolutionary processes that are most easily recoverable with genetic approaches frequently occur at large, regional spatial scales spanning hundreds-to-thousands of kilometers and evolutionary time, while conservation management and planning is most relevant to finer spatial scales across tens of kilometers and decadal time scales. In this study, we explicitly studied local patterns in the landscape genomics of the western spadefoot, a declining amphibian of high conservation value. Even at this very limited spatial scale, we identified discrete population clusters that should function as management units and more subtle population structure consistent with strong philopatry to individual breeding sites. Combined with small effective population sizes and severe habitat loss across the region, our work suggests that *S. hammondii* in its Orange County stronghold is both at risk and a strong candidate for active management. Fortunately, artificial vernal pool breeding sites appear to support relatively healthy spadefoot populations (Baumberger et al. 2020), and we strongly recommend additional pond construction and upland restoration as important strategies for future conservation actions. Whether or not assisted migration among these management units that mimics historical, but now impossible, levels of gene flow is beneficial is an important concern for future management.

Population Structure

We identified 2 major genetic clusters in Orange County spadefoots (Figure 2, $K = 2$), and these should comprise the primary management units for the species in the region. There was also strong evidence for a third group, "Santena," that has affinities to both the Coastal (Figure 2) and Inland (Figures 3 and 4) populations; further resolution of the distinctiveness and affinities of these populations is an important future goal for the species. However, a clear distinction exists between spadefoots in the coastal San Joaquin Hills (Coast cluster) and those in the Los Angeles Basin and foothills of the Santa Ana Mountains (Inland cluster). A larger scale analysis identified the Coast cluster as 1 of 3 primary groupings across the entire southern *S. hammondii* clade (Neal 2019), confirming its importance as a primary management unit. Other species show similar patterns of genetic isolation of a coastal Orange County group (Barr et al. 2015; Fraser et al. 2019; Kozakiewicz et al. 2019; Vandergast et al. 2019) further suggesting that this currently isolated island of coastal upland habitat is an important conservation target for the region.

In addition to this fundamental population subdivision, other analyses, including FastStructure with higher values of *K*, PCA, fineRADstructure, and global and pairwise *F*-statistics *all provided evidence of additional differentiation, often among adjacent ponds. This* pond-level structure is consistent with our understanding of spadefoot biology. Spadefoots are strongly philopatric, and telemetry studies indicate that they rarely move more than a few hundred meters from their breeding pond (Baumberger et al. 2019). In contrast, a study of a congeneric spadefoot species, *Spea multiplicata*, found it to have the lowest levels of genetic structure compared to 2 other pond-breeding amphibians in a floodplain in Arizona, suggesting that spadefoots (or at least *S. multiplicata*) may actually be fairly vagile on flat, open terrain (Mims et al. 2014). The pattern seen in both genetics (this study) and telemetry (Baumberger et al. 2019)

suggest that *S. hammondii* movement ecology is fundamentally different, and more movement-limited. The finer-scale genetic patterns and limited movement based on telemetry may reflect the refugial nature of remaining populations in southern California to marginal habitat in hilly terrain, although the deeper genetic split between Coastal and Inland populations reflects long-term evolutionary divergence. However, consistent with our landscape resistance analysis (Figure 5b), pre-urbanization southern California spadefoots may have also experienced well-connected pond networks following rains in the flat Los Angeles Basin that were more similar to the large-scale gene flow seen in lowland Arizona.

Landscape Resistance and Connectivity

ResistanceGA identified the areas with the lowest resistance to gene flow as having the greatest depth to bedrock and lowest slope, characteristic of relatively flat basins and valleys where spadefoots are now virtually extirpated in Orange County. A study of the historical extent of vernal pools in neighboring San Diego County revealed a once-extensive distribution of vernal pools and vernal pool complexes in lowlands with high depth to bedrock and low slope (Bauder and McMillan 1998), and a similar network of vernal pools existed in the Los Angeles Basin (Mattoni et al. 1997; Mattoni and Longcore 1997). A Maxent environmental niche model (with no anthropogenic land-use variables included) for southern-clade *S. hammondii* also showed some of the highest habitat suitability in the Los Angeles Basin (Neal et al. 2018). Human development has impacted this entire region, eliminating the natural landscape and disrupting genetic connectivity for most native species (Vandergast et al. 2008). Most sampled ponds in this study are restricted to the lower-suitability, higher-resistance uplands bordering the Los Angeles Basin, suggesting that the remaining breeding ponds in Orange County may represent refugial populations in ecologically marginal, but still intact habitat conserved to ensure biodiversity and ecological integrity (County of Orange 1996). However, given the greater landscape resistance of the remaining preserved habitat, installing and maintaining additional breeding ponds between existing ponds should be an important management strategy to ensure that connectivity and metapopulation dynamics persist within these relatively high-resistance landscapes.

Genetic Diversity and Effective Population Size

We found very low effective population sizes and numbers of effective breeders in sampled ponds—among the lowest recorded for pond-breeding amphibians, including the declining California black toad (*Bufo exsul*) (Wang 2009), and the federal and state-listed California red-legged frog (*Rana draytonii*) (Richmond et al. 2013; Richmond et al. 2014) and California tiger salamander (*Ambystoma californiense*) (Wang et al. 2011). Based on results from the California tiger salamander (Wang et al. 2011; Toffelmier and Shaffer, unpublished data) and eastern tiger salamander (McCartney-Melstad et al. 2018b), we predicted that measures of genetic diversity and effective population size (LDN_e, N_{eb}) would be positively correlated with pond area. While we found significant positive correlations between genetic diversity and pond area for all sites considered together and for natural Inland sites, no such correlations exist with effective population size or effective number of breeders for any genetic cluster (Supplementary Figure S6). Several possible reasons may explain this puzzling result. First, even with thousands of SNPs, N_e estimates are often not very precise, especially for LDN_e. The jackknife 95% confidence intervals on LDN_e were often very large and had an infinite upper bound for most ponds (Supplementary Table S9). The jackknife 95% confidence intervals for molecular coancestry-estimated effective number of breeders (N_{eb}) were more reasonable, roughly 10% above and below the estimated values. This imprecision does not seem to be driven by high relatedness among tadpoles in pond samples, and removing putative siblings based on KING relatedness (Manichaikul et al. 2010) did not change the infinite upper bounds on jackknife confidence intervals (results not shown). (Given that removing close relatives risks erasing the true evolutionary signal of small populations (Waples and Anderson 2017), we retained such relatives in these and other analyses.) Second, N_e may be better explained by other site characteristics that are not associated with pond area such as age, hydroperiod, catchment/upland area, predatory macroinvertebrates, or a number of other hydrological or biological variables (Semlitsch et al. 2015). Third, California was recently in a multi-year drought from roughly 2011–2017 that likely reduced survival and fecundity for several consecutive breeding seasons (Fisher et al. 2018). In typical climate years, larger ponds should hold water longer and extend the period before metamorphosis is necessary, leading to enhanced metamorph fitness (Morey 1998; Morey and Reznick 2004) and, over time, higher effective population sizes. However, if droughts are severe enough that even large ponds rarely held water for the 21-day minimum necessary for spadefoots to reach metamorphosis, or lead to a more haphazard recruitment pattern that depends on local conditions of shading, slope, and exposure, then ponds could have experienced bottlenecks leading to generally small effective population sizes independent of pond area. To narrow down the causes of the observed low effective population sizes, we strongly recommend follow-up sampling to study temporal changes in effective population size with respect to climate and landscape change.

Artificial Pond Contribution to Spadefoot Conservation

The existing constructed artificial breeding ponds at IM and Shoestring Canyon (SHOE) were stocked with tadpoles from a single cluster of ponds spanning less than 1 km that we assume were essentially a single genetic source pond (Glenn Lukos Associates, Inc. 2005, 2006; Baumberger et al. 2020). This source site shared similar characteristics with the natural ponds sampled in this study in that it was in hilly terrain at the edge of the Los Angeles Basin and isolated from other known breeding ponds by areas of high resistance. The placement of the artificial ponds was in presumed low-resistance habitat that has the greatest potential for migration and admixture with nearby natural ponds. In the decade since first tadpole introduction (2005–2006) and the current study's sampling (2015–2017), SHOE already shows small amounts of admixture with natural Inland populations, confirming both that spadefoots breed and persist in these artificial ponds, and that there is sufficient movement among ponds to establish low-level gene flow and, potentially, reestablish metapopulation dynamics. With only natural sites as focal nodes, Circuitscape (Figure 5) reconstructed a low-resistance corridor between natural and these artificial sites, validating the choice of location for the construction of these ponds.

Unfortunately, stocking all of the constructed artificial ponds with individuals from essentially a single population limits their potential to contribute to higher levels of genetic variation within the regional metapopulation. We observed hundreds of tadpoles

in the artificial ponds during sampling from 2015 to 2017 (personal observations, and Baumberger et al. 2020), but found low levels of genetic diversity and small effective population sizes, consistent with their initial founding from a single gene pool. When combined and considered as a single population, the group of 9 artificial ponds still had lower genetic diversity than most natural Inland and some Coast ponds. To take full advantage of the existing artificial ponds, or when considering the construction of new ponds, we recommend stocking with tadpoles more broadly from Inland or Coast sites (but not both) to establish genetically diverse populations that capture aspects of the gene flow that presumably existed before major human impacts in the Los Angeles Basin. Whether Coast and Inland populations should be admixed to further increase local genetic diversity is an open question: the observation that the Coast population cluster is a significant genetic unit even at a range-wide level for *S. hammondii* (Neal 2019) argues against such admixture, but natural admixture of Coast and Inland genotypes at SANCAN suggests that some additional admixture might replicate historical levels of genetic mixing. A conservative step forward for the IM/SHOE ponds would be to first introduce tadpoles from additional, closely related Inland ponds and monitor the populations for outbreeding depression or other negative impacts before considering bringing in more distantly related cohorts to further boost genetic variation.

Regardless of the details of post-construction introductions, our data support building additional artificial breeding ponds to enhance natural metapopulation dynamics by establishing a network of "stepping stones" that traverse low-resistance corridors and allow some gene flow between natural and constructed ponds. This innovative, proactive conservation action has already been initiated by the Natural Communities Coalition on one property occupied by the Coastal population, and initial results, including observations of successful breeding and metamorphosis in their first year of operation, are extremely promising.

Conclusions and Next Steps

The landscape resistance models and low effective population sizes found in this study suggest that *S. hammondii* in Orange County persists in ecologically intact, but marginal upland habitat from a landscape connectivity perspective. Spadefoot populations are small, isolated, and likely at demographic and genetic risk, particularly during drought conditions that are predicted to increase under future climate conditions. However, there have also been important efforts to protect large tracts of habitat, and as a consequence, Orange County remains a stronghold for the species in southern California. We strongly recommend a combined approach of 1) assisted migration to counteract the negative effects of inbreeding depression (Frankham et al. 2019) among both natural and artificial ponds within Inland and Coast regions, and 2) construction of additional, artificial breeding sites with sufficiently long predicted hydroperiods (Pyke 2004) to enhance the ability of spadefoots to naturally establish local metapopulations as effective strategies to maintain this declining vernal pool specialist on the southern California landscape.

Supplementary Material

Supplementary material can be found at *Journal of Heredity* online.

Funding

This work was supported by grants from the Natural Communities Coalition (Agreement 16-12), the National Science Foundation (DEB 1257648), the U.S. Fish and Wildlife Service (F16PX02290), and the University of California Conservation Genomics Consortium to HBS, the UCLA Department of Ecology and Evolutionary Biology, the La Kretz Center for California Conservation Science, and the Sea and Sage Audubon to KMN, and the Ecosystems Mission Area of USGS to RNF.

Acknowledgments

We thank Gary Bucciarelli, Robert Cooper, Tara Luckau, Evan McCartney-Melstad, Peter Scott, Erin Toffelmier, and other past and present Shaffer Lab members for help with various discussions, feedback, and advice. We also thank Kathy Baumberger for assistance with sampling. We are grateful to Tom Gillespie, Kirk Lohmueller, and Victoria Sork for their revisions and feedback as part of Neal's doctoral dissertation committee, and the California Department of Fish and Wildlife, Orange County Parks, and Crystal Cove State Park for permitting and site access. This manuscript is contribution 770 of the USGS Amphibian Research and Monitoring Initiative. Any use of trade names or specific product is for descriptive purposes only and does not imply endorsement of the U.S. Government.

Data Availability

We have deposited the primary data (sampling locations, genotype data, environmental raster layers) underlying these analyses and Supplementary Tables on Figshare at https://doi.org/10.6084/m9.figshare.12721019.v1.

References

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat Rev Genet*. 17:81–92.

Barr KR, Kus BE, Preston KL, Howell S, Perkins E, Vandergast AG. 2015. Habitat fragmentation in coastal southern California disrupts genetic connectivity in the cactus wren (*Campylorhynchus brunneicapillus*). *Mol Ecol*. 24:2349–2363.

Bauder ET, McMillan S. 1998. Current distribution and historical extent of vernal pools in southern California and northern Baja California, Mexico. Ecology, Conservation and Management of Vernal Pool Ecosystems–Proceedings from a 1996 Conference. Presented at the California Native Plant Society, Sacramento, CA.

Baumberger KL, Backlin AR, Gallegos EA, Hitchcock CJ, Fisher RN. 2020. Mitigation ponds offer drought resiliency for *Spea hammondii* (western spadefoot) populations. *Southern California Academy of Sciences*. 119:6–17.

Baumberger KL, Eitzel MV, Kirby ME, Horn MH. 2019. Movement and habitat selection of the western spadefoot (*Spea hammondii*) in southern California. *PLoS One*. 14:e0222532.

Bayona-Vásquez NJ, Glenn TC, Kieran TJ, Pierson TW, Hoffberg SL, Scott PA, Bentley KE, Finger JW, Watson PR, Louha S, *et al.* 2019. Adapterama III: quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *BioRxiv*. 205799. doi: 10.1101/205799

Beebee TJC, Griffiths RA. 2005. The amphibian decline crisis: a watershed for conservation biology? *Biol Conserv*. 125:271–285.

Blaustein A, Wake D, Sousa W. 1994. Amphibian declines: judging stability, persistence, and susceptibility of populations to local and global extinctions. *Conserv Biol*. 8:60–71. doi: 10.1046/j.1523-1739.1994.08010060.x

Bryant D, Moulton V. 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol*. 21:255–265.

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

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34:i884–i890.

Corey SJ, Waite TA. 2008. Phylogenetic autocorrelation of extinction threat in globally imperiled amphibians. *Divers Distrib*. 14:614–629.

County of Orange. 1996. Natural community conservation plan and habitat conservation plan. County of Orange Central and Coastal Subregion. Available from: https://www.wildlife.ca.gov/Conservation/Planning/NCCP/Plans/Orange-Coastal

Cushman SA. 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 CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.*; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.

Davidson C, Shaffer HB, Jennings MR. 2002. Spatial tests of the pesticide drift, habitat destruction, UV-B and climate-change hypotheses for California amphibian declines. *Conserv Biol*. 16:1588–1601.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 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. 2019. Interactive assembly and analysis of RAD-seq data sets: dereneaton/ipyrad [Jupyter Notebook]. Available from: https://github.com/dereneaton/ipyrad (Original work published 2015).

Eaton DA. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*. 30:1844–1849.

Fisher RN, Brehme CS, Hathaway SA, Hovey TE, Warburton ML, Stokes DC. 2018. Longevity and population age structure of the arroyo southwestern toad (*Anaxyrus californicus*) with drought implications. *Ecol Evol*. 8:6124–6132.

Frankham R, Ballou JD, Ralls K, Eldridge M, Dudash MR, Fenster CB, Lacy RC, Sunnucks P. 2019. *A practical guide for genetic management of fragmented animal and plant populations*. Oxford (UK): Oxford University Press.

Fraser DL, Ironside K, Wayne RK, Boydston E. 2019. Connectivity of mule deer (*Odocoileus hemionus*) populations in a highly fragmented urban landscape. *Landsc Ecol*. 34:1097–1115.

Garner BA, Hand BK, Amish SJ, Bernatchez L, Foster JT, Miller KM, Morin PA, Narum SR, O'Brien SJ, Roffler G, *et al.* 2016. Genomics in conservation: case studies and bridging the gap between data and application. *Trends Ecol Evol*. 31:81–83.

Glenn Lukos Associated, Inc. 2005. East Orange Planned Community Orange County, California. Western spadefoot toad (WST) (*Spea hammondii*) WST relocation and habitat creation program. Report prepared for The Irvine Company, by Glenn Lukos Associated, Lake Forest (CA), p. 8.

Glenn Lukos Associated, Inc. 2006. East Orange Planned Community (EOPC) translocation and monitoring report for the western spadefoot toad (*Spea hammondii*). Final Report prepared for The Irvine Company, by Glenn Lukos Associated, Lake Forest (CA), pp. 11.

GBIF.org. 2017. *Spea hammondii (Baird, 1859) GBIF occurrence download*. doi:10.15468/dl.uggrh3

Goldberg CS, Waits LP. 2010. Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Mol Ecol*. 19:3650–3663.

Gosselin T. 2019. RADseq data exploration, manipulation and visualization using R: thierrygosselin/radiator [R]. Available from: https://github.com/thierrygosselin/radiator (Original work published 2017)

Goudet J, Jombart T. 2015. hierfstat: estimation and tests of hierarchical F-statistics (Version 0.04–22). Available from: https://CRAN.R-project.org/package=hierfstat

Grant EHC, Miller DAW, Schmidt BR, Adams MJ, Amburgey SM, Chambert T, Cruickshank SS, Fisher RN, Green DM, Hossack BR, *et al.* 2016. Quantitative evidence for the effects of multiple drivers on continental-scale amphibian declines. *Sci Rep*. 6:25625.

Hamer AJ, McDonnell MJ. 2008. Amphibian ecology and conservation in the urbanising world: a review. *Biol Conserv*. 141:2432–2449.

Hendricks S, Anderson EC, Antao T, Bernatchez L, Forester BR, Garner B, Hand BK, Hohenlohe PA, Kardos M, Koop B, *et al.* 2018. Recent advances in conservation and population genomics data analysis. *Evol Appl*. 11:1197–1211.

Hengl T, Mendes de Jesus J, Heuvelink GB, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, *et al.* 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One*. 12:e0169748.

Homer CH, Fry JA, Barnes CA. 2012. USGS fact sheet 2012–3020: the National Land Cover Database. Available from: https://pubs.usgs.gov/fs/2012/3020/

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*. 23:254–267.

Jennings MR, Hayes MP. 1994. *Amphibian and reptile species of special concern in California*. Sacramento (CA): California Department of Fish and Game.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24:1403–1405.

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

Jombart T, Kamvar ZN, Collins C, Lustrik R, Beugin MP, Knaus BJ, Solymos P, Mikryukov V, Schliep K, Maié T, *et al.* 2018. adegenet: exploratory analysis of genetic and genomic data (Version 2.1.1). Available from: https://cran.r-project.org/web/packages/adegenet/index.html

Jost L. 2008. G(ST) and its relatives do not measure differentiation. *Mol Ecol*. 17:4015–4026.

Jost L, Archer F, Flanagan S, Gaggiotti O, Hoban S, Latch E. 2018. Differentiation measures for conservation genetics. *Evol Appl*. 11:1139–1148.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Sci Data*. 4:170122.

Kozakiewicz CP, Burridge CP, Funk WC, Salerno PE, Trumbo DR, Gagne RB, Boydston EE, Fisher RN, Lyren LM, Jennings MK, *et al.* 2019. Urbanization reduces genetic connectivity in bobcats (*Lynx rufus*) at both intra- and interpopulation spatial scales. *Mol Ecol*. 28:5068–5085.

Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of population structure using dense haplotype data. *PLoS Genet*. 8:e1002453.

Li YL, Liu JX. 2018. StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour*. 18:176–177.

Linck E, Battey CJ. 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol Ecol Resour*. 19:639–647.

Lischer HE, Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*. 28:298–299.

Malinsky M, Trucchi E, Lawson DJ, Falush D. 2018. RADpainter and fineRADstructure: population inference from RADseq Data. *Mol Biol Evol*. 35:1284–1290.

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

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 26:2867–2873.

Marsh DM, Trenham PC. 2001. Metapopulation dynamics and amphibian conservation. *Conserv Biol*. 15:40–49.

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

Mattoni R, Longcore T. 1997. The Los Angeles coastal prairie, a vanished community. *Crossosoma*. 23:71–102.

Mattoni R, Longcore T, George J, Rich C. 1997. Down memory lane: the Los Angeles coastal prairie and its vernal pools. Poster Presentation at 2nd Interface Between Ecology and Land Development in California. Presented at the Occidental College.

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

McCartney-Melstad E, Vu JK, Shaffer HB. 2018b. Genomic data recover previously undetectable fragmentation effects in an endangered amphibian. *Mol Ecol*. 27:4430–4443.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol*. 66:526–538.

McMahon BJ, Teeling EC, Höglund J. 2014. How and why should we implement genomics into conservation? *Evol Appl*. 7:999–1007.

Mims MC, Phillipsen IC, Lytle DA, Kirk EEH, Olden JD. 2014. Ecological strategies predict associations between aquatic and genetic connectivity for dryland amphibians. *Ecology*. 96:1371–1382.

Morey SR. 1998. Pool duration influences age and body mass at metamorphosis in the western spadefoot toad: implications for vernal pool conservation. Ecology, Conservation, and Management of Vernal Pool Ecosystems-Proceedings from a 1996 Conference. California Native Plant Society, Sacramento, CA, 86–91.

Morey SR, Reznick DN. 2004. The relationship between habitat permanence and larval development in California spadefoot toads: field and laboratory comparisons of developmental plasticity. *Oikos*. 104:172–190.

Murphy MA, Dezzani R, Pilliod DS, Storfer A. 2010. Landscape genetics of high mountain frog metapopulations. *Mol Ecol*. 19:3634–3649.

Nakagawa S, Schielzeth H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol*. 4:133–142.

NCCP Plan Summary – County of Orange (Central/Coastal) NCCP/HCP. n.d. Available from: https://www.wildlife.ca.gov/Conservation/Planning/NCCP/Plans/Orange-Coastal

Neal KM. 2019. *An integrative population and landscape genomic approach to conservation of a threatened California Amphibian at multiple spatial scales* [Doctoral dissertation]. Los Angeles (CA): University of California.

Neal KM, Johnson BB, Shaffer HB. 2018. Genetic structure and environmental niche modeling confirm two evolutionary and conservation units within the western spadefoot (*Spea hammondii*). *Conservation Genetics*. 19:937–946.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol*. 19:153–170.

Nomura T. 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evol Appl*. 1:462–474.

Olson DH, Aanensen DM, Ronnenberg KL, Powell CI, Walker SF, Bielby J, Garner TW, Weaver G, Fisher MC; Bd Mapping Group. 2013. Mapping the global emergence of Batrachochytrium dendrobatidis, the amphibian chytrid fungus. *PLoS One*. 8:e56802.

Ortiz EM. 2019. Tools for data conversion and results visualization for fineRADstructure (http://cichlid.gurdon.cam.ac.uk/fineRADstructure.html): edgardomortiz/fineRADstructure-tools [Python]. Available from: https://github.com/edgardomortiz/fineRADstructure-tools (Original work published 2017)

Pembleton LW. 2017. StAMPP: statistical analysis of mixed ploidy populations (Version 1.5.1). Available from: https://CRAN.R-project.org/package=StAMPP

Pembleton LW, Cogan NOI, Forster JW. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol Ecol Resour*. 13:946–952.

Peterman WE. 2018. ResistanceGA: an R package for the optimization of resistance surfaces using genetic algorithms. *Methods Ecol Evol*. 9:1638–1647.

Peterman WE, Connette GM, Semlitsch RD, Eggert LS. 2014. Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol Ecol*. 23:2402–2413.

Peterman WE, Winiarski KJ, Moore CE, Carvalho CS, Gilbert AL, Spear SF. 2019. A comparison of popular approaches to optimize landscape resistance surfaces. *Landsc Ecol*. 34:2197–2208.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*. 7:e37135.

Polich RL, Searcy CA, Shaffer HB. 2013. Effects of tail-clipping on survivorship and growth of larval salamanders. *J. Wildlife Management*. 77:1420–1425.

Puechmaille SJ. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol Ecol Resour*. 16:608–627.

Pyke CR. 2004. Simulating vernal pool hydrologic regimes for two locations in California, USA. *Ecol Model*. 173:109–127.

R Core Team. 2018. R: a language and environment for statistical computing. Retrieved from https://www.R-project.org/

Raj A, Stephens M, Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 197:573–589.

Richmond JQ, Backlin AR, Tatarian PJ, Solvesky BG, Fisher RN. 2014. Population declines lead to replicate patterns of internal range structure at the tips of the distribution of the California red-legged frog (*Rana draytonii*). *Biol Conserv*. 172:128–137.

Richmond JQ, Barr KR, Backlin AR, Vandergast AG, Fisher RN. 2013. Evolutionary dynamics of a rapidly receding southern range boundary in the threatened California Red-Legged Frog (*Rana draytonii*). *Evol Appl*. 6:808–822.

Rohr JR, Raffel TR. 2010. Linking global climate and temperature variability to widespread amphibian declines putatively caused by disease. *Proc Natl Acad Sci U S A*. 107:8269–8274.

Scheele BC, Pasmans F, Skerratt LF, Berger L, Martel A, Beukema W, Acevedo AA, Burrowes PA, Carvalho T, Catenazzi A, *et al*. 2019. Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science*. 363:1459–1463.

Semlitsch RD, Peterman WE, Anderson TL, Drake DL, Ousterhout BH. 2015. Intermediate pond sizes contain the highest density, richness, and diversity of pond-breeding amphibians. *PLoS One*. 10:e0123055.

Séré M, Thévenon S, Belem AMG, De Meeûs T. 2017. Comparison of different genetic distances to test isolation by distance between populations. *Heredity*. 119:55–63.

Shaffer HB, Gidiş M, McCartney-Melstad E, Neal KM, Oyamaguchi HM, Tellez M, Toffelmier EM. 2015. Conservation genetics and genomics of amphibians and reptiles. *Annu Rev Anim Biosci*. 3:113–138.

Shah V, McRae B. 2008. Circuitscape: a tool for landscape ecology. 7:62–66.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 31:3210–3212.

Smith KG, Lips KR, Chase JM. 2009. Selecting for extinction: nonrandom disease-associated extinction homogenizes amphibian biotas. *Ecol Lett*. 12:1069–1078.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP. 2007. Putting the 'landscape' in landscape genetics. *Heredity*. 98:128–142.

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues AS, Fischman DL, Waller RW. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science*. 306:1783–1786.

Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*. 144:389–399.

Thomson RC, Wright AN, Shaffer HB. 2016. *California amphibian and reptile species of special concern*. Oakland (CA): UC Press.

Title PO, Bemmels JB. 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*. 41:291–307.

Unglaub B, Steinfartz S, Drechsler A, Schmidt BR. 2015. Linking habitat suitability to demography in a pond-breeding amphibian. *Front Zool*. 12:9.

U.S. Fish and Wildlife Service. 2015. Endangered and threatened wildlife and plants; 90-day findings on 31 petitions. *Fed Regist*. 80:80583–80614.

Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN. 2008. Are hotspots of evolutionary potential adequately protected in southern California? *Biol Conserv*. 141:1648–1664.

Vandergast AG, Kus BE, Preston KL, Barr KR. 2019. Distinguishing recent dispersal from historical genetic connectivity in the coastal California gnatcatcher. *Sci Rep*. 9:1355.

Wake DB, Vredenburg VT. 2008. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proc Natl Acad Sci USA*. 105:11466–11473.

Walls SC, Barichivich WJ, Brown ME. 2013. Drought, deluge and declines: the impact of precipitation extremes on amphibians in a changing climate. *Biology (Basel)*. 2:399–418.

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

Wang IJ, Johnson JR, Johnson BB, Shaffer HB. 2011. Effective population size is strongly correlated with breeding pond size in the endangered California tiger salamander, *Ambystoma californiense*. *Conservation Genetics*. 12:911–920.

Wang IJ, Savage WK, Shaffer HB. 2009. Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). *Mol Ecol*. 18:1365–1374.

Waples RK, Larson WA, Waples RS. 2016. Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity (Edinb)*. 117:233–240.

Waples RS, Anderson EC. 2017. Purging putative siblings from population genetic data sets: a cautionary view. *Mol Ecol*. 26:1211–1224.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*. 38:1358–1370.

Willson JD, Hopkins WA. 2013. Evaluating the effects of anthropogenic stressors on source-sink dynamics in pond-breeding amphibians. *Conserv Biol*. 27:595–604.

Winter D, Green P, Kamvar Z, Gosselin T. 2017. mmod: modern measures of population differentiation (Version 1.3.3). Available from: https://CRAN.R-project.org/package=mmod

Winter DJ. 2012. MMOD: an R library for the calculation of population differentiation statistics. *Mol Ecol Resour*. 12:1158–1160.

Zamberletti P, Zaffaroni M, Accatino F, Creed IF, De Michele C. 2018. Connectivity among wetlands matters for vulnerable amphibian populations in wetlandscapes. *Ecol Model*. 384:119–127.