

ARTICLE



Landscape genetics of a sub-alpine toad: climate change predicted to induce upward range shifts via asymmetrical migration corridors

Paul A. Maier ^{1,2}✉, Amy G. Vandergast ³, Steven M. Ostoja⁴, Andres Aguilar ⁵ and Andrew J. Bohonak ¹

© The Author(s), under exclusive licence to The Genetics Society 2022

Climate change is expected to have a major hydrological impact on the core breeding habitat and migration corridors of many amphibians in the twenty-first century. The Yosemite toad (*Anaxyrus canorus*) is a species of meadow-specializing amphibian endemic to the high-elevation Sierra Nevada Mountains of California. Despite living entirely on federal lands, it has recently faced severe extirpations, yet our understanding of climatic influences on population connectivity is limited. In this study, we used a previously published double-digest RADseq dataset along with numerous remotely sensed habitat features in a landscape genetics framework to answer two primary questions in Yosemite National Park: (1) Which fine-scale climate, topographic, soil, and vegetation features most facilitate meadow connectivity? (2) How is climate change predicted to influence both the magnitude and net asymmetry of genetic migration? We developed an approach for simultaneously modeling multiple toad migration paths, akin to circuit theory, except raw environmental features can be separately considered. Our workflow identified the most likely migration corridors between meadows and used the unique cubist machine learning approach to fit and forecast environmental models of connectivity. We identified the permuted modeling importance of numerous snowpack-related features, such as runoff and groundwater recharge. Our results highlight the importance of considering phylogeographic structure, and asymmetrical migration in landscape genetics. We predict an upward elevational shift for this already high-elevation species, as measured by the net vector of anticipated genetic movement, and a north-eastward shift in species distribution via the network of genetic migration corridors across the park.

Heredity (2022) 129:257–272; <https://doi.org/10.1038/s41437-022-00561-x>

INTRODUCTION

Climate change is likely to have an important influence on genetic connectivity of many organisms this century (reviewed in Littlefield et al. 2019). Changes in climate have already caused phenological and range shifts as well as genetic modulation in many organisms (reviewed in Parmesan 2006), and specifically in amphibians (Alexander and Eischeid 2001; Pounds 2001; Pounds et al. 2006). Yosemite toads are a federally threatened species that is reputed to be declining in both distribution (Brown and Olsen 2013; Drost and Fellers 1994; Jennings and Hayes 1994; but see Lee et al. *in prep.*), and abundance (Sherman and Morton 1993). Climate change is predicted to have a disproportionately large impact on meadow hydrology, and projected to dramatically reduce the geographic range for Yosemite toads by 2100 (Smith and Tirpak 1988; Viers et al. 2013; US Fish & Wildlife Service 2014; Reich et al. 2018).

Yosemite toad tadpoles are obligate shallow pond specialists (mean depth 4.35 cm; Liang et al. 2017), and almost exclusively breed in mesic mountain meadows (Grinnell and Storer 1924; Karlstrom 1962; Ratliff 1985), making survival dependent upon

snowpack runoff and spring recharge. Significant larval mortality is often observed even during years with above-average snowpack (Sherman 1980; Sherman and Morton 1984, 1993; Brown et al. 2015). Freezing and early desiccation in ultra-shallow ponds often causes mass die-offs, exacerbated by water mold infections (*Saprolegnia diclina*) and flatworm (*Turbellaria*) predation (Sadinski et al. 2020). Adults and post-metamorphic juveniles are sensitive to seasonal temperatures, which directly determine their over-wintering body condition, fecundity, and ability to disperse (Morton 1981). All these climatic features and interactions may play a role in determining Yosemite toad movement and breeding patterns across the landscape, and hence the magnitude and net vector of migration. Therefore, a model of fine-resolution climate associations on Yosemite toad network connectivity may help address how and where climate change will fragment meadows in the future, and lead to a more specific understanding of how their distribution may shift.

In this study, we sought to analyze patterns of environmentally mediated genetic connectivity among Yosemite toad populations in Yosemite National Park (YOSE). Genetic connectivity fosters

¹Department of Biology, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182, USA. ²FamilyTreeDNA, Gene by Gene, 1445 N Loop W, Houston, TX 77008, USA. ³U.S. Geological Survey, Western Ecological Research Center, San Diego Field Station, San Diego, CA 92101, USA. ⁴USDA California Climate Hub, Agricultural Research Service, John Muir Institute of the Environment, University of California, Davis, 1 Shields Ave., Davis, CA 95616, USA. ⁵Department of Biological Sciences, California State University, Los Angeles, 5151 State University Dr., Los Angeles, CA 90032, USA. Associate editor: Sam Banks. ✉email: paulm@genebygene.com

Received: 16 May 2022 Revised: 24 August 2022 Accepted: 25 August 2022

Published online: 8 September 2022

long-term persistence of a species with discrete populations, by mitigating inbreeding depression, replenishing adaptive genetic variation, and promoting heterosis (Gotelli 1991; Hansson 1991; Whitlock et al. 2000; Gaggiotti 2003; Lowe and Allendorf 2010). From previous research, we broadly know that topography and climate influence migration (Wang 2012), and that meadow quality depends upon the network of nearby meadows (Berlow et al. 2013; Maier et al. 2022), but we lack an understanding of how specific climatic features such as snowmelt timing influence connectivity, or how previously under-sampled populations at lower elevation are affected. In addition, no study of Yosemite toad connectivity has considered environmental factors influencing toad occupancy of meadows, which may relate to larval and natal requirements, separately from migration corridor environment affecting only adult toads. While tadpoles develop and metamorphose in meadow pools and flooded areas, adult toads forage, hibernate, and can disperse into intervening habitat over 1 km per season (Martin 2008; Liang 2010). For biphasic species, estimates of connectivity can benefit from modeling between-site and at-site effects separately (Pflüger and Balkenhol 2014).

Our goals in this study were: (1) to identify specific climatic, topographic, soil, and vegetation features influencing connectivity, and (2) to forecast a spatial representation of future connectivity and net movement based on a “business-as-usual” climate change model. We also present several new approaches to address some general challenges of landscape genetics, including the challenge of defining migration environment. Our approach models the most likely migration path separately from modeling the influence of environment along that path. This allows a much richer consideration of hypothesized features and their combinations to be explored in a modeling framework. A similar approach has been taken with straight-line transects (Murphy et al. 2010a; van Strien et al. 2012; van Strien 2013), but we instead perform model testing to choose more realistic migration paths. For the most likely migration path, we then optimize the most likely corridor bandwidth of environmental influence since toads likely take multiple routes between any two meadows. This is analogous to circuit theory (McRae 2006; McRae and Beier 2007), except that it allows the separate modeling of paths and environmental influence.

We also introduce several novel approaches to address the challenge of building multivariate models of future genetic change. We apply a unique machine learning approach in our environmental models of connectivity that avoids overfitting, and effectively extrapolates future values. Machine learning approaches have been successfully applied to several studies (Murphy et al. 2010b; Hether and Hoffman 2012; Sylvester et al. 2018; Pless et al. 2021; Kittlein et al. 2022), helping to ameliorate problems identified with the assumptions of linear models (Peterman and Pope 2021). We include a “lineage effect” to account for known phylogeographic structure (Maier et al. 2019; Fig. 1A). Finally, we consider the effect of asymmetrical migration, which has been found at a smaller scale in Yosemite toads (Maier et al. 2022). We separately model the magnitude of genetic connectivity from its asymmetry, to allow an estimate of net asymmetrical shift (i.e., range shifts) under climate change.

Our results provide information that can be used to develop effective mitigation for habitat and climate related disturbance. In addition, our novel workflow provides a useful framework for characterizing genetic network structure in patch-limited species, and this approach may be extended to other systems.

MATERIALS AND METHODS

Spatial extent of sample collection

Population boundaries were highly correlated with meadow boundaries in a previous study (Maier et al. 2022). Thus, we sampled meadows within Yosemite National Park (YOSE) to maximize representation across all

known breeding locations from a recent 6-year survey effort (Lee et al. *in prep.*), and overlap with previous studies (Shaffer et al. 2000; Wang 2012; Berlow et al. 2013; Maier 2018; Maier et al. 2019, 2022; Fig. 1A). YOSE includes ~33% of known Yosemite toad meadows. Tadpoles were sampled across all available egg clutches, ponds, meadows, and two separate years (2012–2013) to maximize inclusion of available genetic diversity and reduce potential bias of oversampling close relatives relative to the total local population. A minimum of five samples was used per meadow if additional meadows were included within 1 km; otherwise 10 samples per meadow were used, unless insufficient samples were available. This scheme maximized intra- and inter-meadow sampling representation across the study area.

Molecular methods, ddRAD sequencing, and bioinformatics

We used a previously generated double-digest Restriction Site Associated DNA Sequencing (ddRADseq) haplotype dataset (described in Maier 2018; Maier et al. 2019) for all analyses of genetic differentiation and structure. The dataset was compiled and analyzed with Stacks v1.19 (Catchen et al. 2011, 2013) and filtered using a minimum 10× locus coverage, 75% complete loci, and minor allele frequency of 0.005. SNPs were called using the multinomial likelihood algorithm of Stacks, which calculates the likelihood of the observed genotypes given the per-base sequencing error rate. The final dataset included 2318 polymorphic loci for 535 individual Yosemite toad tadpoles from 90 populations in YOSE (Table S1). We included all individuals, as previously published analyses did not find any differences in genetic differentiation after removal of close relatives (Maier et al. 2022).

Genetic differentiation and asymmetry estimates

We used AMOVA-based F_{ST} described by Weir (1996) and implemented in Arlequin (Excoffier et al. 1992) and STACKS (Catchen et al. 2013) to represent bidirectional genetic differentiation:

$$F_{ST} = \frac{\sum_i n_i (\bar{p}_i - \bar{p})^2 / (r - 1) \bar{n}}{\bar{p}(1 - \bar{p})}$$

where \bar{p}_i is the estimated allele frequency for each biallelic SNP in the i th population of a pair, n_i is the count of alleles observed in that population, and r is the count of populations (2).

For asymmetrical migration estimates, we used the method of Sundqvist et al. (2016), which estimates directional G_{ST} (Nei and Chesser 1983) for a pair of populations A and B, with allele frequency vectors \mathbf{a} and \mathbf{b} , as follows:

$$G_{ST}^{A \rightarrow B} = \frac{|\mathbf{a} - \mathbf{c}^*|^2}{4 - |\mathbf{a} + \mathbf{c}^*|^2}$$

and:

$$G_{ST}^{B \rightarrow A} = \frac{|\mathbf{b} - \mathbf{c}^*|^2}{4 - |\mathbf{b} + \mathbf{c}^*|^2}$$

where \mathbf{c}^* represents the allele frequency vector for a hypothetical shared gene pool. The shared gene pool has allele frequencies that are the normalized geometric means of \mathbf{a} and \mathbf{b} for each allele k :

$$f(\mathbf{c}^*) = \frac{\sqrt{a_k b_k}}{\sum_k \sqrt{a_k b_k}}$$

Migration rates are then calculated using $N_e m \approx ((1/G_{ST}) - 1)/4$ (Wright 1931). We followed Maier et al. (2022) and only used the differential between emigration and immigration, yielding a relative term δM . Net immigration ($\delta M < 0$) or net emigration ($\delta M > 0$) have the advantage of not modeling any explicit parameter, and being less sensitive to island model assumptions, such as drift-migration equilibrium.

Optimizing migration paths

Landscape genetic studies often integrate the hypothesis-testing of environmental features with the hypothesis-testing of migration paths (Manel et al. 2003; Manel and Holderegger 2013). However, this may vastly simplify the number, combinations, and weights of feature space explored, since it requires each hypothesized resistance raster to be encoded as one combination of all covariates. We took a different approach, first optimizing the most likely migration paths with ResistanceGA using two simple features

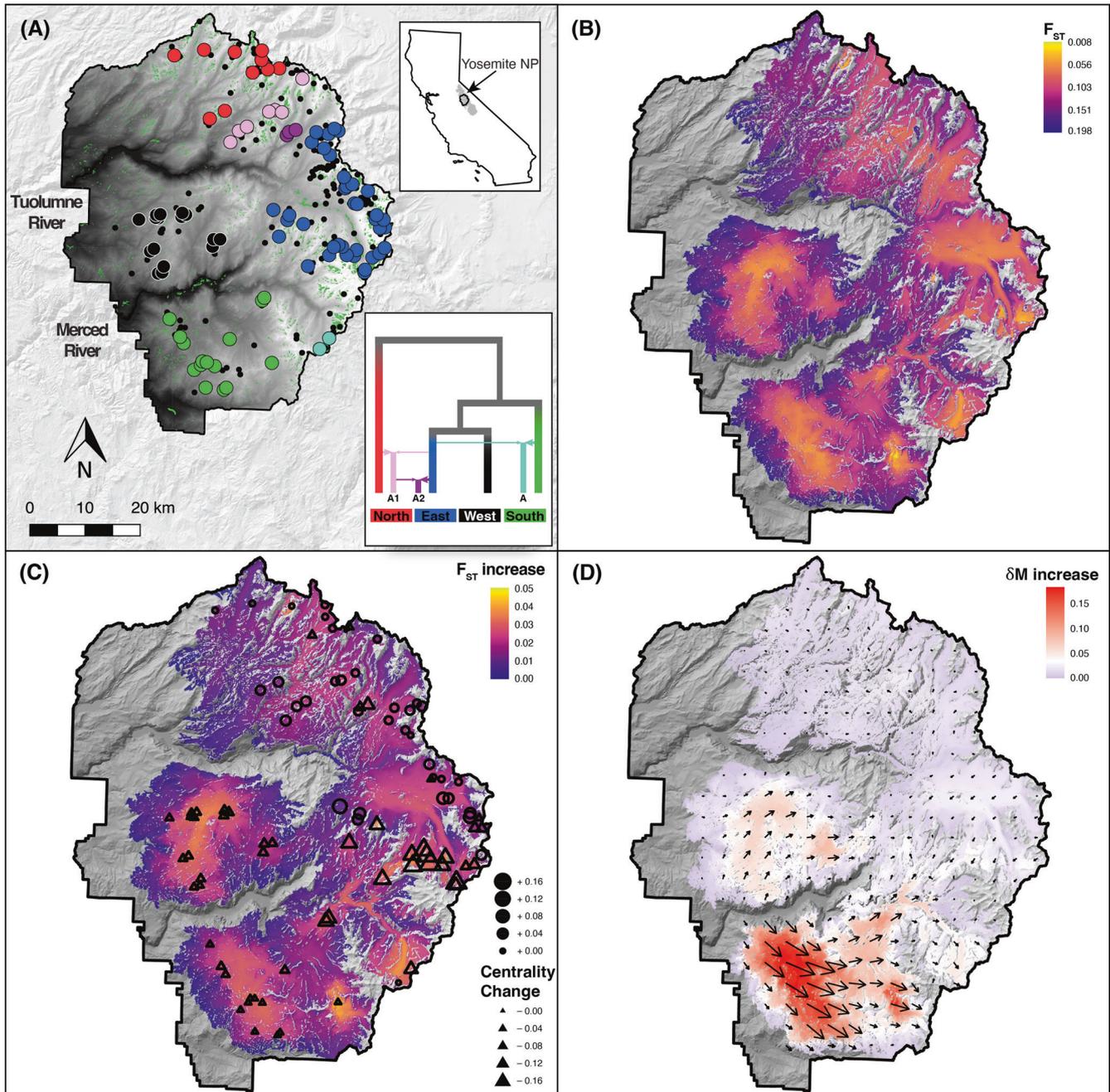


Fig. 1 Projected migrational shift after climate change. **A** Primary study area in Yosemite NP (YOSE), CA includes ~33% of known Yosemite toad populations. Top right inset shows the range of Yosemite toads in gray, and the boundary of YOSE black. Green polygons are all meadows within the park. Solid black circles indicate all known Yosemite toad meadows identified between 1915 and the present. Large circles indicate the meadows sampled and sequenced in the present study ($n = 90$). Colors correspond to phylogenetic lineages shown by the inset (Maier et al. 2019). Random noise is added to locations to protect the locations of this threatened species. **B** Mean per pixel F_{ST} predicted from present-day environmental features, based on all underlying migration corridors. **C** Projected increase in F_{ST} for the time interval 2070–2099 based on the future climate scenario RCP 8.5 (“business-as-usual”). Shift in the graph theoretical “centrality” depicts the increase (circles) or decrease (triangles) in how central a meadow will become to the network. **D** Projected asymmetrical shift in migration (δM) in the future time interval. Arrow length represents the underlying magnitude of asymmetrical shift, and arrow direction represent the net vector of change. Neutral colors are the mean value.

already known to influence Yosemite toad migration, then extracting mean environmental values to be used in a modeling framework.

Topographic complexity and moisture availability are already known to broadly influence genetic connectivity and site occupancy in the Yosemite toad (Martin 2008; Wang 2012; Berlow et al. 2013; Brown et al. 2015; Maier 2018; Maier et al. 2022). We hypothesized that a combination of these two environmental features may best explain migration paths between meadows. High-resolution (10 m) rasters of SRTM-derived slope

(Rabus et al. 2003) and vegetation class from ground-truthed aerial imagery (Keeler-Wolf et al. 2012) were selected to inform these two resistance surfaces, and were resampled to 30 m for computational feasibility. Raw pixel values of slope from 0–90° were used to inform the topographic resistance surface, whereas vegetation pixel values were first scored as ordinal resistance values between 1–10 based on moisture availability (Table S2). Both rasters were then transformed individually into resistance surfaces using the `SS_optim` function of ResistanceGA v4.2.2

(Peterman 2018). This procedure builds linear mixed models between F_{ST} and resistance distance through an environmental surface, then uses a genetic algorithm to explore which parameters optimize the model likelihood. After preliminary testing, a monomolecular curve was chosen to represent each transformation function, because both slope and decreasing moisture should monotonically increase resistance. The costDistance method (gdistance v1.3.6; van Etten 2017) was used to generate least cost paths (LCPs). Steep ridgetops were identified using a “gradient metrics” toolbox (Evans et al. 2014) and were scored as impenetrable to prevent unlikely, short routes from being preferred. We used a maximum value of 1×10^9 , and a maximum iteration count of 50, stopping after ten steps of no improvement in the objective function.

We combined slope and vegetation rasters by rescaling to a maximum value of 10, transforming with Resistance.tran function (ResistanceGA), and finally combining in proportions from 0.0 to 1.0 in increments of 0.1 (a total of 11 hypotheses). To choose the LCP model that best represents migration paths, we used the lme4 package v1.1.27.1 (Bates et al. 2015) to build linear mixed models between LCP distance and F_{ST} . Source and destination populations were used as random effects to control for non-independence of measurements. Geometric distance was used as the null model of isolation by straight-line distance, and a 30 km distance cutoff was used following previous research showing insignificant isolation-by-distance beyond that scale (Maier 2018; Maier et al. 2022). We also excluded improbable paths that crossed Yosemite and Hetch Hetchy Valleys. The 12 models were ranked, and the one with highest log likelihood and lowest AICc was selected.

Alternative corridor bandwidths

We chose six alternative corridor bandwidths of the optimal LCP model with which to extract environmental features (Fig. 2). This allowed us to directly test which bandwidth best captures environmental influence on genetic migration (see “Optimizing corridor bandwidths”). We considered two simple buffer distances of 100 and 500 m, values chosen relative to the mean yearly dispersal distance of 275 m (Liang 2010), and we calculated them using the gBuffer function (rgeos v0.5.8). Additionally, we considered four least cost corridor (LCC) bandwidths, based on varying thresholds for the accumulated cost surface that was used to produce the LCP model. In contrast to a 1-dimensional LCP, a LCC is a 2-dimensional raster surface showing the accumulated cost of moving between source and destination, for all pixels. Such corridors may potentially better represent the bandwidth of travel, because a multitude of reasonable paths may be considered instead of just one. The accCost function (gdistance) was applied to each meadow, and then the sum of cost values was taken between each pair of meadows. Thresholds for clipping each raster into a corridor were then defined as a lower quantile of all resistance values in the raster, from conservative to inclusive: 0.001, 0.005, 0.01, and 0.05. The remaining pixels were rescaled from 0 to 1, to be used as weights for environmental extraction. This allows the probability of migration at any given pixel to influence the extent to which that local environment is considered.

Environmental feature extraction

Remotely sensed and survey data were used to parameterize landscape genetic models. Features were taken from nine categories, detailed in Table 1: climate (California Basin Characterization Model v65; Flint et al. 2013), soil moisture (NDMI from LANDSAT 5; Masek et al. 2006), snowmelt (day of year from MODIS; Hall et al. 2002), geology (age of soil; Huber et al. 2003), topography (SRTM; Rabus et al. 2003), vegetation (Keeler-Wolf et al. 2012), fire (frequency; Cal Fire 2022), recreation (trails and road weighted by usage; Yosemite NP), and meadow network (Berlow et al. 2013; this study). Additionally, we extracted the same climate features under the future Community Climate System Model (CCSM v4), using representation concentration pathway (RCP) 8.5 (Riahi et al. 2011). The “business-as-usual” RCP 8.5 scenario models a future where emissions continue to rise throughout the twenty-first century, leading to increase in global mean surface temperature of 2.6–4.8°C by 2100 (IPCC 2014). For LCC bandwidths, gdalwarp (gdalUtils v2.0.3.2) was first used to resample each raster from its native resolution to the LCC resolution of 30 m. Then pixels within each bandwidth were summarized as a mean, sum, or proportion, depending upon the feature (Table 1). For simple buffers, this was achieved using the exact_extract function (exactextract v 0.7.2), and for LCCs, summary values were weighted by migration probability of that pixel, using cellStats (raster v3.5.15). In addition to “between-site” environment, we extracted “at-site” values for these same features by

using meadow polygon boundaries, then taking the difference between source and destination meadows. The meadow network group of features was only measured “at-site,” and the ones from this study include: degree, centrality, and clustering coefficient of each meadow (igraph v1.2.9; Csárdi and Nepusz 2006), as well as meadow area. Finally, two more features were added to account for isolation-by-distance (minimum path distance; “LCPdist”), and phylogeographic structure, encoded as time to most recent common ancestor (Maier et al. 2019) if source and destination are from two lineages, otherwise zero (“LineageCross”).

Optimizing corridor bandwidths

We trained a random forests model on each of the nine feature groups, to choose the optimal corridor bandwidth. This allows the freedom for one feature group (e.g., climate) to potentially have broader spatial influence on migration than another feature group (e.g., geology). Random forests analysis is a classification and regression tree method that uses an ensemble of bootstrap-aggregated decision trees to avoid overfitting (Breiman 2001). It can handle datasets containing many features, their higher-order interactions, and multicollinearity while modeling a response variable. We used 5000 trees for each model of F_{ST} versus environment, and tuned the hyperparameters using the caret package (v6.0.90; Kuhn 2008) with five-fold cross validation. Values of 1, 5, and 10 were explored for min.node.size (minimum leaves in each tree), and values between 1 and the number of features were explored for mtry (number of features sampled at each split). For each feature group, the best corridor bandwidth was chosen based the model with lowest root-mean-square error (RMSE) value.

Removing environmental redundancy

We performed principal component analysis (PCA) within each feature group to remove environmental redundancy and collinearity from the predictors. Although the predictive accuracy in tree-based machine learning approaches is robust to collinearity, correlated features are known to bias variable importance metrics (Strobl et al. 2007; Tološi and Lengauer 2011; Gregorutti et al. 2017). We focused on in-group redundancy in this step because related features from the same data source and resolution are most likely to contain redundant information. Between-group redundancy was reduced in a later step (see “Cubist models of migration and asymmetry”). The prcomp function in base R was used on groups with >2 features, after centering on zero and rescaling by standard deviation.

Cubist models of migration and asymmetry

Cubist regression trees were chosen to predict and forecast F_{ST} and δM . Although tree-based learning methods such as random forests and xgboost have many appealing properties, they may extrapolate poorly, particularly if relationships are linear (Breiman 2001; Heenkenda et al. 2015; Houborg and McCabe 2018; Meyer and Pebesma 2021). Preliminary testing showed this to be the case in models of δM . Cubist models (Quinlan 1992, 1993) mitigate this by replacing simple split rules with linear models at each branch in the decision tree. Effectively, a different regression equation is locally fit to each partition of the data. The two hyperparameters are the number of committees (number of trees with adjusted weights) and neighbors (nearby data points to average), used to adjust model predictions.

We trained a cubist model on both F_{ST} and δM using all non-redundant features from the previous step. The caret package was used to tune the hyperparameters with ten-fold cross validation. Values of 1, 10, 50, 75, and 100 were explored for committees, and values of 0, 1, 5, 7, and 9 were explored for neighbors. Permuted variable importance was calculated using the varImp function (caret), and then collinear features from different groups (VIF > 10) were removed, with a preference to keep those with higher importance. A final cubist training procedure was then rerun with the reduced set of features.

Forecasting future migration after climate change

Future values of both F_{ST} and δM were predicted 89 years in the future (climate features averaged between 1981–2010 in the present, and 2070–2099 in the future). We first projected the PC scores of future climate features onto the loadings of the PCA model using predict.pcomp. Then, we projected F_{ST} and δM into the future using predict.train (caret). To calculate the average of underlying processes at each pixel on the map, we calculated the average change in magnitude and asymmetry of migration

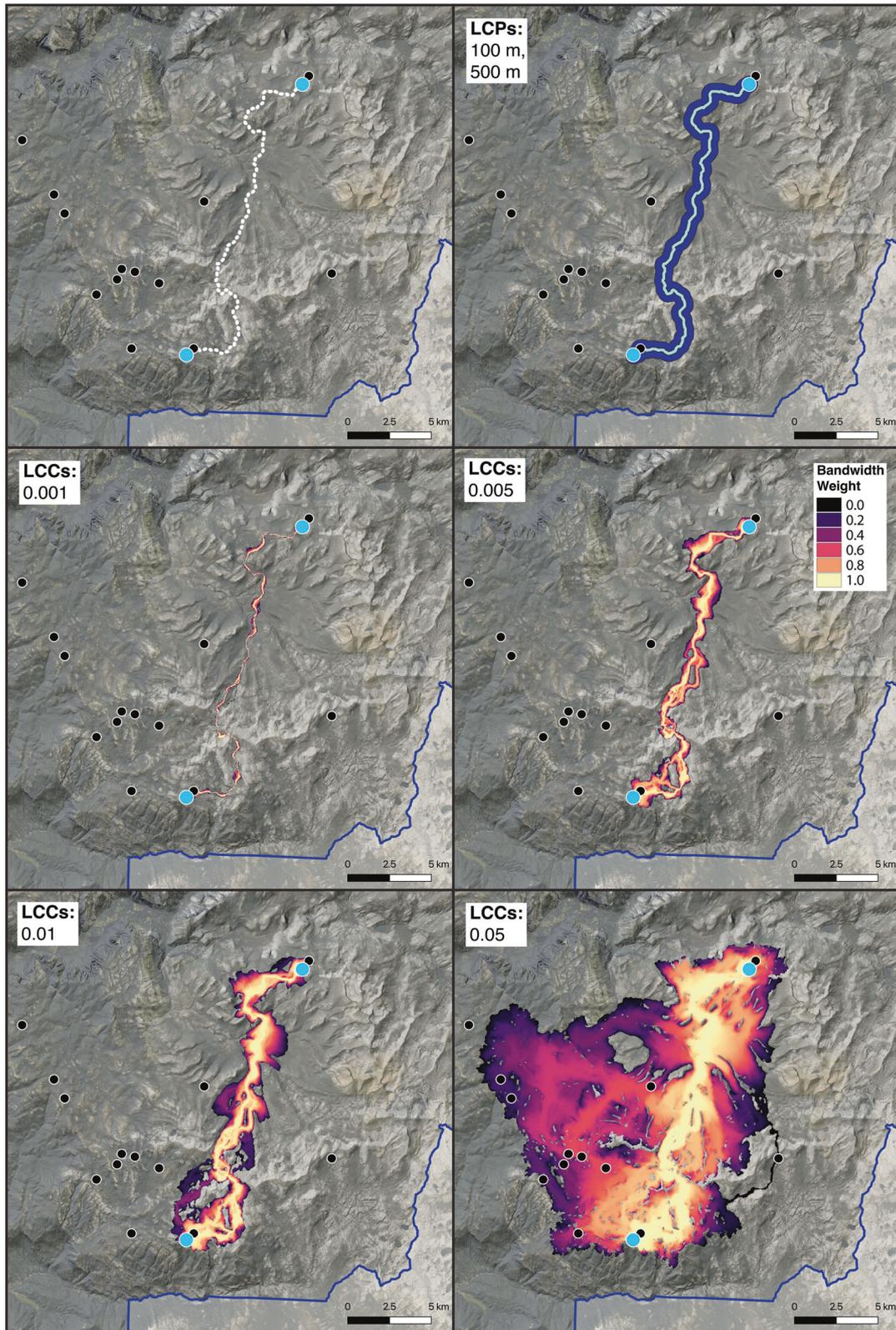


Fig. 2 **Alternative corridor bandwidths.** An example of the six corridor bandwidth types, for one pair of meadows (IDs 359 and 1543). Bandwidths for least cost paths (LCPs) are simple buffer distances of 100 and 500 m, chosen relative to the average seasonal dispersal distance of 275 m. Bandwidths for least cost corridors (LCCs) represent different thresholds for the accumulated cost between source and destination: quantiles of 0.001, 0.005, 0.01, and 0.05 among all pixel values were chosen. Remaining values were then rescaled from 0 to 1.

Table 1. Environmental features.

Feature group	Feature code(s)	Definition	Units	Data time interval	Stat. bt years	Stat. across pixels	Site type	Source	Original resolution	
Climate	aet_av, aet_cv	Actual evapotranspiration	mm	1981–2010	Mean, CV	Mean	at/bt	BCM v65	270 m	
	cwd_av, cwd_cv	Climatic water deficit	mm	2070–2099 ^a						
	pet_av, pet_cv	Potential evapotranspiration	mm							
	ppt_av, ppt_cv	Precipitation	mm							
	rch_av, rch_cv	Recharge	mm							
	run_av, run_cv	Runoff	mm							
	snp_av, snp_cv	Snowpack	mm							
	tmn_av, tmn_cv	Minimum temperature	°K							
	tmx_av, tmx_cv	Maximum temperature	°K							
	Soil moisture	NDMI_MAY_MEAN, NDMI_MAY_STD	Normalized Difference Moisture Index (NDMI): May	n/a	1984–2013	Mean, SD	Mean	at/bt	LANDSAT 5 ^c	30 m
NDMI_SEP_MEAN, NDMI_SEP_STD		Normalized Difference Moisture Index (NDMI): September								
Snowmelt	MODIS_MEAN, MODIS_SD	Day of Year (DOY) for snowmelt	Julian Day	2001–2018	Mean, SD	Mean	at/bt	MODIS ^d	500 m	
Geology	GEOL	Geological age of soil	Kya	n/a	n/a	Mean	at/bt	USGS	10 m	
	DEM	Elevation	m	n/a	n/a	Mean	at/bt	SRTM ^e	10 m	
Topography	SLOPE	Slope	degrees			Mean				
	SLOPE2ND	Slope second derivative	n/a			Mean				
	CTI	Compound topographic index	n/a			Mean				
	CURVE	Curvature	n/a			Mean				
	HLOAD	Heat load index	n/a			Mean				
	ROUGH	Roughness index	n/a			Mean				
	SAR	Surface to area ratio	n/a			Mean				
	SPI	Slope position index	n/a			Mean				
	SRR	Surface relief ratio	n/a			Mean				
	TRASP	Radiation index	n/a			Mean				
	pctEAST, pctNEAST, pctNORTH, pctNWEST, pctWEST, pctS, pctSOUTH, pctSEAST	Aspect (8 cardinal directions)	n/a			Prop.				
	Vegetation	mdwpctIntFlood	Proportion transiently flooded meadow	n/a	n/a	n/a	Prop.	at	Yosemite NP LIDAR	10 m
		mdwpctSemiPermFlood	Proportion semi-permanently flooded meadow				Prop.	at		
mdwpctShSedge		Proportion short-hair sedge in meadow				Prop.	at			
mdwpctWillow		Proportion willow in meadow				Prop.	at			
pctCONHERB		Proportion conifers and herbaceous				Prop.	bt			
pctCONIFER		Proportion coniferous vegetation				Prop.	bt			

Table 1. continued

Feature group	Feature code(s)	Definition	Units	Data time interval	Stat. bt years	Stat. across pixels	Site type	Source	Original resolution
Fire	pctCONSHRUB	Proportion conifers and shrub				Prop.	bt		
	pctCONWOOD	Proportion conifers and hardwood				Prop.	bt		
	pctDMEADOW	Proportion dry meadow				Prop.	bt		
	pctHARDWOOD	Proportion hardwood				Prop.	bt		
	pctRIPARIAN	Proportion riparian				Prop.	bt		
	pctROCK	Proportion rock				Prop.	bt		
	pctSHRUB	Proportion chaparral				Prop.	bt		
	STREAMW ^b	Stream crossings (per pixel, weighted by magnitude)				Weight. Sum	bt		
	pctWATER	Proportion standing water, permanent pool, or stream				Prop.	bt		
	pctWMEADOW	Proportion wet meadow				Prop.	bt		
Fire	FIRE	Fires per year (per pixel); or distance to nearest perimeter	Frequency	1908–2020	n/a	Sum	bt	Cal Fire ^f	10 m
Recreation	ROADW ^b	Road crossings (per pixel, weighted by traffic level)	Frequency	n/a	n/a	Weight. Sum	bt	Yosemite NP	10 m
Meadow network	TRAILW ^b	Trail crossings (per pixel, weighted by traffic level)							
	Degree	Number of neighboring meadow connections	n/a	n/a	n/a	Value	at	This study	n/a
	Centrality	Eigen centrality score within meadow network	n/a					This study	n/a
	ClustCoeff	Probability that meadow neighbors are themselves neighbors	n/a					This study	n/a
	NetworkSim	Network-booster breeding probability	n/a					(Berlow et al. 2013)	n/a
	Prob100GLM	Breeding probability, ignoring network	n/a					(Berlow et al. 2013)	n/a
	Area	Area of meadow	km ²					YNP LiDAR	10 m

^aBCM features sampled from the time interval 1981–2010 were taken as present-day, and those from 2070–2099 were taken as future.
^bFeatures meant to represent linear barriers to dispersal (i.e., streams, roads, and trails) were rasterized with each pixel representing presence or absence of the barrier. Additionally, weights were applied (i.e., roads were weighted by level of traffic, trails by frequency of use, and streams by order).
^cLANDSAT surface reflectance images were corrected using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm, which corrects raw images for water vapor, ozone, shadows, and other atmospheric aberrations (Masek et al. 2006; Schmidt et al. 2013). Only images taken between the months of May and September were used from LANDSAT, to better reflect the active season of Yosemite toads. The NDMI (Normalized Difference Moisture Index) was calculated as $NDMI = (NIR - SWIR1)/(NIR + SWIR1)$, where NIR = pixel values from the near infrared band, and SWIR1 = pixel values from the short-wave infrared 1 band.
^dMODIS images represent daily albedo from a spectroradiometer. Daily MODIS images were derived from previous studies (Dozier et al. 2008; Dozier and Frew 2009; Painter et al. 2009) and summarized in Octave v3.2.3 and ArcGIS v10.1.
^eSRTM data were summarized using ArcGIS Geomorphometry & Gradient Metrics (Evans et al. 2014).
^fFire perimeters were overlaid and converted into frequency data.

Table 2. ResistanceGA results.

Surface	k^a	AIC	AICc	$R^2_m{}^b$	$R^2_c{}^c$	LL ^d	Equation	Shape	Max
Slope	4	-16,094.181	-16,093.710	0.436	0.598	8051.090	Inverse-reverse monomolecular	1.119	736,398.455
Vegetation	4	-16,001.624	-16,001.153	0.421	0.603	8004.812	Inverse-reverse monomolecular	0.515	685,016.210
Distance	2	-15,238.290	-15,238.153	0.174	0.444	7621.145	n/a	n/a	n/a
Null	1	-14,356.485	-14,356.439	0.000	0.312	7179.242	n/a	n/a	n/a

Shape and max are the best parameters after optimizing each model. Table S3 and Fig. S1 contain details of the transformation function compared to original values.

^aDegrees of freedom.

^bMarginal R^2 .

^cConditional R^2 .

^dLog likelihood.

for all overlapping corridors. To do this, we scaled each pairwise estimate by its corridor bandwidth (0–1), with less probable routes receiving less weight. We took the average of overlapping pixel values to estimate mean change in connectivity (F_{ST}). For changes in migration asymmetry (δM), the net direction of change was calculated per pixel by summing all the underlying vectors (dx , dy) into a net vector, then calculating the net direction as $\text{atan2}(dy, dx)$ radians.

To assess how the total network of meadows will shift in the future, we assigned edge attributes for each pair of meadows of $\max(F_{ST}) - F_{ST}$, and calculated Kleinberg's hub centrality score (Kleinberg 1999), which is the principal eigenvector of the weighted adjacency matrix $A^T A$. Hub centrality represents park-wide connectedness, with the highest values receiving the most overall (direct and indirect) gene flow.

RESULTS

Optimizing migration paths

ResistanceGA found Yosemite toad migration to have a gradual monomolecular relationship with slope, with a shape of 1.119 and maximum value of 736,398 (Table 2). This means that a 30° slope elicits a resistance value of <2000, and a 60° slope elicits a resistance value of <40,000, but an inflection point occurs near 70°, which elicits a value of >100,000 (Table S3). This is consistent with observations of toads climbing moderately steep drainages but seldom breeding at meadows isolated by extremely steep terrain (pers obs.). Vegetation showed a much less gradual monomolecular relationship than slope, with a shape of 0.515 and maximum value of 685,016. Although this weighting would favor wet meadow habitat, all vegetation types except dry rock outcrops would be plausible migration paths (Table S2 and Fig. S1).

The LCP model with 0.9 of slope and 0.1 vegetation (including impenetrable ridgeline barriers) had the highest model ranking based on log-likelihood, AIC, and R^2_m metrics (Table 3). The second highest ranked model also gave slope a weight of 0.8, suggesting topography has a stronger influence on Yosemite toad migration paths than vegetation moisture. Hence, we used this LCP model as the most plausible set of toad migration paths for optimizing corridor bandwidths. Due to the topographic complexity of mountain ridgelines bisecting YOSE, this model found only one path connecting each of the two low-elevation lineages (Y-South, Y-West) to their adjacent high-elevation lineage (Y-East; Fig. S2).

Environmental feature extraction

Graph theoretical metrics of degree, centrality, and clustering coefficient highlighted the pattern of meadows in the Y-East lineage as the best-connected ones in the park (Table S1 and Fig. S3). This is partly because values of F_{ST} tend to be lower, and genetic diversity tends to be higher in this high-elevation lineage (Table S1). However, the other reason is spatial configuration:

toads in Y-South, Y-West, and Y-North lineages can only exit those regions by passing through the Y-East lineage, due to topographic constraints of river valleys. After ordinating each group of features onto its respective eigenvectors, the remaining dataset was mostly uncorrelated (Fig. S4).

Optimizing corridor bandwidths

For every group of features except for geology (climate, soil moisture, snowmelt, topography, vegetation, fire, and recreation), the random forests model chose the broadest corridor bandwidth of LCC 0.05 (Table 4). The best bandwidth for geology was LCP 100 m, although with only one variable this relationship was weak ($R^2 = 0.137$), whereas most other feature groups had $R^2 > 0.5$. The LCC 0.05 bandwidth includes the primary LCP, and typically incorporates a multitude of paths with similar probability of migration, as well as peripheral habitat that may occasionally or indirectly influence toads (Fig. 2). This general pattern of LCC 0.05 bandwidths is similar to the idea of circuit theory (McRae 2006; McRae and Beier 2007), except that our method allows each feature to be modeled independently in its raw form.

Cubist models of migration and asymmetry

After hyperparameter tuning, the two cubist models of F_{ST} and δM each attained strong model fit, with ten-fold cross-validated $R^2 = 0.88$ and $R^2 = 0.78$, respectively (Table 5 and Fig. 3). Both models benefited from the maximum number of committees (100) and an intermediate number of neighbors (5; Tables S5 and S6). Variable importance scores showed a pattern of relatively few between-site features with high importance in the F_{ST} model, but many at-site features of similar importance in the δM model (Fig. 4 and Tables S7 and S8). This makes sense given that F_{ST} mostly explains the overall magnitude of connectivity between sites, whereas δM is expected to explain environmental contrasts between meadows which might favor connectivity in one direction.

Based on the loadings of the PCA models (Figs. S5 and S6), we interpreted the most important features influencing Yosemite toad connectivity (F_{ST}) within their corridors. Many climatic features were among the top ranked ones: snow melt-off day-of-year variability, snow runoff and temperature variability, recharge, climatic water deficit mean and variability, maximum temperature, and potential evapotranspiration (Fig. 4). We also found that vegetation type along corridors has a large influence on level of connectivity: rocky and shrubby habitats, proportion of water, proportion of wet meadow, and dry meadow with shrubby cover. Similarly, LANDSAT-derived soil moisture had an influence. The most important topographic features included radiation indices (particularly related to SE, NW, and SW slopes), and general topographic complexity. Interestingly, trail crossings (weighted by traffic level) were found to be a highly important feature.

Table 3. Migration path results.

Slope	Vegetation	R^2_m ^a	R^2_c ^b	AIC	AICc	BIC	LL ^c
0.9	0.1	0.104	0.912	-6176.878	-6176.823	-6151.904	3093.439
0.8	0.2	0.102	0.911	-6163.112	-6163.056	-6138.137	3086.556
0.4	0.6	0.100	0.912	-6159.875	-6159.820	-6134.901	3084.938
0.5	0.5	0.098	0.911	-6152.689	-6152.634	-6127.715	3081.345
1	0	0.090	0.910	-6143.724	-6143.668	-6118.749	3076.862
0.7	0.3	0.097	0.910	-6143.199	-6143.144	-6118.225	3076.600
0.6	0.4	0.096	0.910	-6139.191	-6139.136	-6114.217	3074.596
0.3	0.7	0.094	0.910	-6138.701	-6138.646	-6113.727	3074.350
0.2	0.8	0.093	0.910	-6135.075	-6135.020	-6110.101	3072.538
0	1	0.103	0.908	-6126.692	-6126.637	-6101.718	3068.346
0.1	0.9	0.096	0.905	-6119.832	-6119.777	-6094.858	3064.916
None	None	0.086	0.904	-6101.830	-6101.775	-6076.856	3055.915

Ranked linear mixed models of LCP migration path hypotheses. "None" represents the null hypothesis of straight-line isolation-by-distance.

^aMarginal R^2 .

^bConditional R^2 .

^cLog likelihood.

Table 4. Corridor bandwidth results.

Feature group	Bandwidth	mtry	min.node.size	RMSE ^a	RMSE s.d.	R^2 ^b	R^2 s.d.	MAE ^c	MAE s.d.
Climate	LCCs (0.05)	8	5	0.021	0.002	0.674	0.030	0.016	0.001
Fire	LCCs (0.05)	1	10	0.031	0.001	0.285	0.070	0.024	0.001
Traffic	LCCs (0.05)	1	1	0.026	0.002	0.490	0.044	0.020	0.001
Geology	LCPs (100 m)	1	10	0.034	0.002	0.137	0.024	0.027	0.001
Moisture	LCCs (0.05)	2	5	0.022	0.002	0.620	0.035	0.017	0.001
Meltoff	LCCs (0.05)	1	5	0.024	0.002	0.563	0.041	0.018	0.001
Topography	LCCs (0.05)	7	5	0.021	0.002	0.664	0.031	0.016	0.001
Vegetation	LCCs (0.05)	6	5	0.021	0.001	0.675	0.030	0.016	0.001

The best bandwidth in each feature group is shown. A full model ranking is in Table S4.

^aRoot mean squared error.

^b R^2 value is based on ten-fold cross validation.

^cMean absolute error.

Finally, as expected, isolation-by-distance along the shortest migration path, and lineage effect, both were important features.

Although at-site (contrast between meadows) features were less important in predicting F_{ST} , several are noteworthy: meadow size, meadow vegetation type, and geological age of the soil. Meadow vegetation type is a classification of meadow polygons into four categories: proportion of semi- to permanently flooded, transiently flooded, willow, and short-hair sedge (more xeric). The PCA-projected feature appeared to be a gradient between semi- to permanently flooded meadow and short-hair sedge. Many of the same climatic features, indices of radiation, and topographic complexity were also important at-site.

For migration asymmetry (δM), some of the most important at-site (contrast between meadows) features were in the meadow network group: meadow size (area), degree (count of neighbors), and probability of occupancy, which is also informed by the network of available nearby meadows (Berlow et al. 2013). Many climatic and moisture-related features had similar importance to the model. LANDSAT-derived soil moisture was found to be a large influence, with fall conditions more important than spring conditions. The same gradient in meadow vegetation between semi- to permanently flooded meadow and short-hair sedge was ranked high, as was willow cover to a lesser extent. The most important topographic features included southern-facing aspect,

radiation indices (including SE and NE slopes), elevation, and topographic complexity measures such as slope, roughness, and surface relief ratio. Many climatic features were important, including those relating to snowpack (mean and variability, runoff, recharge), temperature (variability of minimum and maximum, range of extremes), climatic water deficit (mean and variability), and potential evapotranspiration. The geological age of soil was also an important feature.

Forecasting future migration after climate change

Our model predicted a continuous surface of Yosemite toad connectivity that highlights regional corridors of high flow, and "pinch-points" of low connectivity (Fig. 1B). In general, regions of high connectivity tend to occur at or below the lineage level, and landscape features inhospitable to toad occupancy or migration tend to bisect these regions (e.g., ridgelines, low fire-prone areas, areas of sparse meadow habitat). The Y-North lineage, which is balkanized by a myriad of canyons and ridgelines, apparently has low connectivity compared to the others.

Our future model of F_{ST} change suggests greatest reduction in connectivity to Y-West, Y-South, and certain admixed areas in the south, such as Merced Pass (Y-South) and the East-South-A admixed lineage (Fig. 1C). We compared the graph theoretical metric of centrality, which computes the normalized first

Table 5. Cubist models of F_{ST} and δM .

Model	Committees	Neighbors	RMSE ^a	RMSE s.d.	R^2 ^b	R^2 s.d.	MAE ^c	MAE s.d.
F_{ST}	100	5	0.013	0.002	0.880	0.026	0.009	0.001
δM	100	5	0.029	0.003	0.776	0.033	0.021	0.001

The best model for each is shown. A full model ranking from hyperparameter tuning is in Tables S5 and S6.

^aRoot mean squared error.

^b R^2 value is based on ten-fold cross validation.

^cMean absolute error.

eigenvector of how central a meadow is to the entire network of toad migration, and found a shift from south and west, to north and east. This indicates that a higher proportion of future toad migration will occur at higher elevations and latitudes (Y-East and Y-North) compared with the present.

Our future model of δM change strongly suggested a pattern of net asymmetrical movement from west to east, toward higher elevation areas (Fig. 1D). Given the pinch points in southern migration corridors, this would encourage net inter-lineage movement to occur from Y-South to Y-East up the M. Fork Merced River (north of the Clark Range), or up the S. Fork Merced River directly into the higher elevation Clark Range. Similarly, Y-West is only loosely connected to Y-East, with a single likely path up Tenaya Creek into the Tuolumne River watershed. These two vectors were the most strongly supported patterns of predicted range shift, although minor vectors within the high-elevation Y-East and Y-North lineages suggest some net movement northward. Broadly, the F_{ST} and δM results are consistent, suggesting a range shift toward higher elevation and latitude by 2100.

DISCUSSION

Impact of climate change on connectivity

Our models revealed that climate has strongly contributed to Yosemite toad genetic connectivity and forecasted a range shift toward higher elevations and latitudes (Figs. 1 and 4). Climatic features relating to snowpack variability were the most important for both F_{ST} and δM models, and by far the strongest F_{ST} feature overall. Snow runoff and groundwater recharge variability also ranked very high, for both migration corridors and meadows, consistent with occupancy modeling work that showed these features to be good predictors of site presence range-wide (Viers et al. 2013). Broadly, changes to snowpack and associated runoff are expected to have greatest impact on amphibian phenology and persistence, since snow can account for 80% of total runoff during dry summer months (Corn 2003; Stewart 2009). Recently, network methods have shown that the number of Sierra Nevada meadows that offer refugia from climate change (i.e., less climate variability) is projected to decline, and this will reduce connectivity across multiple meadow-dwelling species (Maher et al. 2017). In another meadow-dwelling species, Belding's ground squirrel (*Urocitellus beldingi*), network genetic connectivity among meadows was indeed related to climate refugia (Morelli et al. 2017). Our results are suggestive as to which changing climate attributes could further fragment Yosemite toad meadows but may also offer an opportunity to pinpoint refugial meadows for the species.

In the future, climate change is expected to dramatically change the hydrology of Sierra Nevada meadows. If carbon emissions continue unabated, by the end of the twenty-first century there will be an estimated 7 °C rise in average springtime temperature, 64% drop in springtime snowpack volume, and 50 day phenological shift to earlier runoff of snowmelt (Reich et al. 2018). This effect will not be uniform however, and some meadows may become climate refugia for toads (Smith and Tirpak 1988; Viers et al. 2013; Reich et al. 2018). Our work suggests that lower elevation Yosemite toads in the Y-South and Y-West lineages will

track wetter and more stable conditions as they recede upward and northeastward. Prehistorically, there is evidence that these toads have contracted and expanded their distribution cyclically throughout the Pleistocene (Maier et al. 2019), although possibly at a slower rate. This is also consistent with geologically younger alluvial soils playing a role in connectivity magnitude and direction, given that toad meadow habitat only formed ~10,000 years ago as glaciers receded and deposited these soils (Wood 1975). Our mapped projections of range shifts (Fig. 1C, D) could be used as a guide to protect or even assist migration as it becomes necessary to help toads keep pace of local desiccation.

Impact of topography

We found southern and eastern aspects, along with associated indices of heat load and radiation, to be along the most important topographic features for Yosemite toad connectivity. South-facing slopes have been found to influence Yosemite toad patch suitability in previous work (Liang and Stohlgren 2011), and the influence of solar radiation may play a role in efficient and successful larval development prior to ponds desiccating (Mullally 1953; Mullally and Cunningham 1956; Brattstrom 1962). Although topographic complexity and elevation were important to δM , topographic complexity only ranked 21% as important as the top F_{ST} feature (Table S7). This may be partly explained by the fact that corridors were chosen based on lower slope, and so this "pre-modeling" step partialled out some of the inherent importance of topography. We further discuss this limitation of our approach below (see "Implications for forecasting climate impacts with landscape genetics"). Our LCP model testing did identify the 90% slope 10% vegetation model as best fit to the data, consistent with a similar pattern found using variance partitioning on micro-satellite markers (Wang 2012). Interestingly, the effect of elevation on δM suggests that toads prefer to change elevations, and previous work (Maier et al. 2022) suggests a preference toward meadows with a higher slope position index (closer to ridgetops). Possibly, Yosemite toads are behaviorally biased toward upland dispersal that maximizes the likelihood of encountering suitably wet, sub-alpine breeding meadows. Negative geotaxis, or the tendency toward uphill movements, has previously been reported in other alpine animals (Baur 1986; Peterson 1997).

Impact of vegetation and soil

Rocky and shrubby habitats were the most important vegetation feature found in the F_{ST} model, possibly because they represent more xeric elements of toad corridors (e.g., western whitethorn, talus slopes) that inhibit migration. However, we note that the opposite is possible, as shrubby post-fire habitat counter-intuitively provides necessary shade and facilitates connectivity in two amphibian species in Yellowstone NP (Spear et al. 2005; Murphy et al. 2010b). The importance of a gradient in vegetation type within meadows, from semi-permanently flooded meadow to xeric short-hair sedge, presents an interesting opportunity to leverage this feature for future ecological research. Previous research has shown the Yosemite toad to utilize a gradient of meadow habitat, from immature stages and adult males in lower wet areas, to the more readily dispersing adult females occupying

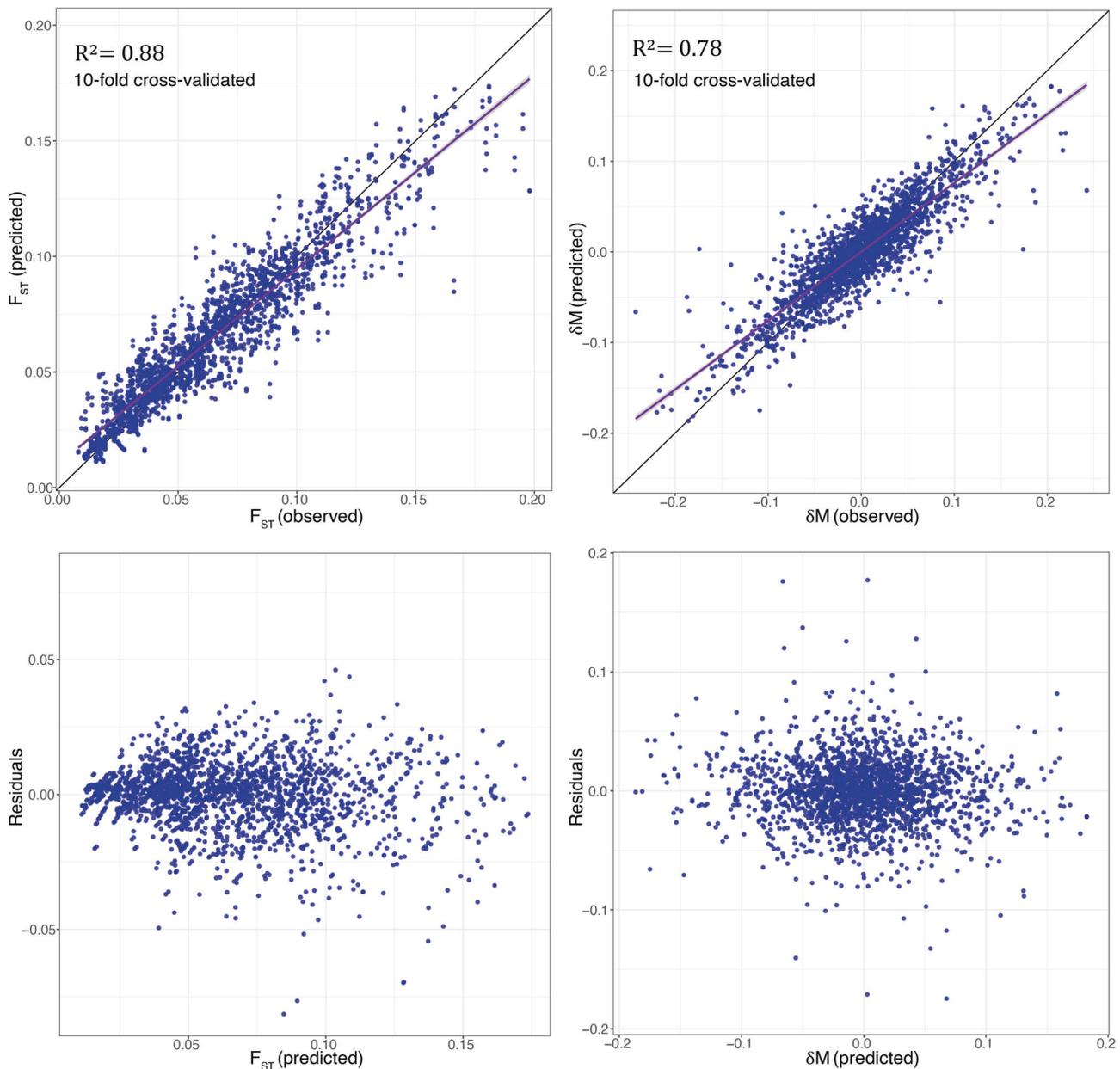


Fig. 3 Cubist model fit for F_{ST} and δM . Model fit of environmentally predicted F_{ST} and δM , using the cubist machine learning method. In the top row: predicted versus observed values, with linear trend lines in purple. In the bottom row: residuals versus predicted values, showing approximate homoscedasticity of residuals. All models were assessed using ten-fold cross validation and had coefficients of determination $R^2 = 0.88$ (F_{ST}), and $R^2 = 0.78$ (δM).

upper and rockier areas (Morton and Pereyra 2010). Future work could assess whether such a gradient influences emigration and immigration rates. Finally, LANDSAT-derived soil moisture was important in both models, with late season moisture playing a stronger role than spring moisture. This hints at an influence of predictably moist soil substrates for migrating adult toads, as they are foraging or searching for hibernacula. Moist surfaces during the dry season are an essential source for toads to absorb water via their pelvic drink patch, and the speed at which they replenish body fluid osmolality is known to influence where they aggregate (Reynolds and Christian 2009).

Impact of foot traffic

Trails as linear barriers were important predictors of Yosemite toad connectivity in the model of F_{ST} . The impact of trails is unsurprising

given the magnitude of recreation YOSE experiences each year. In 2019, 4.4 million people drove into YOSE, 155,578 of whom hiked on trails and camped in the backcountry (National Park Service 2022). This user load has gradually accumulated to 125-fold the number of visitors one century ago. Automobiles began driving over Tioga Road in 1900, and the route became popular for businessmen beginning in the 1920s (Trexler 1975). The trail system was slowly built up from existing Native American trails during the 1860s by the U.S. Geological Survey of California, and later in the 1920s and 1930s by the Sierra Club and National Park Service (Bingaman 1968). Future mark-recapture studies might reveal that these linear barriers also impact demographic connectivity. In the future, our results may be effectively combined with corridor design methods (Chetkiewicz et al. 2006; Beier et al. 2008; Sawyer et al. 2011) to help maintain species persistence.



Fig. 4 Variable importance plots. Final non-redundant features selected for cubist models of F_{ST} (top panel) and δM (bottom). Permuted importance based on RMSE is normalized to a maximum value of 100%. Actual features used, based on PCA scores, are shown on the left. Interpretations of each feature based on PCA loadings (see Figs. S5 and S6) are shown on the right. “At-site” features (taken as the difference between source and destination values) are bold, italicized, and underlined.

Meadow network effects

This landscape genetics study is among the first to isolate source-specific environmental effects on genetic connectivity (Murphy et al. 2010a; Dileo et al. 2014), and the first to accomplish this goal for Yosemite toads. Site contrasts in network features such as meadow area, probability of occupancy (Berlow et al. 2013), and degree (number of adjacently connected meadows) had a strong influence on asymmetry of migration δM . This echoes the finding that Yosemite toad meadows are locally organized into neighborhoods of “hub” and “satellite” meadows, with larger, flatter hub meadows receiving net genetic input from satellites (Maier et al. 2022). In the present study, we found a similar pattern at a larger spatial scale. At-site climatic contrasts were decisively forecasted to invoke net genetic movement toward the more climatically protected high-elevation meadows in the Y-East lineage (Fig. 1D).

Environmentally, our model (Fig. 4) predicted this asymmetry of genetic movement between source and destination meadows to be driven by contrasting meadow vegetation (e.g., vegetation type, willow cover), differing patterns of insolation (e.g., southern aspect, surface relief ratio, heat load), snow-related differences (e.g., snowpack mean, variability, runoff, recharge), and varying soil properties (e.g., fall soil moisture, and geological age of soil). Soil age could influence water retention ability of natal pools, given the recent deposition of alluvium into some (but not all) meadows by glaciation (Wood 1975). At-site feature quality may encourage emigration or immigration, by influencing resource abundance and population density, causing individuals to disperse toward meadows that reduce competition, and fitness costs (Travis et al. 1999; Matthysen 2005; Mathieu et al. 2010; Pflüger and Balkenhol 2014). Toads at more suitable meadows may benefit from higher carrying capacity through more natal pond habitat, more willow hibernacula, higher solar input to expedite tadpole metamorphosis, and less inter-annual variability in pond water supply by snowmelt. As climate change progresses, toads may consolidate at more stable “hub” sites; there is some evidence for a decadal shift in occupancy away from isolated sites, and toward meadows with consistent occupancy (Lee et al. in prep.).

An added benefit of this network perspective on meadows is that graph theory can be applied to entire networks of species connectivity, to elucidate emergent properties of indirect gene flow such as “centrality” or “modularity” (Dyer and Nason 2004; Garroway et al. 2008; Dyer et al. 2010). Although F_{ST} is anticipated to increase for nearly all meadows (likely due to population declines and genetic drift isolating sites), the relative differences in that change are expected to re-center the meadow network eastward and northward (Fig. 1C). Effectively, these two network approaches to Yosemite toad connectivity (asymmetrical δM shift from source-destination contrasts, and network-wide shift in F_{ST} centrality) are complimentary and support the same conclusion.

Implications for forecasting climate impacts with landscape genetics

Predictive spatial models in landscape genetics and genomics are an important, burgeoning class of tools for summarizing evolutionary processes and incorporating them into species or landscape conservation planning (Vandergast et al. 2008; Sork et al. 2010; Fitzpatrick and Keller 2015). In this study, we provided a novel network method for individually modeling variables of interest (Fig. 2), which fills an important void in the landscape genetic toolbox. Many studies generate hypothesized “resistance rasters” of environmental variables (McRae 2006; Zeller et al. 2012), but these are not well-suited to multivariate hypotheses, and can be biased by the subjectivity of translating raw values into resistance values (Shirk et al. 2010; Spear et al. 2010). Moreover, this approach often does a poor job of exploring total parameter space, and may converge on the wrong environmental model (Graves et al. 2013). The alternate approach, transect analysis, does extract raw values of variables from rectilinear corridors, but at the

cost of losing biologically meaningful dispersal corridor locations (van Strien et al. 2012; van Strien 2013). Our method combines the biological plausibility of LCP-optimized dispersal corridors (Douglas 1994) with a multivariate modeling process of raw environmental variables, making inferences potentially more realistic. Our corridor hypothesis-testing framework was able to accommodate the multitude of equally probable dispersal paths that Circuit Theory allows (McRae 2006; McRae and Beier 2007), without giving up the flexibility of a multivariate modeling process. Our approach also does not assume that habitat suitability models explain both at-site and between-site environmental preferences as is typically done (Epps et al. 2007; Storfer et al. 2007), including for the Yosemite toad (Wang 2012). Finally, our approach accounts for phylogeographic signal in genetic differentiation estimates, which is often an unaccounted source of variance (Dyer et al. 2010).

One potential weakness of our approach is that it must first develop a simple environmental model of LCPs, before extracting features and building a nuanced model. Thus, if the first model is over simplistic, biased, or uninformed by existing knowledge of broad environmental features, then the corridors may contain some irrelevant landscape values. However, we carefully surveyed the existing literature and used the two most broadly important features (slope and vegetation moisture) to form LCP hypotheses for subsequent testing. The model ranking consistently showed slope to be the more important (but not only) feature relevant to migration paths, and all models outperformed straight-line transects, which are used by default in other transect methods (Murphy, et al. 2010a; van Strien et al. 2012; van Strien 2013). Another weakness is that machine learning approaches are not very transparent about the sign (positive or negative) of each feature’s relationship in the model, and only describe importance in terms of effect on RMSE.

Although the cubist algorithm (Quinlan 1992, 1993) has not previously been applied in landscape genetics, it can have some advantages over random forests, which has been used several times (Murphy et al. 2010b; Hether and Hoffman 2012; Sylvester et al. 2018; Pless et al. 2021; Kittlein et al. 2022). The method recursively applies linear models to each branch in a collection of rule-based decision trees, which has inherent advantages. Unlike random forests, which we found to vastly underpredict extreme values and extrapolate poorly for our model of δM , the cubist algorithm can accurately extrapolate linear conditions outside the observed range. Like random forests and xgboost, cubist is an ensemble method that sidesteps the problems of overfitting, multicollinearity, and non-linearity inherent in the Ordinary Least Squares (OLS) linear regression class of models. Cross-validated performance of the cubist algorithm was higher than linear mixed models, random forests, xgboost, and neural networks in our preliminary experiments. Cubist models should be explored further in other landscape genetic questions.

Conclusions and implications for management

Despite numerous efforts to find causative agents for the decline of Yosemite toads, such as UV radiation (Sadinski 2004), exotic predators (Grasso et al. 2010), meadow grazing (Roche et al. 2012a, b; Matchett et al. 2015), chemical deposition (Bradford and Gordon 1994; Davidson 2004; Sadinski 2004), and chytridiomycosis (Dodge et al. 2012; Lindauer and Voyles 2019; Lindauer et al. 2020), no clear patterns have emerged (Brown et al. 2015). This is in stark contrast to the “smoking-gun” inferences about other amphibian declines and recoveries in California (Knapp and Matthews 2000; Vredenburg et al. 2010; Knapp et al. 2016). However, our results based on modeling over one third of the species range indicate that climate and recent climate change strongly influence connectivity in the Yosemite toad. This is occurring both at the level of breeding meadows and adult migration corridors and may be a central threat to the future persistence of the species. Specifically, we found some evidence

that lower-elevation meadows will become increasingly disconnected to each other in response to changing snowpack and runoff conditions, which could force lower-elevation toads to either adapt or move upward. Unfortunately, both adaptation and migration become much less likely in smaller and more fragmented populations, where genetic drift and inbreeding can contribute to extinction risks (Nunney and Campbell 1993; Sacchi et al. 1998; Spielman et al. 2004; Schlaepfer et al. 2018; Bozzuto et al. 2019). This is consistent with the observation that low elevation meadows have experienced higher rates of extirpations (Drost and Fellers 1994, 1996). We suggest that land managers could account for these movement patterns by prioritizing the protection of likely climate change corridors and refugia.

DATA AVAILABILITY

No original genetic data were produced in this study. Demultiplexed fastq files of double-digest RADseq data are available at NCBI GenBank SRA under BioProject [PRJNA558546](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA558546). Scripts, genotype data, and environmental data used to perform analyses in this study are deposited at Dryad (<https://doi.org/10.5061/dryad.xsj3tx99h>). The U.S. Geological Survey metadata record associated with this study can be found at the USGS Science Data Catalog (<https://doi.org/10.5066/9KABYPU>).

REFERENCES

- Alexander MA, Eischeid JK (2001) Climate variability in regions of amphibian declines. *Conserv Biol* 15:930–942
- Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- Baur B (1986) Patterns of dispersion, density and dispersal in alpine populations of the land snail *Arianta arbustorum* (L.) (Helicidae). *Holarct Ecol* 9:117–125
- Beier P, Majka DR, Spencer WD (2008) Forks in the road: choices in procedures for designing wildland linkages. *Conserv Biol* 22:836–851
- Berlow EL, Knapp R, Ostoja SM, Williams RJ, McKenny H, Matchett JR et al. (2013) A network extension of species occupancy models in a patchy environment applied to the Yosemite toad (*Anaxyrus canorus*). *PLoS ONE* 8:e72200
- Bingaman JW (1968) Pathways: a story of trails and men. End-Kian Publishing Company, Lodi, CA
- Bozzuto C, Biebach I, Muff S, Ives AR, Keller LF (2019) Inbreeding reduces long-term growth of Alpine ibex populations. *Nat Ecol Evol* 3:1359–1364
- Bradford D, Gordon M (1994) Acidic deposition as an unlikely cause for amphibian population declines in the Sierra Nevada, California. *Biol Conserv* 69:155–161
- Brattstrom BH (1962) Thermal control of aggregation behavior in tadpoles. *Herpetologica* 18:38–46
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Brown C, Hayes MP, Green GA, Macfarlane DC, Lind AJ (2015) Yosemite toad conservation assessment. USDA Forest Service report. Sonora, CA
- Brown C, Olsen AR (2013) Bioregional monitoring design and occupancy estimation for two Sierran amphibian taxa. *Freshw Sci* 32:675–691
- Cal Fire (2022) Fire perimeters. FRAP Mapp. <https://frap.fire.ca.gov/mapping/gis-data/>
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3 Genes Genomes Genet* 1:171–182
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22:3124–3140
- Chetkiewicz C-LB, St. Clair CC, Boyce MS (2006) Corridors for conservation: integrating pattern and process. *Annu Rev Ecol Syst* 37:317–342
- Corn PS (2003) Amphibian breeding and climate change importance of snow in the mountains. *Conserv Biol* 17:622–625
- Csárdi G, Nepusz T (2006) The igraph software package for complex network research. *Int J Complex Syst* 1695:1–9
- Davidson C (2004) Declining downwind: amphibian population declines in California and historical pesticide use. *Ecol Appl* 14:1892–1902
- Dileo MF, Siu JC, Rhodes MK, López-Villalobos A, Redwine A, Ksiazek K et al. (2014) The gravity of pollination: integrating at-site features into spatial analysis of contemporary pollen movement. *Mol Ecol* 23:3973–3982
- Dodge C, Cheng T, Vredenburg V (2012) Exploring the evidence of a historical chytrid epidemic in the Yosemite toad by PCR analysis of museum specimens
- Douglas DH (1994) Least-cost path in GIS using an accumulated cost surface and slopelines. *Cartographica* 31:37–51
- Dozier J, Frew J (2009) Computational provenance in hydrologic science: a snow mapping example. *Philos Trans R Soc A Math Phys Eng Sci* 367:1021–1033
- Dozier J, Painter TH, Rittger K, Frew JE (2008) Time-space continuity of daily maps of fractional snow cover and albedo from MODIS. *Adv Water Resour* 31:1515–1526
- Drost C, Fellers G (1994) Decline of frog species in the Yosemite section of the Sierra Nevada. National Park Service report. Davis, CA
- Drost C, Fellers G (1996) Collapse of a regional frog fauna in the Yosemite area of the California Sierra Nevada, USA. *Conserv Biol* 10:414–425
- Dyer RJ, Nason JD (2004) Population graphs: the graph theoretic shape of genetic structure. *Mol Ecol* 13:1713–1727
- Dyer RJ, Nason JD, Garrick RC (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Mol Ecol* 19:3746–3759
- Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. *J Appl Ecol* 44:714–724
- van Etten J (2017) R Package gdistance: distances and routes on geographical grids. *J Stat Softw* 76:1–21
- Evans J, Oakleaf J, Cushman S, Theobald D (2014) An ArcGIS toolbox for surface gradient and geomorphometric modeling
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Fitzpatrick MC, Keller SR (2015) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol Lett* 18:1–16
- Flint LE, Flint AL, Thorne JH, Boynton R (2013) Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. *Ecol Process* 2:1–21
- Gaggiotti OE (2003) Genetic threats to population persistence. *Ann Zool Fennici* 40:155–168
- Garway CJ, Bowman J, Carr D, Wilson PJ (2008) Applications of graph theory to landscape genetics. *Evol Appl* 1:620–630
- Gotelli NJ (1991) Metapopulation models: the rescue effect, the propagule rain, and the core-satellite hypothesis. *Am Nat* 138:768–776
- Grasso RL, Coleman RM, Davidson C (2010) Palatability and antipredator response of Yosemite toads (*Anaxyrus canorus*) to nonnative brook trout (*Salvelinus fontinalis*) in the Sierra Nevada Mountains of California. *Copeia* 2010:457–462
- Graves TA, Beier P, Royle JA (2013) Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Mol Ecol* 22:3888–3903
- Gregorutti B, Michel B, Saint-Pierre P (2017) Correlation and variable importance in random forests. *Stat Comput* 27:659–678
- Grinnell J, Storer TI (1924) Animal life in the Yosemite: an account of the mammals, birds, reptiles, and amphibians in a cross-section of the Sierra Nevada. University of California Press, Berkeley, CA
- Hall DK, Riggs GA, Salomonson VV, Digirolamo NE, Bayr KJ (2002) MODIS snow-cover products. *Remote Sens Environ* 83:181–194
- Hansson L (1991) Dispersal and connectivity in metapopulations. *Biol J Linn Soc* 42:89–103
- Heenkenda MK, Joyce KE, Maier SW, de Bruin S (2015) Quantifying mangrove chlorophyll from high spatial resolution imagery. *ISPRS J Photogramm Remote Sens* 108:234–244
- Hether TD, Hoffman EA (2012) Machine learning identifies specific habitats associated with genetic connectivity in *Hyla squirella*. *J Evol Biol* 25:1039–1052
- Houborg R, McCabe MF (2018) A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. *ISPRS J Photogramm Remote Sens* 135:173–188
- Huber N, Bateman P, Wahrhaftig C (2003) Geologic map of Yosemite National Park and Vicinity, California: a digital database. Menlo Park, CA
- IPCC (2014) Climate change 2014: synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland
- Jennings M, Hayes M (1994) Amphibian and reptile species of special concern in California. California Department of Fish & Game report. Rancho Cordova, CA
- Karlstrom EL (1962) The toad genus *Bufo* in the Sierra Nevada of California: ecological and systematic relationships. *University Calif Publ Zool* 62:1–104
- Keeler-Wolf T, Reyes ET, Menke JM, Johnson DN, Karavidas DL (2012) Yosemite National Park vegetation classification and mapping project report. National Park Service report. Fort Collins, CO
- Kittlén MJ, Mora MS, Mapelli FJ, Austrich A, Gaggiotti OE (2022) Deep learning and satellite imagery predict genetic diversity and differentiation. *Methods Ecol Evol* 13:711–721
- Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. *J ACM* 46:604–632
- Knapp RA, Fellers GM, Kleeman PM, Miller DAW, Vredenburg VT, Rosenblum EB et al. (2016) Large-scale recovery of an endangered amphibian despite ongoing exposure to multiple stressors. *Proc Natl Acad Sci* 113:11889–11894

- Knapp RA, Matthews KR (2000) Non-native fish introductions and the decline of the mountain yellow-legged frog from within protected areas. *Conserv Biol* 14:428–438
- Kuhn M (2008) Building predictive models in R using the caret package. *J Stat Softw* 28:1–26
- Lee SR, Ostoja SM, Maier PA, Matchett JR, McKenny HC, Brooks ML et al. Distribution and spatio-temporal variation of Yosemite toad populations in Sierra Nevada national parks (in preparation)
- Liang CT (2010) Habitat modeling and movements of the Yosemite toad (*Anaxyrus (=Bufo) canorus*) in the Sierra Nevada, California. Ph.D. Dissertation. University of California, Davis
- Liang CT, Grasso RL, Nelson-Paul JJ, Vincent KE, Lind AJ (2017) Fine-scale habitat characteristics related to occupancy of the Yosemite toad, *Anaxyrus canorus*. *Copeia* 105:120–127
- Liang CT, Stohlgren TJ (2011) Habitat suitability of patch types: a case study of the Yosemite toad. *Front Earth Sci* 5:217–228
- Lindauer AL, Maier PA, Voyles J (2020) Daily fluctuating temperatures decrease growth and reproduction rate of a lethal amphibian fungal pathogen in culture. *BMC Ecol* 20:1–9
- Lindauer AL, Voyles J (2019) Out of the frying pan, into the fire? Yosemite toad (*Anaxyrus canorus*) susceptibility to *Batrachochytrium dendrobatidis* after development under drying conditions. *Herpetol Conserv Biol* 14:185–198
- Littlefield CE, Krosby M, Michalak JL, Lawler JJ (2019) Connectivity for species on the move: supporting climate-driven range shifts. *Front Ecol Environ* 17:270–278
- Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Mol Ecol* 19:3038–3051
- Maher SP, Morelli TL, Hershey M, Flint AL, Flint LE, Moritz C et al. (2017) Erosion of refugia in the Sierra Nevada meadows network with climate change. *Ecosphere* 8:1–17
- Maier PA (2018) Evolutionary past, present, and future of the Yosemite toad (*Anaxyrus canorus*): a total evidence approach to delineating conservation units. Ph.D. Dissertation. University of California Riverside
- Maier PA, Vandergast AG, Ostoja SM, Aguilar A, Bohonak AJ (2019) Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). *Evolution* 73:2476–2496
- Maier PA, Vandergast AG, Ostoja SM, Aguilar A, Bohonak AJ (2022) Gene pool boundaries for the Yosemite toad (*Anaxyrus canorus*) reveal asymmetrical migration within meadow neighborhoods. *Front Conserv Sci* 3:1–14
- Manel S, Holderegger R (2013) Ten years of landscape genetics. *Trends Ecol Evol* 28:614–621
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* 18:189–197
- Martin DL (2008) Decline, movement and habitat utilization of the Yosemite toad (*Bufo canorus*): an endangered anuran endemic to the Sierra Nevada of California. Ph.D. Dissertation. University of California, Santa Barbara
- Masek JG, Vermote EF, Saleous NE, Wolfe R, Hall FG, Huemmrich KF et al. (2006) A landsat surface reflectance dataset, 1990–2000. *IEEE Geosci Remote Sens Lett* 3:68–72
- Matchett JR, Stark PB, Ostoja SM, Knapp RA, McKenny HC, Brooks ML et al. (2015) Detecting the influence of rare stressors on rare species in Yosemite National Park using a novel stratified permutation test. *Sci Rep* 5:1–12
- Mathieu J, Barot S, Blouin M, Caro G, Decaëns T, Dubs F et al. (2010) Habitat quality, conspecific density, and habitat pre-use affect the dispersal behaviour of two earthworm species, *Aporrectodea icterica* and *Dendrobaena veneta*, in a mesocosm experiment. *Soil Biol Biochem* 42:203–209
- Matthysen E (2005) Density-dependent dispersal in birds and mammals. *Ecography* 28:403–416
- McRae B (2006) Isolation by resistance. *Evolution* 60:1551–1561
- McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci USA* 104:19885–19890
- Meyer H, Pebesma E (2021) Predicting into unknown space? estimating the area of applicability of spatial prediction models. *Methods Ecol Evol* 12:1620–1633
- Morelli TL, Maher SP, Lim MCW, Kastely C, Eastman LM, Flint LE et al. (2017) Climate change refugia and habitat connectivity promote species persistence. *Clim Chang Responses* 4:8
- Morton M (1981) Seasonal changes in total body lipid and liver weight in the Yosemite toad. *Copeia* 1981:234–238
- Morton M, Pereyra M (2010) Habitat use by Yosemite toads: life history traits and implications for conservation. *Herpetol Conserv Biol* 5:388–394
- Mullally D (1953) Observations on the ecology of the toad *Bufo canorus*. *Copeia* 1953:182–183
- Mullally D, Cunningham J (1956) Aspects of the thermal ecology of the Yosemite toad. *Herpetologica* 12:57–67
- Murphy MA, Dezzani R, Pilliod D, Storfer A (2010a) Landscape genetics of high mountain frog metapopulations. *Mol Ecol* 19:3634–3649
- Murphy MA, Evans JS, Storfer A (2010b) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91:252–261
- National Park Service (2022) National park service visitor use statistics
- Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. *Ann Hum Genet* 47:253–259
- Nunney L, Campbell KA (1993) Assessing minimum viable population size: demography meets population genetics. *Trends Ecol Evol* 8:234–239
- Painter TH, Rittger K, McKenzie C, Slaughter P, Davis RE, Dozier J (2009) Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sens Environ* 113:868–879
- Parnesan C (2006) Ecological and evolutionary responses to recent climate change. *Annu Rev Ecol Syst* 37:637–669
- Peterman WE (2018) Surfaces using genetic algorithms ResistanceGA: an R package for the optimization of resistance. *Methods Ecol Evol* 9:1638–1647
- Peterman WE, Pope NS (2021) The use and misuse of regression models in landscape genetic analyses. *Mol Ecol* 30:37–47
- Peterson MA (1997) Host plant phenology and butterfly dispersal: causes and consequences of uphill movement. *Ecology* 78:167–180
- Pflüger FJ, Balkenhol N (2014) A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Mol Ecol* 23:2146–56
- Pless E, Saarman NP, Powell JR, Caccone A, Amatulli G (2021) A machine-learning approach to map landscape connectivity in *Aedes aegypti* with genetic and environmental data. *Proc Natl Acad Sci USA* 118:1–8
- Pounds J (2001) Climate and amphibian declines. *Nature* 410:639–640
- Pounds JA, Bustamante MR, Coloma LA, Consuegra JA, Fogden MPL, Foster PN et al. (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439:161–167
- Quinlan JR (1992) Learning with continuous classes. *Aust Jt Conf Artif Intell* 92:343–348
- Quinlan JR (1993) Combining instance-based and model-based learning. *Mach Learn Proc* 1993 93:236–243
- Rabus B, Eineder M, Roth A, Bamler R (2003) The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS J Photogramm Remote Sens* 57:241–262
- Ratliff RD (1985) Meadows in the Sierra Nevada of California: state of knowledge. U.S. Forest Service report. Berkeley, CA
- Reich KD, Berg N, Walton DB, Schwartz M, Sun F, Huang X et al. (2018) Climate change in the Sierra Nevada: California's water future. UCLA Center for Climate Science report. Los Angeles, CA
- Reynolds SJ, Christian KA (2009) Environmental moisture availability and body fluid osmolality in introduced toads. *J Herpetol* 43:326–331
- Riahi K, Rao S, Krey V, Cho C, Chirkov V, Fischer G et al. (2011) RCP 8.5—a scenario of comparatively high greenhouse gas emissions. *Clim Change* 109:33–57
- Roche LM, Allen-Diaz B, Eastburn DJ, Tate KW (2012a) Cattle grazing and Yosemite toad (*Bufo canorus*, Camp) breeding habitat in Sierra Nevada meadows. *Rangel Ecol Manag* 65:56–65
- Roche LM, Latimer AM, Eastburn DJ, Tate KW (2012b) Cattle grazing and conservation of a meadow-dependent amphibian species in the Sierra Nevada. *PLoS ONE* 7:e35734
- Sacchi I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature* 392:491–494
- Sadinski W (2004) Amphibian declines: causes. U.S. Geological Survey report. La Crosse, Wisconsin
- Sadinski W, Gallant AL, Cleaver JE (2020) Climate's cascading effects on disease, predation, and hatching success in *Anaxyrus canorus*, the threatened Yosemite toad. *Glob Ecol Conserv* 23:e01173
- Sawyer SC, Epps CW, Brashares JS (2011) Placing linkages among fragmented habitats: do least-cost models reflect how animals use landscapes? *J Appl Ecol* 48:668–678
- Schlaepfer DR, Braschler B, Rusterholz HP, Baur B (2018) Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. *Ecosphere* 9 e02488
- Schmidt G, Jenkerson C, Masek J, Vermote E, Gao F (2013) Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. U.S. Geological Survey report. Reston, VA
- Shaffer H, Fellers G, Magee A, Voss S (2000) The genetics of amphibian declines: population substructure and molecular differentiation in the Yosemite toad, *Bufo canorus* (Anura, Bufonidae) based on single-strand conformation polymorphism analysis (SSCP) and mitochondrial DNA sequence data. *Mol Ecol* 9:245–257
- Sherman CK (1980) A comparison of the natural history and mating system of two anurans: Yosemite toads (*Bufo canorus*) and Black toads (*Bufo exsul*). Ph.D. Dissertation. University of Michigan
- Sherman CK, Morton ML (1984) The toad that stays on its toes. *Nat Hist* 93:72–78

- Sherman CK, Morton ML (1993) Population declines of Yosemite toads in the eastern Sierra Nevada of California. *J Herpetol* 27:186–198
- Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI (2010) Inferring landscape effects on gene flow: a new model selection framework. *Mol Ecol* 19:3603–3619
- Smith JB, Tirpak DA (1988) The potential effects of global climate change on the United States: draft: report to Congress. U.S. Environmental Protection Agency, Office of Policy, Planning and Evaluation, Office of Research and Development
- Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H et al. (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Mol Ecol* 19:3806–3823
- Spear SF, Balkenhol N, Fortin M-J, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Mol Ecol* 19:3576–3591
- Spear SF, Peterson CR, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Mol Ecol* 14:2553–2564
- Spielman D, Brook BW, Briscoe DA, Frankham R (2004) Does inbreeding and loss of genetic diversity decrease disease resistance? *Conserv Genet* 5:439–448
- Stewart IT (2009) Changes in snowpack and snowmelt runoff for key mountain regions. *Hydrol Process* 23:78–94
- Storfer A, Murphy M, Evans J, Goldberg C, Robinson S, Spear S et al. (2007) Putting the 'landscape' in landscape genetics. *Heredity* 98:128–142
- van Strien M (2013) Advances in landscape genetic methods and theory: lessons learnt from insects in agricultural landscapes. Ph.D. Dissertation. ETH Zürich
- van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* 21:4010–23
- Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinform* 8:25
- Sundqvist L, Keenan K, Zackrisson M, Prodöhl P, Kleinhans D (2016) Directional genetic differentiation and relative migration. *Ecol Evol* 6:3461–3475
- Sylvester EVA, Beiko RG, Bentzen P, Paterson I, Horne JB, Watson B et al. (2018) Environmental extremes drive population structure at the northern range limit of Atlantic salmon in North America. *Mol Ecol* 27:4026–4040
- Toloşi L, Lengauer T (2011) Classification with correlated features: Unreliability of feature ranking and solutions. *Bioinformatics* 27:1986–1994
- Travis MJM, Murrell DJ, Dytham C (1999) The evolution of density-dependent dispersal. *Proc R Soc Lond Ser B Biol Sci* 266:1837–1842
- Trexler KA (1975) The Tioga road: a history, 1883–1961. Yosemite Natural History Association, El Portal, CA
- U.S. Fish & Wildlife Service (2014) Endangered and threatened wildlife and plants; endangered status for the Sierra Nevada yellow-legged frog and the northern distinct population segment of the mountain yellow-legged frog, and threatened status for the Yosemite toad: final rule. *Fed Regist* 79:1–56. <https://www.federalregister.gov/documents/2014/04/29/2014-09488/endangered-and-threatened-wildlife-and-plants-endangered-species-status-for-sierra-nevada>
- 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
- Viers JH, Purdy SE, Peek RA, Fryjoff-Hung A, Santos NR, Katz JV et al (2013) Montane meadows in the Sierra Nevada: changing hydroclimatic conditions and concepts for vulnerability assessment. Centre for Watershed Sciences report. Davis, CA
- Vredenburg VT, Knapp RA, Tunstall TS, Briggs CJ (2010) Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proc Natl Acad Sci USA* 107:9689–9694
- Wang JJ (2012) Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. *Divers Distrib* 18:1033–1041
- Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer Associates, Inc., Sunderland, MA
- Whitlock MC, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. *Heredity* 84:452–457
- Wood SH (1975) Holocene stratigraphy and chronology of mountain meadows, Sierra Nevada, California. Ph.D. Dissertation. California Institute of Technology
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159
- Zeller KA, McGarigal K, Whiteley AR (2012) Estimating landscape resistance to movement: a review. *Landsc Ecol* 27:777–797

ACKNOWLEDGEMENTS

We thank Michael Hernandez, Johanne Boulat, Alexa Killion, Mara MacKinnon, Steven Lee, Amy Patten, Heidi Rodgers, Cameron DeMaranville, Keevan Harding and Brenna Blessing for assistance with field specimen collection. Jared Grummer provided technical advice that improved library preparation for sequencing. Rui Hu provided invaluable suggestions and feedback about machine learning. Rachel Shoop contributed feedback that significantly improved the figures and presentation of this study. All animal handling was performed in accordance with SDSU animal care and use protocol #13-03-001B and adhered to the regulations of NPS research permits. We thank Jonathan Q. Richmond, Brian J. Halstead, and Leonard Nunney for their comments that significantly improved the writing of this manuscript. This research was supported by U.S. Geological Survey, Ecosystems Mission Area, Natural Resource Preservation Program Research grant awarded to Steven Ostojka, Eric Berlow, and Matthew Brooks. In addition, PAM received funding from the Harold & June Memorial, Jordan D. Covin, and ARCS scholarships. The UC Merced Sierra Nevada Institute provided housing and accommodations for fieldwork. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

AUTHOR CONTRIBUTIONS

PAM, SMO, AA, AGV, and AJB designed the study; PAM and SMO collected tissue samples; PAM performed the laboratory work and conducted the analyses, with guidance from AJB, AGV, and AA. PAM wrote the manuscript, with input from AGV, AJB, SMO, and AA.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41437-022-00561-x>.

Correspondence and requests for materials should be addressed to Paul A. Maier.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.