

Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*)

Paul A. Maier,^{1,2,3}  Amy G. Vandergast,⁴  Steven M. Ostoja,⁵  Andres Aguilar,⁶ 
 and Andrew J. Bohonak¹ 

¹Department of Biology, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182

²FamilyTreeDNA, Gene by Gene, 1445 N Loop W, Houston, TX 77008

³E-mail: maierpa@gmail.com

⁴U.S. Geological Survey, Western Ecological Research Center, San Diego Field Station, 4165 Spruance Road, Suite 200, San Diego, CA 92101

⁵USDA California Climate Hub, Agricultural Research Service, John Muir Institute of the Environment, University of California, Davis, 1 Shields Ave., Davis, CA 95616

⁶Department of Biological Sciences, California State University, Los Angeles, 5151 State University Dr, Los Angeles, CA 90032

Received March 13, 2018

Accepted October 14, 2019

Species endemic to alpine environments can evolve via steep ecological selection gradients between lowland and upland environments. Additionally, many alpine environments have faced repeated glacial episodes over the past two million years, fracturing these endemics into isolated populations. In this “glacial pulse” model of alpine diversification, cycles of allopatry and ecologically divergent glacial refugia play a role in generating biodiversity, including novel admixed (“fused”) lineages. We tested for patterns of glacial pulse lineage diversification in the Yosemite toad (*Anaxyrus [Bufo] canorus*), an alpine endemic tied to glacially influenced meadow environments. Using double-digest RADseq on populations densely sampled from a portion of the species range, we identified nine distinct lineages with divergence times ranging from 18 to 724 thousand years ago (ka), coinciding with multiple Sierra Nevada glacial events. Three lineages have admixed origins, and demographic models suggest these fused lineages have persisted throughout past glacial cycles. Directionality indices supported the hypothesis that some lineages recolonized Yosemite from east of the ice sheet, whereas other lineages remained in western refugia. Finally, refugial niche reconstructions suggest that low- and high-elevation lineages have convergently adapted to similar climatic niches. Our results suggest glacial cycles and refugia may be important crucibles of adaptive diversity across deep evolutionary time.

KEY WORDS: Ecological selection, glacial refugia, glacial pulse, hybridization, Pleistocene, secondary contact.

Vicariance by glacial expansion and retreat has long been thought to drive speciation (Mayr 1942; Anderson 1948; Stebbins 1950; Anderson 1953; Haffer 1969). Pleistocene ice ages have fragmented and promoted speciation for a high proportion of boreal organisms over the past two million years (Bernatchez and Wilson 1998; Weir and Schluter 2004). Species endemic to alpine or montane environments are prone to new lineage formation because their narrow distributions and specialized habitat require-

ments are closely linked to the process of glaciation (Knowles, 2000, 2001; Shepard and Burbrink, 2008, 2009; McCulloch et al. 2010; Qiu et al. 2011; Liberal et al. 2014; Wallis et al. 2016). For example, endemic alpine species of butterfly (Schoville and Roderick 2009), chipmunk (Rubidge et al. 2014), and amphibian (Rovito 2010; Schoville et al. 2011) are habitat specialists of glacially deposited lake, meadow, or talus habitats; repeated glacial episodes have forced these species to retreat into refugia,

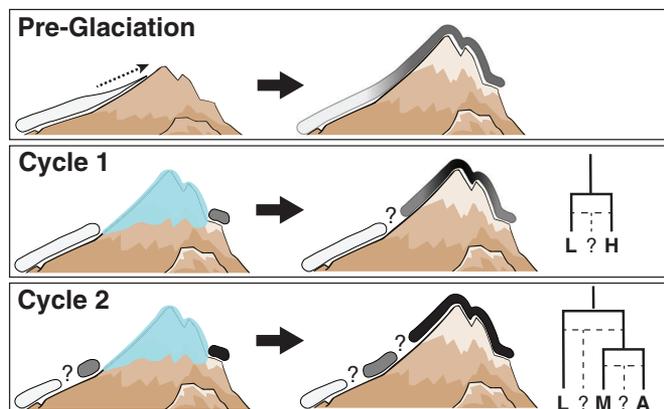


Figure 1. The glacial pulse model of alpine lineage formation. The profile of a mountain range (e.g., west to east) over time is shown. Prior to Pleistocene glacial cycles, mountains may not have reached modern stature. A hypothetical ancestral species (white) colonizes the young mountain range and eventually shows signs of ecological adaptation over an elevational gradient (indicated by grayscale cline). During the first glacial maximum (“Cycle 1”), vicariant populations experience divergent climates in isolation, due to physiographic differences caused by the mountains (e.g., rain shadow effect). Refugia on the east side may resemble higher elevation habitat than refugia on the west side, promoting further ecological and genetic divergence. Refugial populations maintain partial reproductive isolation during secondary contact. Low to moderate levels of divergence may allow a third hybrid lineage (denoted by “?”) to form. The high (H) lineage repeats the process during Cycle 2, bisecting into montane (M), alpine (A), and possibly another fused hybrid (“?”) lineage. This process may continue recursively, subdividing species into ecologically distinct lineages until ecological or genetic differences are too minor to maintain permanent barriers.

where they fractured into new lineages before recolonizing the mountains. Many alpine endemics are heterogeneously adapted to their environments along an elevational cline (Chabot and Billings 1972; Billings 1974), and time spent in glacial refugia may reinforce adaptive differences. The linear nature of glacial barriers and rain shadow effect provide ample potential for climatic differentiation, since orographic patterns of rainfall promote dramatically different refugia on opposite sides of a mountain range (Mulch et al. 2008). Thus, alpine endemics may be prone to repeated glacial “pulses” of isolation, divergent adaptation, and secondary contact (Hewitt 1996). We suggest that alpine speciation and intraspecific lineage divergence is best understood under this model, whereby pulses of glacial action and climatic differentiation bisect lineages, which can then recursively subdivide during subsequent glacial cycles (Fig. 1). This general framework explains ubiquitous patterns of species and lineage endemism in alpine zones, and provides specific testable hypotheses about the fate of new lineages where secondary contact zones emerge.

Secondary contact zones in alpine systems worldwide tend to share concordant patterns but different locations among species (Remington 1968; Petit et al. 2003; Swenson and Howard 2005; Wallis et al. 2016). This suggests that common glacial barriers may affect entire ecosystems, but species-specific traits such as dispersal ability may play a role. What is less clear is whether these zones typically maintain lineage integrity, either by lowered or augmented hybrid fitness (Fig. 1), or homogenize lineages back together by unimpeded gene flow. The outcome depends upon the extent of reproductive isolation, ecological divergence, and whether admixed populations survive into the next glacial cycle. Patchworks of multiple contact zones that reticulate cyclically across time are not uncommon, further confusing inference about population history (Hewitt, 1988, 1996, 1999, 2001, 2004, 2011). Following secondary contact, one possible outcome is that divergent lineages will produce hybrids with intrinsic genetic incompatibilities via epistasis, regardless of hybrid environment. Such a “tension zone” tends to be restricted in space, owing to a balance between migration and selection (Harrison 1993; Harrison and Larson 2014). This has been observed for two closely related lineages of newt in the Sierra Nevada, as evidenced by high inbreeding, elevated linkage disequilibrium, and heterozygote deficit at secondary contact (Kuchta 2007). Another possible outcome is for admixture to occur along an ecotone where hybrid fitness surpasses parental fitness, but only in that specific environmental context. This is known as a “bounded hybrid superiority zone,” and is especially common along alpine gradients (Abbott and Brennan 2014). Two closely related lineages of *Lycaeides* butterfly reach secondary contact at lower montane elevations in the Sierra Nevada, where a reproductively isolated hybrid lineage has developed alpine adaptations, such as high host specificity for alpine-endemic plants, a loss of egg adhesion, and novel color patterns (Gompert et al. 2006). If recombinant hybrid genotypes maintain complete isolation from parental types due to local adaptation, it is considered homoploid hybrid speciation (Buerkle et al. 2000; Abbott et al. 2013).

Between the extremes of tension zones and homoploid speciation, there is mounting evidence for a third intermediate outcome. If intraspecific hybrids experience moderate vigor and reduced outcrossing with parental lineages, they may form a semi-isolated hybrid population via lineage fusion. Although only a few instances of homoploid hybrid speciation have been described in animals (e.g., Mavárez and Linares 2008; reviewed in Schumer et al. 2014), observational bias may limit examination to pairs of distantly related species. By contrast, detailed accounts of intraspecific lineage fusion exist (e.g., Behm et al. 2010; Webb et al. 2011; Vonlanthen et al. 2012; Garrick et al. 2014; Rudman et al. 2016; Lamichhaney et al. 2018), often with evidence that lineages are temporally stable (e.g., Li et al. 2016). Regardless of whether early-generation hybrids have reduced fertility or reproductive

success, subsequent generations can still form a stable lineage (Arnold et al. 1999), and provide an important source of adaptive novelty (Hedrick 2013). Lineage fusion might be common for alpine endemics, because the temporal scale of glacial cycles is short enough that lineages may have accumulated some adaptive genetic differences without having become intrinsically incompatible. Hence, this process is particularly likely when refugial environments are similar enough to minimize postzygotic barriers (Seehausen 2006; Seehausen et al. 2008), and is one hypothesized source of lineage diversity under the glacial pulse model (Fig. 1).

The Sierra Nevada Mountains of California are an ideal system for testing the hypotheses of the glacial pulse model. A complex history of geological and climatic changes has dynamically altered California's landscape for millions of years, and driven sharp genetic discontinuities in plants and animals during the late tertiary and early quaternary periods (Calsbeek et al. 2003; Swenson and Howard 2005). California's juxtaposition of rising mountains, shifting plate tectonics, active volcanism, and resulting sharp ecotones has produced unparalleled levels of biodiversity and endemism (Myers et al. 2000; Lapointe and Rissler 2005). Over 73% of amphibian and squamate species that transect the Sierra Nevada exhibit lineage divergence in this ecoregion, although the biogeographic boundaries often do not coincide (Rissler et al. 2006). Endemism of lineages in California is especially high in patches of the western Sierra Nevada foothills and along the eastern crest, likely a consequence of glacial advance and retreat (Swenson and Howard 2005; Rissler et al. 2006). The earliest known glaciation (McGee) is thought to have occurred 2.5–1.5 million years ago during the early Pleistocene, by which time westward tilting (beginning about 5 million years ago, Pliocene) had caused the Sierra Nevada escarpment to attain sufficient height for glacial formation (Huber 1981; Unruh 1991; Gillespie and Clark 2011). The onset of Sierra Nevada glaciation matches well with major divergence times for many resident species (Calsbeek et al. 2003). Subsequent glacial cycles during the Pleistocene have played an important role in shaping biodiversity patterns in the region, by repeatedly isolating and reuniting populations (Avise et al. 1998; Hewitt 2004).

In this study, we tested for patterns of glacial pulse lineage formation in the Yosemite toad (*Anaxyrus canorus*), a Sierra Nevada endemic that is closely associated with Pleistocene glacial cycles. It breeds almost exclusively in the shallow, transient water bodies of mountain meadows (Ratliff 1985), which make up <3% of the landscape. However, currently existing meadows only formed about 10 thousand years ago (ka), as receding glaciers and rising snowpack levels helped retain more alluvium in stream valleys (Wood 1975). Hence, location of suitable habitat has likely shifted during glacial maxima. In addition, the species utilizes hydrologically distinct meadow types along an elevational gradient, from montane to subalpine (Ratliff, 1982, 1985; Weixelman

et al. 2011; Viers et al. 2013). Given its evolutionary history of dependence on glacial cycles, and adaptation along an elevational cline, the Yosemite toad is an ideal subject for studying patterns of glacial pulse lineage formation. It has even been hypothesized that the entire species is a product of adaptation to boreal conditions during mountain uplift in the Pliocene, then repeated glacial vicariance, and eventual competitive exclusion of its sister species, the Western toad (*A. boreas*; Karlstrom 1962). If glacial pulses explain the origin of the species and continued generation of novel lineages, (1) species and lineage divergence times should correspond to dates of Pleistocene glacial oscillations, (2) most sister lineages should occupy divergent ecological niches that are similar to hypothesized refugial locations, and (3) size of secondary contact zones should depend on niche distinctness of the two lineages; narrow contact zones should form between ecologically distinct lineages, and wider contact zones should form between lineages with more similar ecologies. Lineage fusion may occur anywhere that secondary admixture occurs, if hybrid genotypes have moderate vigor or advantage, and remain semi-distinct from parental lineages.

Our goals were to: (1) delineate the phylogeographic boundaries and timing of past lineage isolation for the species (most recent common ancestor of Yosemite and Kings Canyon National Parks), and more specifically within Yosemite National Park; (2) reconstruct refugial niches and locations of post-glacial recolonization, to test the hypothesis that climatically distinct refugial niches reinforced lineage divergence; and (3) resolve geographic boundaries of contact zones between lineages to test hypotheses of age, symmetry, and viability of interlineage admixture. We used a spatially dense sampling scheme in Yosemite rather than sparsely sample the entire species, because our goal was to resolve population processes at likely lineage boundaries (Shaffer and Fellers 2000; Stephens 2001; Goebel and Ranker 2009), rather than discover all lineages. We also employed a double-digest RADseq genomic sampling scheme to provide thousands of genome-wide nuclear markers that can accurately reconstruct species and population history. The glacial pulse model provides a novel framework for shedding light on patterns of alpine speciation and endemism worldwide, by hypothesizing that endemic species and their young lineages may have arisen by a common, cyclical process.

Materials and Methods

STUDY REGION

Yosemite toads (*Anaxyrus canorus*) are meadow-breeding specialists restricted to the central Sierra Nevada of California, and found between 1950 and 3444 m (Mullally and Cunningham 1956; Karlstrom 1962). The species breeds almost exclusively in shallow, snowmelt ponds within meadows, which seasonally

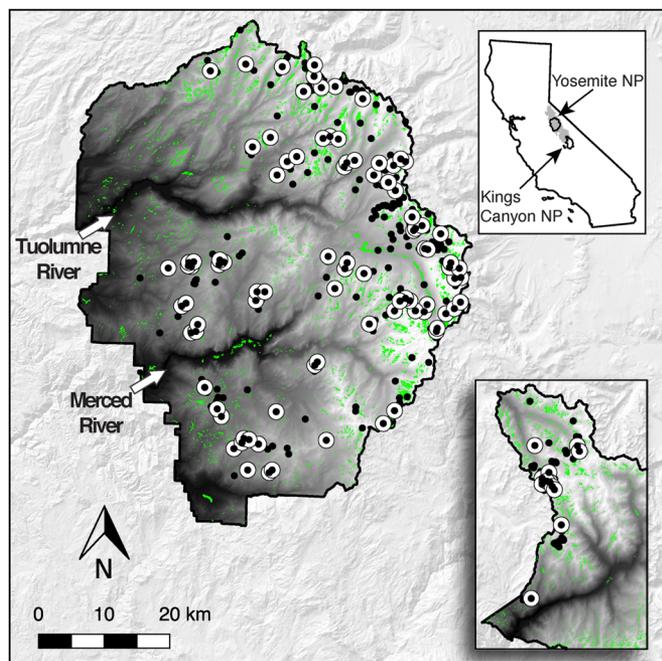


Figure 2. Study region. Primary study area in Yosemite NP (YOSE), CA includes approximately 33% of known Yosemite toad sites, whereas the outgroup area in Kings Canyon NP (KICA), CA (bottom right inset) includes the southernmost 4% of known sites. Top right inset shows the range of Yosemite toads in gray, and the boundaries of YOSE and KICA in black. Green polygons are all meadows within the parks. Solid black circles indicate all known Yosemite toad meadows identified between 1915 and the present. White circles indicate the meadows sampled and sequenced in the present study, in YOSE ($n = 90$) and KICA ($n = 12$). Eight samples of *Anaxyrus boreas* and one sample of *A. punctatus* from throughout California are not shown. Random noise is added to locations in order to protect the locations of this threatened species.

dry up. Lower elevation sites are typically spring-fed mesic or hydric meadows characterized by adjacent stands of montane red fir (*Abies magnifica*) and lodgepole pine (*Pinus contorta* var. *murrayana*), whereas subalpine and alpine meadows are typically larger, snow-fed, more xeric, and surrounded by whitebark pine (*Pinus albicaulis*) or boulders (Keeler-Wolf et al. 2012; Viers et al. 2013; Fig. 2). Yosemite National Park (YOSE; 3,027 km²) was chosen as the primary study area because it overlaps with previous studies (Shaffer and Fellers 2000; Wang 2012; Berlow et al. 2013), likely contains genetic discontinuities (Shaffer and Fellers 2000; Stephens 2001; Goebel and Ranker 2009), and is representative of the ecological and elevational conditions experienced by the species. We also sampled northern Kings Canyon National Park (KICA; 467 km²) at the southern terminus of the species range, because it likely contains the most distantly related outgroup lineage within the species.

TISSUE SAMPLING AND SITE SELECTION

Tail tissue was collected from larval toads using a sterilized razor blade during summers of 2011–2013. All samples were preserved in 95% EtOH and stored at -20°C within one week. An existing USGS meadow layer based on a vegetation map with 0.5 ha resolution was used to delimit sampling sites (Keeler-Wolf et al. 2012; Berlow et al. 2013). Sites (= meadows) were chosen to maximize representation across all known breeding locations from a recent six-year survey effort (Ostoja et al., in prep.) and overlap with previous studies. Tadpoles were sampled haphazardly across all available egg clutches, pools, and multiple years to maximize inclusion of available genetic diversity in each meadow. Although egg clutches were not directly observed, we were conscious of phenotypic indicators suggesting different clutches (e.g., average size and stage), and we sampled tadpoles from all likely clutches. A minimum of five samples was used per meadow if additional meadows were sequenced 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. Toe clips from eight *Anaxyrus boreas* samples spanning California (including samples from YOSE) and a single *A. punctatus* sample (next closest relative) were obtained as outgroups for phylogenetic analysis from the Museum of Vertebrate Zoology (University of California, Berkeley) and USGS (Western Ecological Research Center).

MOLECULAR METHODS

A total of 653 samples were chosen for sequencing (535 samples from 90 meadows in YOSE, 109 from 12 meadows in KICA, and nine outgroup samples). Genomic DNA was extracted using a combination of 96-well glass fiber plate (Ivanova et al. 2006) and DNeasy blood and tissue spin column (Qiagen) protocols. We constructed double-digest RADseq libraries following the protocol of Peterson et al. (2012; S1 Protocol), and then sequenced them using 2×100 bp reads on an Illumina HiSeq 2500. Details are described in the Supporting Information.

BIOINFORMATIC DATA PROCESSING AND SUMMARY STATISTICS

Raw data were filtered and processed using Stacks version 1.19 (Catchen et al., 2011, 2013). A detailed bioinformatic pipeline is described in the Supporting Information. Briefly, two datasets were generated from a total of 3261 loci: (1) SNPs (one SNP kept per locus), and (2) full RAD-locus haplotypes (concatenated paired-end reads were used where possible, otherwise Read 1 was used) (Table S1). Some analyses required SNP data, whereas others could accommodate the full diversity present among multiple SNPs at a locus due to intra-locus recombination. Therefore, we built and implemented a custom python script (`fasta2genotype.py`, github.com/paulmaier/fasta2genotype) to output RAD haplotypes

in addition to SNPs. After applying a locus genotype coverage threshold of 10 reads, loci were removed from meadows if absent from >25% individuals and removed from the dataset if absent from >25% individuals overall. Only alleles with a minor allele frequency (MAF) of 0.005 or greater were kept. All markers in the haplotype dataset were initially tested for random mating genotype (Hardy-Weinberg) frequencies at the meadow scale using the adegenet package (Jombart 2008) in R version 3.3.3 (R Core Team 2019). Population genetic parameters were estimated using the populations.pl script of STACKS, and N_e values were estimated using the single-sample linkage disequilibrium method in NeEstimator version 2.01 (Do et al. 2014). All processing was performed on a high-performance biocluster at the Institute for Integrative Genome Biology, UC Riverside, CA.

LINEAGE BOUNDARIES AND DIVERGENCE HISTORY

We applied four analyses to estimate phylogeographic structure. First, a spatial principal components analysis (sPCA; Jombart et al. 2008) was performed to reveal the extent and location of cryptic phylogeographic discontinuities in YOSE. This method works by creating orthogonal synthetic variables that optimize the product of genetic variance and spatial autocorrelation (measured by Moran's I). We used the haplotype dataset (2318 loci; 6473 haplotypes), and a spatial connection network of the $k = 50$ nearest neighbors for each sample location in YOSE.

Phylogenetic structure among meadows was first estimated by creating a concatenated alignment of loci, and using the GTR + Γ nucleotide model in RAxML version 8.2.9 (Stamatakis 2014) and BEAST version 2.5.2 (Drummond and Rambaut 2007). This evolutionary model was chosen using jModelTest version 2.1.4, (Posada 2008) by selecting the model with lowest AIC score and significant likelihood ratio test. In concatenated analyses, choosing one allele at random from heterozygous sites can produce incorrect and biased topologies (Degnan and Rosenberg 2009; Weisrock et al. 2012; Lischer et al. 2014; Andermann et al. 2019). Excluding these sites altogether can bias branch lengths or other parameter estimates (Sota and Vogler 2003; Garrick et al. 2010). Encoding them with IUPAC ambiguity codes can also potentially bias branch lengths if the phylogenetic program treats them as missing data (Lischer et al. 2014); however, this can be avoided by using ambiguities as equally likely values in the likelihood calculations. We chose to use meadows (not representative individuals) as operational taxon units (OTUs), to avoid these sources of phylogenetic bias, and to fully account for all mutations detected in each local population. Therefore, any SNPs that were polymorphic within OTUs were summarized by IUPAC ambiguity codes.

For the RAxML analysis, we concatenated one SNP per locus and applied the Stamatakis method of ascertainment bias correction to account for the known number of invariable sites, and

base frequencies. This ensured that bootstrap replicates sampled variable sites and provided reasonable estimates of node confidence. *A. punctatus* was used as the formal outgroup, and all eight samples of *A. boreas* were also included. A heuristic search was performed with RAxML for 10 independent runs using TBR branch swapping and 1000 bootstrap replicates for each. The tree with highest likelihood was compared with the combined bootstrap tree to check for differences in topology. In order to further assess the reciprocal monophyly of *A. boreas* and *A. canorus*, we repeated this analysis with a data matrix that balanced representation of *A. boreas* ($n = 8$) and *A. canorus* ($n = 12$; two individuals per major lineage).

We estimated divergence dates to determine if species and lineage divergence times correspond to dates of Pleistocene glacial oscillations. We used BEAST on a concatenated alignment of full sequences from all *A. canorus* individuals in the initial RAxML analysis. We used a coalescent Bayesian skyline tree model, because it had a higher harmonic mean marginal likelihood than either Yule or constant population models during preliminary runs. We used a strict clock, and calibrated the clockRate parameter with a normal prior bounded by the range of amphibian nuclear DNA rates (9.24×10^{-10} to 1.53×10^{-9} substitutions per site per year) described by Crawford (2003). Preliminary runs using a relaxed lognormal clock failed to converge, and uclsd.stdev values were close to zero, consistent with a clock-like pattern of evolution across lineages. Divergence times with and without a strict clock were nearly identical. We chose the option to utilize IUPAC ambiguity code information in tree likelihood calculations. The MCMC chain was run for 5×10^7 generations with 10^7 burnin, logging every 10^3 , and checked for convergence across two independent runs using Tracer 1.6.

Finally, we estimated the phylogeny under the multispecies coalescent (MSC) using SNAPP (Bryant et al. 2012). Unlike other MSC methods, SNAPP estimates species tree posterior probabilities directly from unlinked biallelic markers, which also avoids the need to integrate over all possible gene trees. We sampled one SNP per locus, and removed SNPs absent from one or more lineages, which resulted in a dataset with 2645 SNPs. To reduce computation time, we sampled four representative meadows per major lineage. Following Stange et al. (2018), we also fixed θ (population size) to be equal across lineages, with a uniform prior. Based upon the BEAST results, a strong Gaussian prior (mean = 1.967 million years ago, SD = 0.075 million years ago) was placed on the root age to approximate the same 95% HPD, and the clock rate was estimated. Two independent runs were performed with a chain length of 1×10^6 , sampled every 250 steps, and checked for convergence using Tracer 1.6.

For all phylogenetic analyses, meadows with strong evidence of interlineage admixture were excluded, because they would violate basic model assumptions and bias branch lengths. These

meadows were detected using (1) sPCA scores that were intermediate between lineages, (2) NewHybrids estimates of two-generation genotype classes (Anderson and Thompson 2002), and (3) conflicting signals of ancestry, based on preliminary BEAST runs (Figs. S1–S3). For (3), the posterior sample of trees was examined with DensiTree version 2.5.2, and meadows with conflicting ancestry were identified. Although (3) may denote topological uncertainty for reasons other than admixture, we considered evidence from (1), (2), and (3) to be strong evidence of interlineage admixture.

Results from the foregoing analyses suggested that three separate contact zones exist between lineages, with distinct admixed individuals found in two contact zones (see Results). Such admixed lineages may originate by several processes. For example, lineage fusion may result if hybridization occurred briefly, and produced a genetically distinct lineage that persists through time. Another possibility is that hybrids have lowered fitness, but ongoing introgression toward the contact zone replenishes a distinct admixed lineage. These possibilities assume that admixture follows secondary contact; if lineages have never been fully isolated, then primary contact (i.e., isolation with migration) may explain the pattern. Finally, the null hypothesis is that admixture has not happened, and the lineage in question may reflect simple divergence.

We tested these demographic models of admixed lineage origins using fastsimcoal version 2.6 (Excoffier et al. 2013). Fastsimcoal can leverage the observed minor allele site frequency spectrum (SFS) to calculate composite likelihoods of complex models by simulating an expected SFS with coalescent simulations. This method is particularly suitable for large SNP datasets. Multidimensional folded SFSs were generated for fastsimcoal analysis using a custom python script (github.com/isaacovercast/easySFS). For each contact zone, SFSs were downsampled to equal sample sizes among lineages while also maximizing the number of segregating sites to increase statistical power, as recommended by Gutenkunst et al. (2009). All parameters were given uniform priors except for time parameters, which were given log-uniform priors. Each model was run for 10^5 coalescent simulations with 40 expectation-maximization (ECM) cycles, and replicated for 100 independent runs, to determine parameter values leading to the maximum likelihood. Likelihood ratios or AIC statistics are imprecisely estimated using composite likelihoods; hence, we used the qpcR package (Ritz and Spiess 2008) in R to calculate Akaike's weight of evidence, based on the highest overall likelihoods across 100 independent estimations of each model.

Confidence intervals for parameters were estimated by generating 100 non-parametric bootstrap datasets. Pseudoreplicate datasets were made by sampling SNPs with replacement using the vcfR (Knaus and Grünwald 2017) package in R. Each model was

rerun 100 times, using the maximum likelihood model parameters as starting values. Parameter estimates of highest likelihood from each pseudoreplicate run were used to construct 95% confidence intervals.

REINFORCED LINEAGE DIVERGENCE FROM CLIMATICALLY DISTINCT REFUGIA

Next, we tested the hypothesis that sister lineages tend to occupy divergent climatic niches, and that glacial refugia may have reinforced their ecological divergence. Hypothesized Pleistocene refugia were reconstructed using environmental niche models (ENMs), because they provide a more spatially explicit and less subjective hypothesis of refugial distribution than coalescent approaches (Waltari et al. 2007). ENMs were constructed using Maxent version 3.4 (Phillips et al. 2004), and bioclimatic variables were obtained from WorldClim version 1.4 corresponding to present day (1960–1990) and the last glacial maximum (LGM; approximately 22 ka) (Hijmans et al. 2005). We used the Model for Interdisciplinary Research on Climate Earth System Model (MIROC-ESM), at a resolution of 2.5 arc-minutes. Sampling points for all non-admixed lineages were derived from occurrence data from 1915 to present. The number of points used was 222–1046, depending on lineage, with 25% set aside as test points. Present-day environmental niche models were constructed for each lineage separately using all 19 variables, removing those with permutation importance values of zero to reduce multicollinearity, and then rerunning the model. The final number of variables was 5–7 (Supporting Information). Each model was projected using the corresponding LGM dataset. We considered suitable refugial habitat to be locations with >0.8 probability of occurrence.

We tested the hypothesis that each lineage of toads recolonized meadows from Pleistocene refugia, in two ways. First, we used Fu's F_S (Fu 1997) to test for recent demographic expansion, using the SNP dataset ($n = 2318$ loci) in Arlequin version 3.5.2 (Excoffier and Lischer 2010). We used 10,000 coalescent simulations to assess significance. Fu's F_S estimates the probability of the observed number of alleles (k) given the number of pairwise sequence differences θ_π , with a significant ($P < 0.02$) negative value of the test statistic F_S indicating recent expansion. Next, we explicitly tested the hypothesis that a spatial range expansion had taken place from our Maxent-reconstructed refugia, using the method of Peter and Slatkin (2013) on the SNP dataset. This method infers a directionality index (ψ), based on pairwise allele frequency asymmetries that result from drift fixing lower frequency alleles during an expansion. Based on observed ψ values, the origin is inferred using time difference of arrival methods (TDOA). A key assumption of the method is even sample sizes, so five individuals were randomly chosen from every meadow. A block-jackknife approach was used to generate a null distribution (isolation by distance) to test significance of the model. Only pure

lineages were considered because admixed sample sizes were insufficient.

Climatic niche overlap of present-day lineages within YOSE was assessed using two methods. First, a linear discriminant analysis (LDA) on lineages was performed for all lineages (pure/admixed) using the six bioclimatic variables with highest loadings, with the MASS package (Venables and Ripley 2002) in R. Second, Schoener's D measuring niche overlap was calculated for each pair of non-admixed lineages, and a niche equivalency test was performed (Broennimann et al. 2012). All GIS extraction and manipulation was automated in R.

CONTACT ZONE WIDTH AND DYNAMICS

Finally, we tested the hypothesis that narrower contact zones tend to result from ecologically distinct lineages. For each of three contact zones, we estimated the geographic limits of admixture by estimating the ancestry proportion from the two parental lineages along a spatial transect, using STRUCTURE version 2.3.4 (Pritchard et al. 2000). Using the haplotype dataset ($n = 2318$ loci), we ran STRUCTURE for 10 replicates each at $K = 2$ for individuals along a spatial transect, using 10^5 steps and 10^4 burn-in. Results were combined using CLUMPP version 1.1.2 and visualized with R.

We estimated the approximate number of filial generations since admixture, along with hybrid index, to further test the stability of fused lineages. Presence of advanced hybrid genotypes is one indication of admixed population fitness, since any reproductive incompatibilities would likely curtail the extent of admixture to one or a few generations. We jointly estimated admixture index (S ; the proportion of ancestry from one of the two parental populations) and interlineage heterozygosity (H_I ; the observed heterozygosity for markers with one allele from each parental population). H_I can either be estimated as a discrete genotype class (e.g., F1), or as a continuous value from 0 to 1, where $H_I = 1$ is the maximal interlineage heterozygosity (i.e., F1 class). This latter type of analysis is more realistic in situations where hybrids are more than two generations old (Fitzpatrick 2012).

Joint estimation of S and H_I requires ancestry-informative markers that approach fixation within the two parental lineages. We chose ancestry-informative markers by subsampling the genomic data to those within the top 1% of markers contributing to spatial PC1, PC2, and PC3, which corresponded perfectly to the three clines. We chose diagnostic markers using sPCA loadings and not F_{ST} ; sPCA is model-free, whereas F_{ST} is strongly influenced by intralocus diversity, and therefore a poor predictor of marker divergence (Noor and Bennett 2009; Cruickshank and Hahn 2014). We used the R package HIest (Fitzpatrick 2012) to estimate hybrid index S and H_I with the haplotype dataset. We optimized the maximum likelihood estimate (MLE) for each individual using 1000 simulated annealing iterations, with a starting

grid size of 100. We determined approximate number of filial generations (e.g., F2, F3, F4, etc.) by contrasting empirical results with simulated values of S and H_I , assuming a population of size 50, no gene flow following admixture, and 25 ancestry-informative markers. We also identified "advanced" (>2 generations) hybrids using the HItest function in HIest, which compares the MLE of H_I and S to the MLE of a simple two-generation model, and rejects the simple model using a likelihood ratio test.

We used a principal components analysis (PCA) to assess the distinctiveness of hybrid genotypes compared to parental types. Advanced hybrids (based on S and H_I) forming distinct genetic clusters would be further evidence for lineage fusion. The prcomp package in R was applied to the haplotype dataset, after centering the data and scaling by the standard deviation.

Results

LINEAGE BOUNDARIES AND DIVERGENCE HISTORY

The current distribution of Yosemite toads within YOSE can primarily be traced back to four historic lineages. Y-North was found to be sister to the remaining YOSE lineages, with Y-South sister to a clade containing Y-East and Y-West, based on concatenated analyses (Fig. 3). Y-East was paraphyletic with respect to Y-West in the RAxML (but not the BEAST) tree (Fig. 3; Figs. S4 and S5), which likely reflects the recentness of that divergence. The SNAPP MSC tree recovered an identical topology to the concatenated trees, with Bayesian posterior probabilities of 1.0 for all clades (Fig. S6). The spatial PCA axes recapitulated the same phylogeny across geographic space (Fig. S1): the largest geo-genetic gradient was between Y-North and the remaining lineages (sPC1), followed by a gradient distinguishing Y-South (sPC2), and finally a gradient separating Y-East and Y-West (sPC3). These phylogenetic and spatial discontinuities (i.e., white areas of rapid sPC score turnover; Fig. S1) were concordant with the major glacial barriers: the main Pleistocene ice sheet, and the Merced and Tuolumne River gorges, which were also filled with glaciers (see glacial extent in Fig. 4). Additionally, we observed a pattern of increasing genetic diversity from low to high elevation (Tables S2 and S3; see Supporting Information Results).

In addition to the four primary lineages in YOSE, three distinct lineages with admixed ancestry ("East-North-A1," "East-North-A2," and "East-South-A") were detected by sPCA scores intermediate between major lineages, NewHybrids genotype classes, and conflicting signals of ancestry shown by DensiTree (Figs. S1–S3). Hence, these meadows were excluded from subsequent phylogenetic analysis. Outside of YOSE, two more lineages were found in the *A. canorus* outgroup sample of KICA ("Godard" and "Evolution").

Lineages in YOSE began to diverge from those in KICA approximately 1.96 ± 0.14 million years ago (Fig. 3; Table S4). If

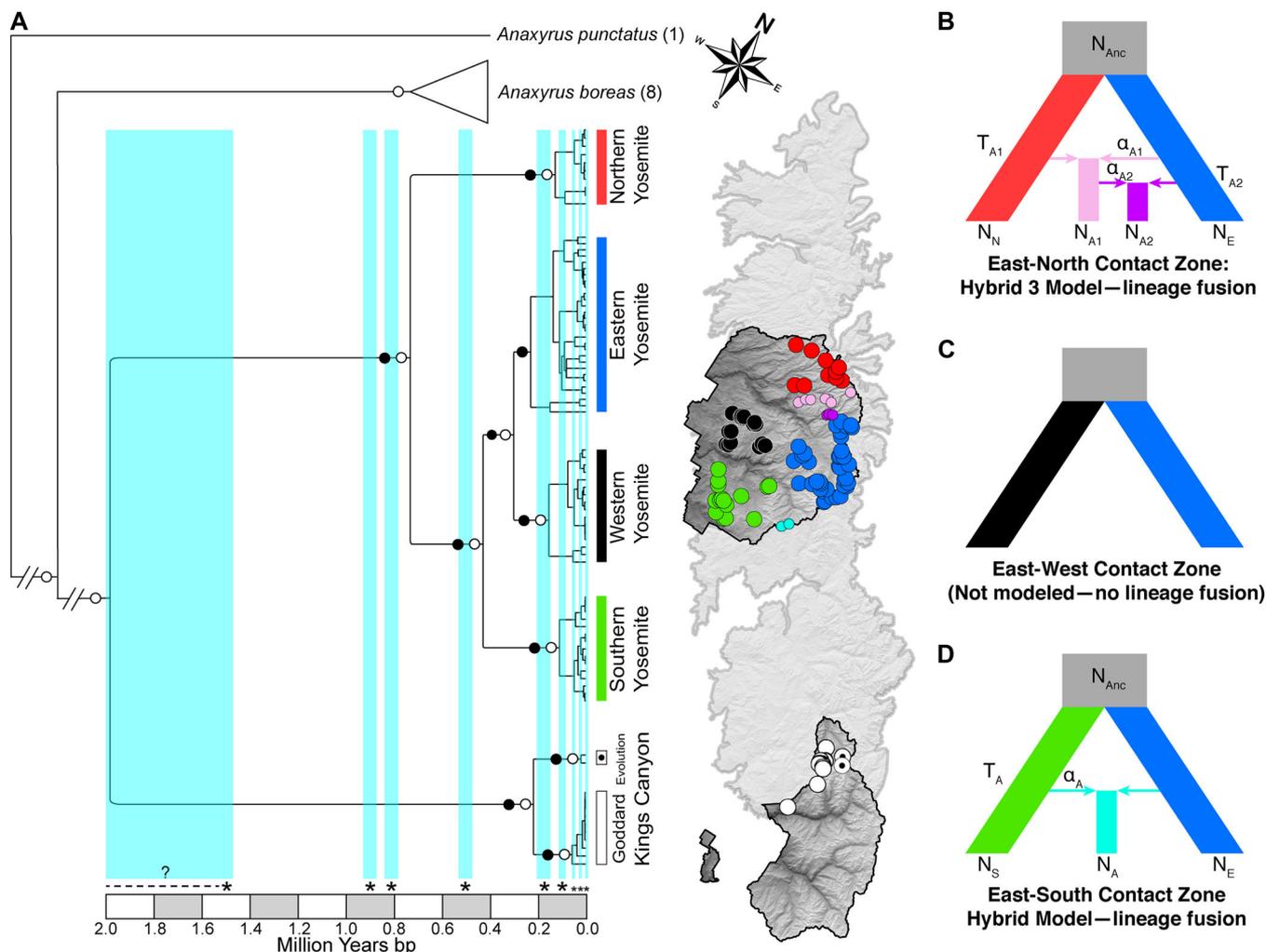


Figure 3. Lineage divergence and fusion. (A) Maximum clade credibility BEAST chronogram, with rate calibration based on a molecular clock of 9.24×10^{-10} to 1.53×10^{-9} substitutions per site per year, following Crawford (2003). A maximum likelihood phylogram from RAXML produced the same ingroup topology for major lineages, with the exception that Y-East was paraphyletic with respect to Y-West (Fig. S4). Outgroups to *Anaxyrus canorus* were only included for RAXML analysis. Black circles indicate posterior probabilities of 1.0, and open circles indicate bootstrap support values >95 for major clades. Outgroup branches for *A. boreas* and *A. punctatus* individuals are broken for clarity. Maps of both national parks are overlaid on the full distribution of *A. canorus*. Asterisks on the timeline denote ages of known glacial maxima based on Gillespie and Clark (2011), which are overlain on the tree by vertical blue stripes (see Table S4 for names and dates). (B–D) Most likely model of admixture history for each of the three contact zones. Lineage fusion was the most likely scenario in (B) and (D) based on fastsimcoal results, but no admixture was detected in (c), hence, no models were tested. See Figure S8 for a graphical representation of all 17 models tested. Parameters are defined in Table 2.

the two parks include the major lineages of the species, then the high-elevation *A. canorus* first speciated from *A. boreas* sometime during the McGee glaciation, in the early Pleistocene. The second oldest divergence between Y-North and the remainder of YOSE was estimated at 724 ± 51 ka. Only the youngest lineage (Evolution) had a crown date (18 ± 9 ka) corresponding to the most recent glaciation (Tioga; advance 21–20 ka, retreat 15–14 ka). Most divergences predated the Tioga, and had estimated dates similar to the McGee, Sherwin, Walker Creek, and Bloody Canyon glaciations (Fig. 3; Table S4). Additionally, all

eight *A. boreas* samples from both northern and southern California (including from YOSE) formed one monophyletic lineage, sister to all *A. canorus* samples. *A. canorus* and *A. boreas* were reciprocally monophyletic regardless of taxon sampling and ascertainment used (Fig. S7). The estimated clockRate was 1.507×10^{-9} substitutions per site per year, which we used in subsequent fastsimcoal analyses.

Due to the configuration of Merced and Tuolumne River canyons, three contact zones exist between the four primary YOSE lineages: East-North, East-West, and East-South (Figs. 2–4). We

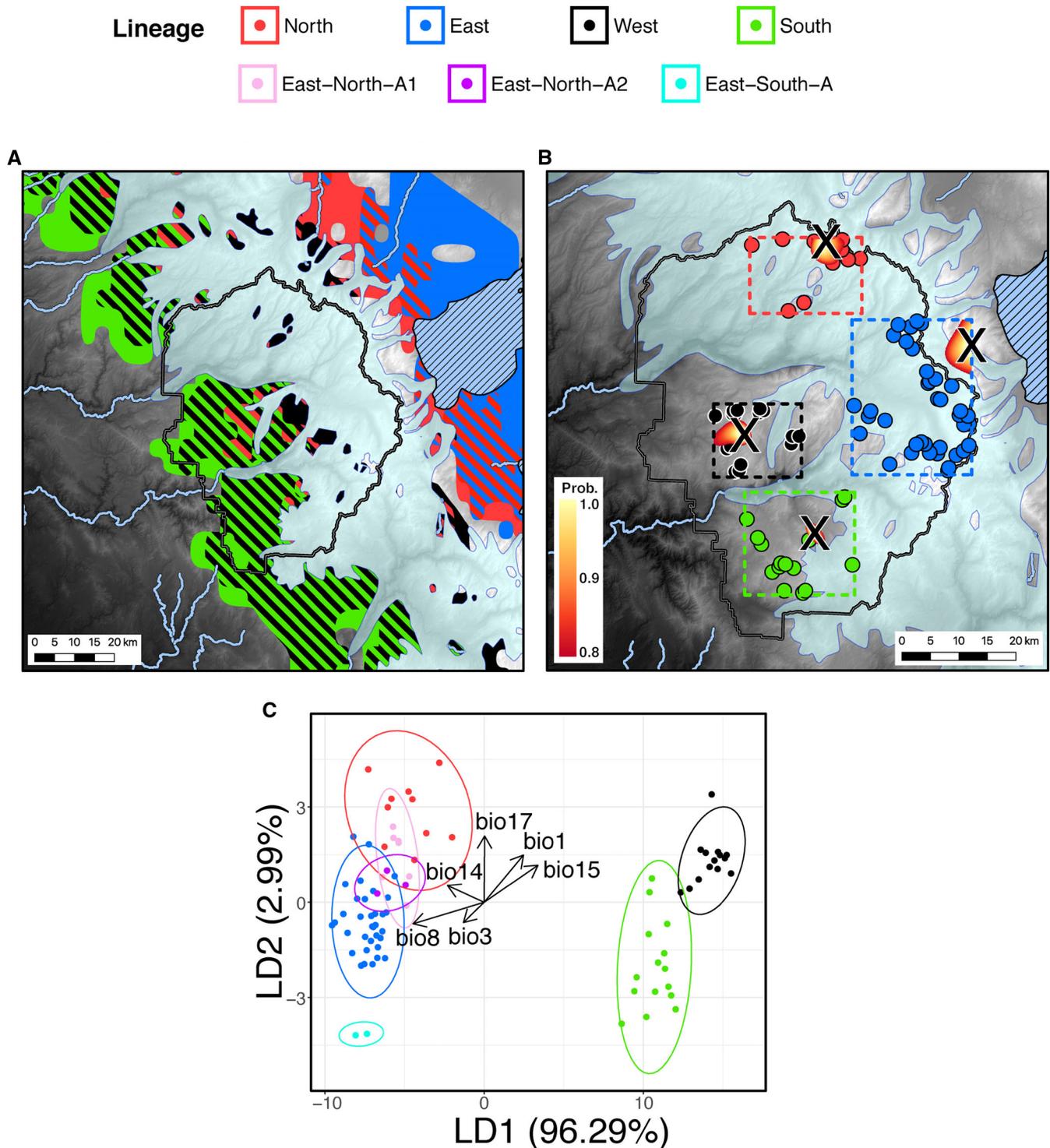


Figure 4. Climatically distinct refugia reinforced lineage divergence. (A) Potential LGM niche space for each lineage in Yosemite NP, with climatic niche suitability scores 0.8–1.0. Areas suitable for more than one lineage are indicated by diagonal hatching with multiple colors. (B) Origins of post-Pleistocene colonization events for each lineage, indicated by "X," inferred by genetic directionality index (ψ). Heatmaps underneath each "X" show its most probable location, ranging from 0.8–1.0 normalized probability. The inferred origin is limited to the area inside the dashed rectangle, so an "X" location near the rectangle boundary may suggest an origin outside of it. For (A and B), only non-admixed lineages are included due to sample size. The light blue polygon denotes the maximum extent of glaciation in the Sierra Nevada, the hatched blue polygon denotes a pluvial lake, and the blue lines denote rivers. (C) Linear discriminant biplot of six BioClim variables shows the climatic differentiation between all lineages in Yosemite NP. See Table S5 for variable definitions, and Table S6 for results of formal niche overlap and equivalency tests.

detected admixture in the East-North zone (East-North-A1 and East-North-A2 lineages) and East-South zone (East-South-A lineage). We tested a total of 17 hypothesized models for admixed lineage history in northern ($n = 10$) and southern ($n = 7$) contact zones to determine whether patterns of admixture reflect true interlineage fusion, lineage introgression, or isolation-with-migration (see Fig. S8 for all models).

The models with highest likelihood were “Hybrid 3” (northern), and “Hybrid” (southern), both with AIC weights of ≈ 1.0 (Table 1). These models inferred an instantaneous origin of admixed lineages T generations ago, after a period of time with no gene flow. The Hybrid 3 model contains two fusion events: East-North-A1 was created from Y-East and Y-North, and then East-North-A2 was created from East-North-A1 and Y-East. We used the divergence date estimates and an assumed generation time of five years to calibrate those times, and estimated admixture dates (in ka) of 473 [203–610] for East-North-A1, 254 [0.055–322] for East-North-A2, and 366 [327–377] for East-South-A (Table 2). With the exception of East-North-A2, these events predate the last three glacial episodes. The East-North-A1 lineage had higher genetic input (α) from Y-North (0.44 [0.22–0.82]), whereas East-North-A2 was much closer to Y-East (0.14 [0.14–0.86]). The East-South-A lineage had higher admixture from Y-South (0.61 [0.17–0.77]). These findings were consistent with the Hlest results (see below), but large 95% confidence intervals preclude extensive interpretation (e.g., whether backcrossing occurred).

REINFORCED LINEAGE DIVERGENCE FROM CLIMATICALLY DISTINCT REFUGIA

Pleistocene refugial models of the four major YOSE lineages showed largely east-of-ice-sheet distributions for Y-North and Y-East, and west-of-ice-sheet distributions for Y-South and Y-West (Fig. 4). The refugial variables with highest permutation importance were: annual mean temperature (Y-North), precipitation seasonality (Y-East), annual mean temperature (Y-South), and mean temperature of wettest quarter (Y-West; Table S5). Area under the curve (AUC) for test data was >0.996 for all models.

Fu's F_S estimates were consistent with demographic expansion (Table 3); all values were significantly negative ($P < 0.02$) except for the two young admixed lineages (East-North-A2, East-South-A). Estimates of pairwise ψ supported a pattern of spatial expansions in the pure lineages of YOSE, consistent with origins from the hypothesized refugia (Table 4; Figs. 4 and S9). Interestingly, TDOA-estimated origin points were east of the ice sheet for high-elevation Y-North and Y-East (whose current distributions are underneath the former ice extent), but low-elevation Y-South and Y-West had origin points inside their current distributions, west of the maximum distribution of Pleistocene ice sheets.

The LDA on six climatic variables showed that the two low-elevation lineages (Y-South/Y-West) are completely distinct from

the others across LD1 (96.29% variance), which encompasses a complex gradient of temperature and precipitation variables (Fig. 4C). The second axis explained a small minority of the variance (LD2: 2.99% variance), based upon differing summer precipitation levels between Y-North/Y-East and Y-West/Y-South. Based on Schoener's D statistics, two pairs of lineages were found to have niche overlap while accounting for total available niche space: Y-South/Y-West (46.7% overlap), and Y-North/Y-East (10.6% overlap; Table S6). A test for niche equivalency found all climatic niches to be significantly divergent except for Y-South/Y-West ($P = 0.051$), despite not being sister lineages.

CONTACT ZONE WIDTH AND DYNAMICS

Geographic extent of admixture was greatest for the high-elevation East-North contact zone, and narrower for the two contact zones of mixed elevation (Figs. 5A and S2). Based on STRUCTURE estimates, the number of meadows with $>2\%$ mean ancestry from both lineages was: 18 (East-North), 10 (East-South), and zero (East-West). The number of meadows specifically included in fused lineages was: 10 (East-North), two (East-South), and zero (East-West). The East-North contact zone spanned an approximately 20 km section of northern YOSE, from Kerrick Canyon to Virginia Canyon, whereas the East-South contact zone spanned approximately 5 km at the headwaters of the Merced River. No predominantly admixed meadows were found in the East-West contact zone.

Hlest suggested most admixed individuals are the result of more than two generations of admixture (Figs. 5B and S10). Based on simulations with $N = 50$ and no gene flow, the observed patterns are consistent with origins between five and 500 generations or more (Fig. S10), although much older admixture (as suggested by fastsimcoal; Table 2) is possible if these parameters are underestimated. In addition, the simple two-generation model was rejected in favor of advanced admixture in the majority of admixed individuals (Table S7; Fig. 5B and C).

PC1 and PC2 scores showed distinct genotyping clustering of individuals in the fused lineages, East-North-A1/A2 and East-South-A, as true-breeding recombinants would be expected to show (Fig. 5C). In addition, one meadow (2498) near Polly Dome in the East-West contact zone showed some evidence of recent (<2 generations) admixture and genotypic distinctiveness.

Discussion

Our results provide strong evidence that Yosemite toad lineages have evolved in response to repeated glacial vicariance. Lineage divergence times generally correspond to known dates of glacial maxima, spanning the Pleistocene. Our demographic and spatial demographic reconstructions support the hypothesis that toads have recently recolonized the mountains, and specifically from

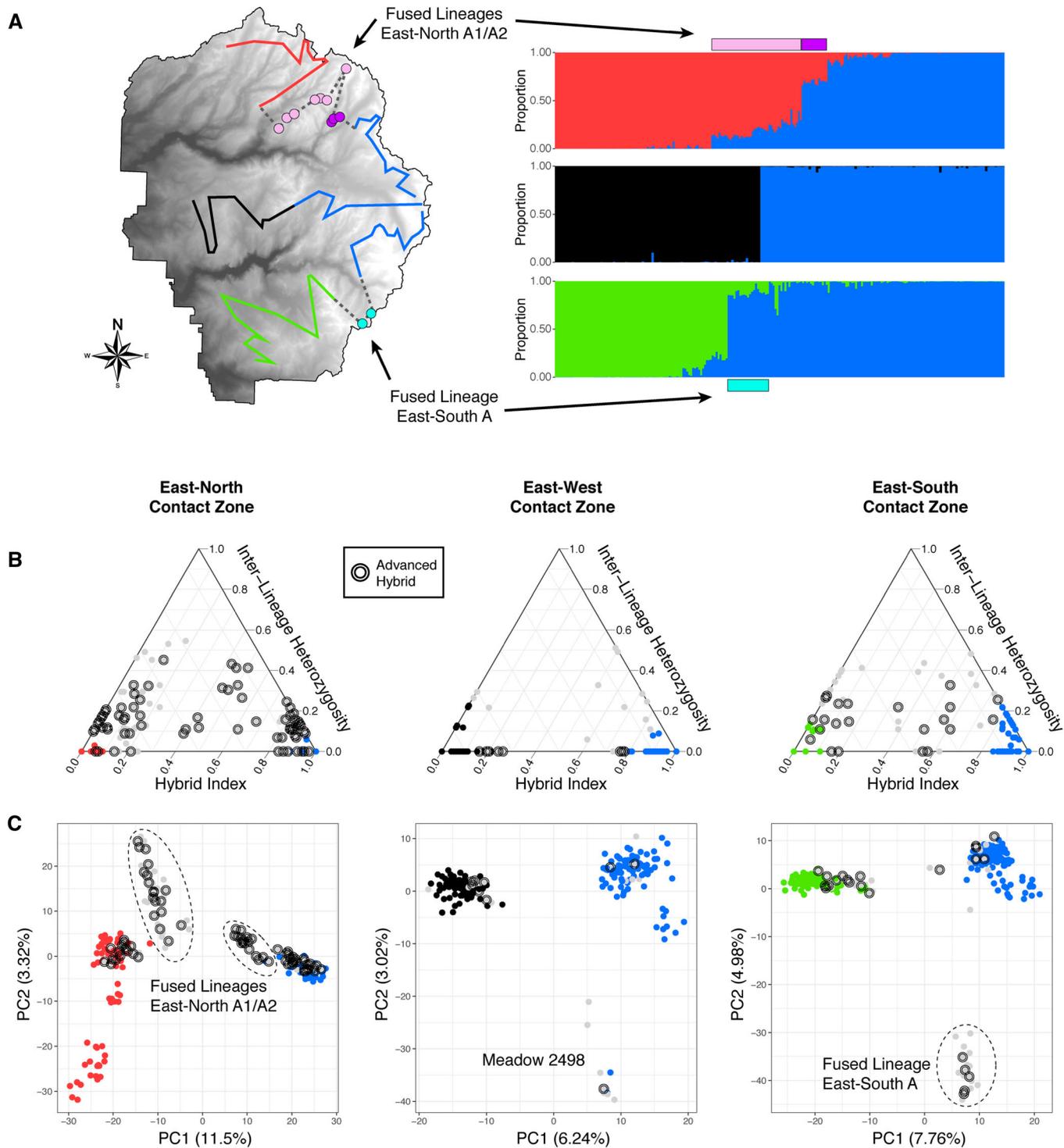


Figure 5. Contact zone width and dynamics. (A) STRUcTURE barplots ($K = 2$) and transect maps for each contact zone, showing the spatial width of each zone. Meadows from fused lineages are indicated with the same colors from Figures 3 and 4. (B) Estimation of hybrid index (S ; proportion of ancestry from each lineage) and interlineage heterozygosity (H_I ; heterozygosity at lineage-diagnostic markers) for three contact zones in YOSE, estimated by Hlest. Circles represent individual genotypes, halos represent individuals for which two-generation hybrid classes have been rejected in favor of advanced (older) hybrid classes. Colors represent pure (non-admixed) genotypes. (C) PCA biplots of pure and admixed individuals, showing separate clustering of genotypes in fused lineages.

Table 1. Results of model testing using fastsimcoal on two admixture zones in Yosemite NP, East-North, and East-South. "Isolation" = simple isolation/divergence of lineages; "Iso./Migration" = isolation with migration; "Hybrid" = lineage fusion by hybridization; "Intro." = ongoing introgression into an admixed lineage. A graphical depiction of models and model parameters is shown in Figures 3 and S8. Each model was run independently 100 times using 100,000 coalescent simulations 40 expectation-maximization cycles each time. The most likely model is shown in bold.

| Zone | Model | Maximum AIC | No. Param. | Δ AIC | AIC Weight |
|-------------------|------------------|------------------|------------|--------------|-------------------------|
| East-North Admix. | Isolation | -3771.12 | 7 | -51.744 | 2.49×10^{-22} |
| | Iso./Migration | -3904.74 | 15 | -185.37 | 7.75×10^{-84} |
| | Hybrid 1 | -3729.53 | 9 | -10.157 | 3.88×10^{-5} |
| | Hybrid 2 | -3748.28 | 9 | -28.906 | 2.79×10^{-13} |
| | Hybrid 3 | -3719.37 | 9 | 0 | ~1.00 |
| | Hybrid 4 | -3733.56 | 8 | -14.185 | 1.88×10^{-6} |
| | Intro. 1 | -3854.50 | 9 | -135.131 | 2.06×10^{-59} |
| | Intro. 2 | -3792.58 | 9 | -73.206 | 1.61×10^{-32} |
| | Intro. 3 | -3846.50 | 9 | -127.124 | 6.18×10^{-56} |
| | Intro. 4 | -3767.68 | 8 | -48.302 | 2.86×10^{-21} |
| East-South Admix. | Isolation 1 | -4073.765 | 5 | -23.481 | 1.72×10^{-10} |
| | Isolation 2 | -4064.046 | 5 | -13.762 | 2.87×10^{-6} |
| | Iso./Migration 1 | -5784.45 | 11 | -1734.16 | ~0.00 |
| | Iso./Migration 2 | -5700.31 | 11 | -1650.03 | ~0.00 |
| | Hybrid | -4050.284 | 6 | 0 | ~1.00 |
| | Intro. 1 | -4182.295 | 6 | -132.011 | 4.66×10^{-58} |
| | Intro. 2 | -4704.026 | 6 | -653.742 | 1.21×10^{-284} |

glacial refugia. Lower-elevation toads have repeatedly been isolated in situ west of Pleistocene ice sheets, while their higher-elevation brethren have been forced eastward into Owens Valley, only to recolonize later. This has resulted in a phylogenetic pattern of sister lineages residing in divergent climatic niches. Niche evolution may have reinforced lineage incompatibility at secondary contact zones. If drift alone was responsible for levels of reproductive isolation, then lineage incompatibility would increase with time. However, the most extensive admixture observed was between Y-East and Y-North, the contact zone with the oldest divergence, yet highest niche overlap. Conversely, the two narrowest contact zones are between lineages with least niche overlap. Furthermore, at least two contact zones have apparently generated novel lineages through the process of lineage fusion, the extent of which also correlates with niche overlap. These results are consistent with the model of glacial pulses (Fig. 1), which may apply to other alpine endemics.

ROLE OF NICHE DIVERGENCE IN THE GLACIAL PULSE MODEL

Recently, there has been a debate about the role of niche evolution in speciation. Proponents of niche conservatism argue that species innately have low tolerance to adapt outside their fundamental niche, meaning that climate shifts can fragment the distribution of that niche, leading to allopatric speciation (Wiens 2004; Wiens and Graham 2005; Wiens et al. 2010). Under this model, ecolog-

ical divergence is the result (not the cause) of lineage formation. However, this process does not convincingly explain niche divergence for adjacent species pairs occupying steep mountainous gradients. These are commonly observed in animals (e.g., Graham et al. 2004) and plants (e.g., Chabot and Billings 1972). In contrast, proponents of ecological speciation argue that natural selection can overpower gene flow to produce local adaptation, and hence, niche divergence directly spawns new lineages (Dobzhansky 1951; Schluter 2001; Rundle and Nosil 2005). Although there are numerous examples of divergent selection driving incipient ecological speciation, advocates readily admit that the process is often incomplete (Nosil et al. 2009).

The glacial pulse model combines elements of niche divergence with the classical and widely accepted model of allopatric speciation and associated reproductive incompatibilities (Mayr 1942; Coyne and Orr 1989; Barraclough and Vogler 2000). It unifies the two principal explanations for speciation in Sierra Nevada alpine endemics: (1) colonization of mountains followed by physiological or ecological divergence (Went 1948; Chabot and Billings 1972; Billings 1974), and (2) isolation of formerly continuous species in glacial or interglacial refugia (Rovito 2010; Rubidge et al. 2014). Our results suggest these both happen, cyclically, and they are mutually reinforcing (Fig. 1). In this study, we found evidence that Yosemite toads have repeatedly become allopatric, and have diverged in climatic niche while isolated into distinct (west vs. east) refugia. The fact that high- and

Table 2. Demographic parameter estimates for models with highest weighted AIC scores from the East-North and East-South admixture zones. See Figure 3 for graphical depictions of models and parameters. Estimates are the highest maximum likelihood scores from 100 replicate fastsimcoal runs.

| Model | Parameter [†] | | | | | | | | | |
|----------------------|------------------------|---|---|---|---|---|---|--------------------|---|--|
| | N_{Anc}^{\ddagger} | N_N | N_{A1} | N_{A2} | N_E | α_{A1} | α_{A2} | T_{A1}^{\S} | T_{A2} | |
| East-North Admix. | 1.17×10^7 | 9.65×10^5 | 7.11×10^5 | 4.15×10^5 | 7.17×10^5 | 0.44 | 0.14 | 4.73×10^5 | 2.54×10^5 | |
| | 95% CI | $[1.17 \times 10^7 - 1.77 \times 10^7]$ | $[4.55 \times 10^5 - 9.75 \times 10^5]$ | $[8.29 \times 10^5 - 7.11 \times 10^5]$ | $[0.95 \times 10^5 - 4.29 \times 10^5]$ | $[6.34 \times 10^5 - 8.10 \times 10^5]$ | $[0.22 - 0.82]$ | $[0.14 - 0.86]$ | $[2.03 \times 10^5 - 6.10 \times 10^5]$ | |
| East-South Admix. | 1.02×10^7 | 3.75×10^5 | 3.34×10^5 | 6.12×10^5 | 0.61 | 3.66×10^5 | | | | |
| | 95% CI | $[9.53 \times 10^6 - 1.21 \times 10^7]$ | $[1.97 \times 10^5 - 4.05 \times 10^5]$ | $[2.74 \times 10^5 - 5.50 \times 10^5]$ | $[3.99 \times 10^5 - 6.92 \times 10^5]$ | $[0.17 - 0.77]$ | $[3.27 \times 10^5 - 3.77 \times 10^5]$ | | | |

[†]Parameters: population sizes for North (N_N), East-North-A1 (N_{A1}), East-North-A2 (N_{A2}), East (N_E), South (N_S), and ancestral (N_{Anc}) lineages; admixture proportions from East \rightarrow East-North-A1 (α_{A1}), East-North-A2 (α_{A2}), and South \rightarrow East-North-A1 (α_{A1}), East-North-A2 (α_{A2}), and East-South-A (T_A) lineages.

[‡]Population sizes are in haploid units.

[§]Admixture times are in units of years, with an assumed generation time of five years.

^{||}The 95% confidence intervals were estimated using 100 pseudoreplicate SFS datasets and rerunning each model 100 times, using the best model parameters as starting values.

Table 3. Results of Fu's F_S test of demographic expansion in Yosemite NP. Large negative values of Fu's F_S indicate recent demographic expansion, with significance level considered to be 0.02.

| Lineage | Number of alleles | Expected number of alleles | θ_π | Fu's F_S | Fu's F_S P -value |
|---------------|-------------------|----------------------------|--------------|------------|-----------------------|
| North | 150.00 | 93.54 | 105.26 | -24.85 | *** |
| East-North-A1 | 116.00 | 78.78 | 106.76 | -25.12 | *** |
| East-North-A2 | 28.00 | 24.93 | 105.02 | -3.28 | 0.05 |
| East | 388.00 | 141.29 | 79.53 | -23.97 | *** |
| East-South-A | 42.00 | 37.07 | 147.45 | -5.35 | 0.03 |
| South | 176.00 | 96.73 | 87.34 | -24.50 | *** |
| West | 170.00 | 102.20 | 107.34 | -24.82 | *** |

*** Significance of demographic expansion at the 0.001 probability level.

Table 4. Inferred origins of post-Pleistocene colonization events in the four pure lineages of Yosemite NP. Admixed lineages were excluded due to low sample sizes.

| Lineage | Longitude [†] | Latitude [†] | q^\ddagger | r_1^\S | r_{10}^\S | r_{100}^\S | d | R^2 [¶] | p -value |
|---------|------------------------|-----------------------|--------------|----------|-------------|--------------|-------|--------------------|------------|
| North | -119.5173 | 38.11973 | 0.015 | 0.971 | 0.772 | 0.253 | 0.343 | 0.600 | *** |
| East | -119.2167 | 37.96289 | 0.008 | 0.985 | 0.869 | 0.398 | 0.669 | 0.273 | *** |
| South | -119.5223 | 37.66362 | 0.012 | 0.976 | 0.804 | 0.291 | 0.414 | 0.237 | *** |
| West | -119.6657 | 37.8162 | 0.020 | 0.961 | 0.710 | 0.196 | 0.247 | 0.540 | *** |

[†]Inferred origin of spatial range expansion.

[‡]Strength of colonization, given as regression slope in km^{-1} .

[§]Decrease in diversity over 1, 10, and 100 km given as N_e^{founder}/N_e .

[¶]Adjusted coefficient of determination for the nonlinear model.

*** Significance of spatial range expansion at the 0.001 probability level.

low-elevation niches have evolved more than once (i.e., convergence between Y-South and Y-West) suggests a strong influence of climatic niche selection. The observation that ecological divergence—not genetic divergence—is most correlated with narrow contact zones suggests that niche divergence may accelerate reproductive isolation. Although the Sierra Nevada are unique among mountain ecosystems in having a Mediterranean climate (Rundel 2011; Rundel and Millar 2016), the patterns found here may be widespread. For example, similar patterns (i.e., repeated ecological selection, glacial allopatry, and recolonization) have been observed in European grasshoppers (Hewitt 1996).

One assumption of the model is heterogeneity of refugial climates, because post-glacial recolonization must proceed into non-overlapping niches. High-Sierra and eastern desert (Owen's Valley) ecosystems have similar growing seasons (Went 1948), which may explain why the latter was a suitable refugium for high-elevation and not low-elevation Yosemite toad lineages. Climatic conditions in Owen's Valley over the past 155,000 years were also wetter and cooler than at present (Menking et al. 1997; Koehler et al. 2005), and might have supported large population sizes. This might explain the much higher genetic diversity

of high-elevation lineages. Another Sierra Nevada endemic amphibian (*Hydromantes platycephalus*) reaches its highest diversity in its eastern lineage, near Owen's Valley, which seems to corroborate this pattern (Rovito 2010). Diversity patterns may also reflect current habitat: although high-elevation lineages must periodically retreat into refugia, their meadows are rejuvenated with new alluvial deposits that may positively influence hydrology for tadpoles (Wood 1975). Conversely, low-elevation lineages that remain in situ may experience little benefit from glacial action, and higher rates of conifer encroachment into meadows as trees recolonize middle elevations (Woolfenden 1996; Lubetkin et al. 2017).

PHYLOGEOGRAPHIC BOUNDARIES AND DATES OF PAST LINEAGE ISOLATION

We detected substantial phylogeographic structure within YOSE and between the two parks. Previous studies have found two major mitochondrial clades of *A. canorus*, often polyphyletic with respect to *A. boreas*, with alternate samples from Ireland Lake (in Y-East) falling into either clade (Shaffer and Fellers 2000; Stephens 2001; Goebel and Ranker 2009). Although their spatial

sampling was limited, our East-North contact zone is likely the boundary of those previously described clades. The increased resolution of our sampling revealed that two other contact zones exist within the park, and implicated both the longitudinal and transverse ice sheets of the Pleistocene as likely barriers. The Merced and Tuolumne River gorges (Fig. 4) contained formidable ice sheets, and structured lineage boundaries in other Sierra amphibians such as *Rana sierrae* (Poorten et al. 2017) and *Hydromantes* salamanders (Rovito 2010). These barriers provided two disconnected low elevation refugia, allowing multiple glacial pulses to churn out low elevation lineages (Y-West and Y-South).

Using a published molecular clock for anuran non-synonymous nuclear DNA, we estimated that the time to most recent common ancestor (TMRCA) for the species—based on the divergence date for the two parks—was 1.96 ± 0.14 million years ago, corresponding to the approximate onset of Sierra Nevada glaciation (Calsbeek et al. 2003). The median estimates for all subsequent divergences ranged between 724 and 215 ka. Beginning in the late Pleistocene (about 900 ka), glacial periodicity switched from 41 to 100 ka cycles, with longer duration, shorter inter-glacials, and colder average temperature (Mudelsee and Schulz 1997; Tziperman and Gildor 2003). Hence, intra-park divergences could reflect stronger isolating events during this period. Numerous other arctic and alpine temperate species exhibit deep divergences dating back to the mid-Pleistocene, indicating that they retained distinctiveness over several ice ages, and that lineage diversification proceeds by repeated allopatry (i.e. “glacial pulse” mechanism) (Hewitt, 1996, 2004; Avise et al. 1998). All of our divergence estimates (1.96 million years ago to 215 ka) are much older than the five most recent glaciation events (Gillespie and Clark 2011), which is consistent with remaining distinct during interglacial time. This suggests these lineages have either survived through glacial maxima in high-elevation nunatak or peripheral refugia (Holderegger and Thiel-Egenter 2009), or remained reproductively isolated while in lowland refugia.

Although theory predicts that cold-adapted species should expand their ranges during ice ages (Haffer 1969), montane and alpine amphibians may counter this pattern by remaining in refugia within their current distributions. For example, phylogeographic and isolation by distance analyses showed that pygmy salamanders (*Desmognathus wrighti*) remained fragmented at high elevation during the Pleistocene, and this restriction was likely driven by ecological interactions (Crespi et al. 2003). Interestingly, another Sierra Nevada endemic amphibian (the sympatric *Rana sierrae*) has three gene pools with almost identical spatial extent to *A. canorus* lineages in YOSE (Poorten et al. 2017). If divergence times turn out to be similar, this would be strong support that common glacial barriers and refugia have structured each species. For our foregoing conclusions, we note the caveat that using a rate calibration from other species assumes

similar rates of evolution. This assumption may be violated; for example, a relatively low historical N_e could accelerate fixation, biasing divergence estimates upward. With increasing genome-wide studies of phylogeography, genomic rates of evolution will be better assessed and calibrated across diverse taxa.

GENETIC CONSEQUENCES OF INTERLINEAGE ADMIXTURE

Our analyses suggest that secondary contact following isolation has fused together new and persistent lineages of admixed origin. Admixture tended to be asymmetrical, suggesting that backcrossing may be important during this process, although the confidence intervals of fastsimcoal estimates were too broad to corroborate this pattern. Lineage fusion presents an intriguing research opportunity for admixed fitness and conservation. Traditionally, hybridization is discussed in the context of interspecific gene flow or introgression between native and exotic species. In these cases of anthropogenic hybridization, two negative consequences are likely: outbreeding depression follows from maladaptive allele combinations and results in low hybrid vigor or fertility, or genetic assimilation erodes locally adapted populations and species boundaries (Ellstrand and Elam 1993; Weber and D’Antonio 2000). For example, introduction of an exotic ambystomatid salamander into southern California has sent a rapid influx of morphological and genetic changes across populations of *Ambystoma californiense*, where hybrids have outcompeted natives of that and other species (Ryan et al. 2009; Fitzpatrick et al. 2010). However, a more recent view of hybridization posits it is more ubiquitous and beneficial than previously assumed (Arnold et al. 1999; Mallet 2005). Many recent examples of lineage fusion have been described, due to anthropogenic (Hendry et al. 2006; Behm et al. 2010; Vonlanthen et al. 2012; Rudman et al. 2016) or natural (Vergilino et al. 2011; Webb et al. 2011; Cui et al. 2013; Talavera et al. 2013; Garrick et al. 2014; Grant and Grant 2014; Lamichhaney et al. 2018) breakdown of postmating barriers, with evidence of temporal stability in several cases (e.g., Li et al. 2016). Intraspecific lineage fusion is particularly likely when environmental heterogeneity between lineages is minimal (Seehausen 2006; Seehausen et al. 2008).

Periodic lineage fusion could be an essential source of adaptive genetic novelty and hybrid vigor in species such as the Yosemite toad, given its small average N_e of 29.13 (Table S3; consistent with Wang 2012). Preliminary analysis of Yosemite toad evolutionary rates has shown a strong negative relationship with estimated census population size, indicating genetic drift might overpower adaptation from standing variation and new mutations (Maier et al. 2016; Maier 2018). Although recombination in admixed individuals can disrupt co-adapted genotypes underlying polygenic traits, it has the advantage of generating novel adaptive alleles at an accelerated rate (Hedrick 2013). Since

admixture is a recurrent process, the continual flow of adaptive alleles can raise the initial frequency of them considerably, and lead to much faster fixation (Hedrick 2013). This process, in concert with possible heterosis experienced by F1s and other early filial individuals, can dramatically expedite adaptive evolution. For example, hybridization between several closely related species of Darwin's finches on Daphne Island has contributed at least 10% additive genetic variance in beak shape, a trait under intense and episodic selection (Grant and Grant, 2008, 2010, 2014; Grant 2015). Subsequent fluctuations in food availability during an El Niño event allowed hybrids to survive longer, and backcrossed individuals eventually formed a morphologically, ecologically, and reproductively distinct lineage.

In the present study, we detected three lineage fusions of moderate age, whose descendants appear to maintain independent evolutionary trajectories. With ongoing admixture zones involving independently evolving lineages, more extreme traits are possible, including very rapid or slow tadpole growth rates (Parris 1999; Reyer 2008). Under conditions where ponds are shallow and tend to desiccate, admixture in pond-breeding anurans is more widespread (Kingsolver et al. 2002; Pfennig 2007), and natural selection can more efficiently replace old traits with new, successful ones because they are already at high frequencies (Hedrick 2013). There are many examples of intraspecific admixture increasing the viability of translocated larvae and adults (e.g., Japanese common toads: Hase et al. 2013; European toads: Zeisset and Beebe 2013). Given the extremely short and variable hydroperiods Yosemite toads experience during larval development, it seems likely that tadpole physiology would be under strong selection. Thus, any heritable Yosemite toad fitness traits that have diverged between lineages may be studied for conservation value, to ascertain whether (and in what environmental context) pure versus admixed individuals would make superior stock for translocation efforts (Hamilton and Miller 2016). Although the longevity and number of lineage fusions hints at their population fitness, our evidence does not directly suggest hybrid advantage, a subject that future studies should examine at the phenotypic level.

Conclusions

In this study, we have provided evidence that pulses of glacial action bisect lineages into divergent climatic niches, which then can form further lineages by fusion. This mechanism is unique from other described mechanisms for lineage diversification, and should be considered in phylogeographic hypothesis testing for other alpine systems. The process of lineage fusion in secondary contact zones is particularly intriguing. Given that Yosemite toads are federally threatened and particularly susceptible to climatic fluctuations (USFWS 2014), it may be beneficial to elucidate

any adaptive or deleterious effects of interlineage admixture before attempting conservation measures such as translocations. We suggest further research into the possibility that such a zone is a crucible for valuable adaptive diversity to combat stressors such as disease and periodic climatic extremes, such as intense drought. All nine lineages in our study area may be considered potential evolutionarily significant units (ESUs; sensu Moritz 1994), since they appear to be evolving independently, and occupy distinct climatic niches in most cases. The future for alpine endemics is grim; one study estimated that up to 66% of California endemic plants will experience large (>80%) range reductions by year 2100 (Loarie et al. 2008). However, our results are promising in that lineages have evolved differing climatic optima, and hence should have a diversified response to climate change. Future studies may expand upon our results by examining the fitness differences between pure/admixed, and low-/high-elevation lineages, to anticipate future changes.

AUTHOR CONTRIBUTIONS

P.A.M., S.M.O., A.A., A.G.V., and A.J.B. designed the study; P.A.M. and S.M.O. collected tissue samples; P.A.M. performed the laboratory work and conducted the analyses, with guidance from A.J.B., A.G.V., and A.A. P.A.M. wrote the manuscript, with input from A.G.V., A.J.B., S.M.O., and A.A.

ACKNOWLEDGMENTS

We thank Michael Hernandez, Johanne Boulat, Alexa Killion, Mara MacKinnon, Alexa Lindauer, Ross Maynard, Nathalie Aall, Corrina Kamoroff, Steven Lee, Amy Patton, 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. 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, Leonard Nunney, Rachel Shoop, and three anonymous reviewers 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 Ostoja, 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, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

DATA ARCHIVING

Demultiplexed fastq files of double-digest RADseq data are deposited at NCBI GenBank SRA under BioProject PRJNA558546 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA558546>). The python program `fasta2genotype.py`, user manual, and example dataset are available for download online: <https://github.com/paulmaier/fasta2genotype>. R scripts and genotype/GIS data necessary for reproducing phylogenetic, demographic, niche overlap, and hybrid ancestry analyses are available at Dryad (<https://doi.org/10.5061/dryad.xsj3tx99h>). Information about all

samples used in this study, in addition to supplemental tables, figures, and methods/results, are available in the online Supporting Information.

JOURNAL CLUB SLIDES

Journal club slides for presenting the article have been posted on FigShare (<https://www.doi.org/10.6084/m9.figshare.10260362>).

LITERATURE CITED

- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, et al. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229–246.
- Abbott, R. J., and A. C. Brennan. 2014. Altitudinal gradients, plant hybrid zones and evolutionary novelty. *Philos. Trans. R. Soc. B Biol. Sci.* 369:1–12.
- Andermann, T., A. M. Fernandes, U. Olsson, M. Töpel, B. Pfeil, B. Oxelman, A. Aleixo, B. C. Faircloth, and A. Antonelli. 2019. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Syst. Biol.* 68:32–46.
- Anderson, E. 1948. Hybridization of the habitat. *Evolution* 2:1–9.
- . 1953. Introgressive hybridization. *Biol. Rev.* 28:280–307.
- Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160:1217–1229.
- Arnold, M. L., M. R. Bulger, J. M. Burke, and A. L. Hempel. 1999. Natural hybridization: how low can you go and still be important? *Ecology* 80:371–281.
- Avise, J. C., D. Walker, and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. B Biol. Sci.* 265:1707–1712.
- Barracough, T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* 155:419–434.
- Behm, J. E., A. R. Ives, J. W. Boughman, J. E. Behm, A. R. Ives, and J. W. Boughman. 2010. Breakdown in postmating isolation and the collapse of a species pair through hybridization. *Am. Nat.* 175:11–26.
- Berlow, E. L., R. Knapp, S. M. Ostoja, R. J. Williams, H. McKenny, J. R. Matchett, Q. Guo, G. M. Fellers, P. Kleeman, M. L. Brooks, 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.
- Bernatchez, L., and C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Mol. Ecol.* 7:431–452.
- Billings, W. D. 1974. Adaptations and origins of alpine plants. *Arct. Alp. Res.* 6:129–142.
- Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G. Yoccoz, W. Thuiller, M. J. Fortin, C. Randin, N. E. Zimmermann, et al. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob. Ecol. Biogeogr.* 21:481–497.
- Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. Roychoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* 29:1917–1932.
- Buerkle, C. A., R. J. Morris, M. A. Asmussen, and L. H. Rieseberg. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84:441–451.
- Calsbeek, R., J. N. Thompson, and J. E. Richardson. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Mol. Ecol.* 12:1021–1029.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22:3124–40.
- Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3 Genes, Genomes, Genet.* 1:171–82.
- Chabot, B. F., and W. D. Billings. 1972. Origins and ecology of the Sierran alpine flora and vegetation. *Ecol. Monogr.* 42:163–199.
- Coyne, J., and H. Orr. 1989. Patterns of speciation in *Drosophila*. *Evolution* 43:362–381.
- Crawford, A. J. 2003. Relative rates of nucleotide substitution in frogs. *J. Mol. Evol.* 57:636–641.
- Crespi, E. J., L. J. Rissler, and R. A. Browne. 2003. Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Mol. Ecol.* 12:969–984.
- Cruikshank, T. E., and M. W. Hahn. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23:3133–3157.
- Cui, R., M. Schumer, K. Kruesi, R. Walter, P. Andolfatto, and G. G. Rosenthal. 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution* 67:2166–2179.
- Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24:332–340.
- Do, C., R. S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett, and J. R. Ovenden. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* 14:209–214.
- Dobzhansky, T. 1951. *Genetics and the origin of species*. 3rd ed. Columbia Univ. Press, New York, NY.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:1–8.
- Ellstrand, N. C., and D. R. Elam. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annu. Rev. Ecol. Syst.* 24:217–242.
- Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet.* 9:e1003905.
- Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–567.
- Fitzpatrick, B. M. 2012. Estimating ancestry and heterozygosity of hybrids using molecular markers. *BMC Evol. Biol.* 12:131.
- Fitzpatrick, B. M., J. R. Johnson, D. K. Kump, J. J. Smith, S. R. Voss, and H. B. Shaffer. 2010. Rapid spread of invasive genes into a threatened native species. *Proc. Natl. Acad. Sci. U. S. A.* 107:3606–3610.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Garrick R. C., P. Sunnucks, and R. J. Dyer. 2010. Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evol. Biol.* 10:118.
- Garrick, R. C., E. Benavides, M. A. Russello, and C. Hyseni. 2014. Lineage fusion in Galápagos giant tortoises. *Mol. Ecol.* 23:5276–5290.
- Gillespie, A. R., and D. H. Clark. 2011. Glaciations of the Sierra Nevada, California, USA. Pp. 447–462 in J. Ehlers, P. L. Gibbard, and P. D. Hughes, eds. *Developments in quaternary science*. Elsevier, Amsterdam, the Netherlands.
- Goebel, A., and T. Ranker. 2009. Mitochondrial DNA evolution in the *Anaxyrus boreas* species group. *Mol. Phylogenet. Evol.* 50:209–225.
- Gompert, Z., A. M. Shapiro, and C. C. Nice. 2006. Homoploid hybrid speciation in an extreme habitat. *Science* 314:1923–1925.

- Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58:1781–1793.
- Grant, B. R. 2015. Introgressive hybridization and natural selection in Darwin's finches. *Biol. J. Linn. Soc.* 117:812–822.
- Grant, P. R., and B. R. Grant. 2010. Conspecific versus heterospecific gene exchange between populations of Darwin's finches. *Philos. Trans. Biol. Sci.* 365:1065–1076.
- . 2014. Synergism of natural selection and introgression in the origin of a new species. *Am. Nat.* 183:671–681.
- Grant, P. R., and R. Grant. 2008. How and why species multiply: the radiation of Darwin's finches. Princeton Univ. Press, Princeton, NJ.
- Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5:e1000695.
- Haffer, J. 1969. Speciation in Amazonian forest birds. *Science* 165:131–137.
- Hamilton, J. A., and J. M. Miller. 2016. Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conserv. Biol.* 30:33–41.
- Harrison, R. G. (ed). 1993. Hybrid zones and the evolutionary process. Oxford Univ. Press, Ithaca, NY.
- Harrison, R. G., and E. L. Larson. 2014. Hybridization, introgression, and the nature of species boundaries. *J. Hered.* 105:795–809.
- Hase, K., N. Nikoh, and M. Shimada. 2013. Population admixture and high larval viability among urban toads. *Ecol. Evol.* 3:1677–1691.
- Hedrick, P. W. 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* 22:4606–4618.
- Hendry, A. P., P. R. Grant, B. R. Grant, H. A. Ford, M. J. Brewer, and J. Podos. 2006. Possible human impacts on adaptive radiation: beak size bimodality in Darwin's finches. *Proc. R. Soc. B Biol. Sci.* 273:1887–1894.
- Hewitt, G. M. 1988. Hybrid zones—natural laboratories for evolutionary studies. *Trends Ecol. Evol.* 3:158–167.
- . 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58:247–276.
- . 1999. Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* 68:87–112.
- . 2001. Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Mol. Ecol.* 10:537–549.
- . 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. B Biol. Sci.* 359:183–195.
- . 2011. Quaternary phylogeography: the roots of hybrid zones. *Genetica* 139:617–638.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965–1978.
- Holderegger, R., and C. Thiel-Egenter. 2009. A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. *J. Biogeogr.* 36:476–480.
- Huber, N. K. 1981. Amount and timing of late Cenozoic uplift and tilt of the central Sierra Nevada, California: evidence from the upper San Joaquin River basin. Geological Survey professional paper, 1197. Washington, D.C.
- Ivanova, N. V., J. R. Dewaard, and P. Hebert. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol. Ecol. Notes* 6:998–1002.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.
- Jombart, T., S. Devillard, A. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101:92–103.
- Karlstrom, E. L. 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., E. T. Reyes, J. M. Menke, D. N. Johnson, and D. L. Karavidas. 2012. Yosemite National Park vegetation classification and mapping project report. Natural Resource Technical Report NPS/YOSE/NRTR—2012/598. National Park Service, Fort Collins, CO.
- Kingsolver, J. G., D. W. Pfennig, and M. R. Servedio. 2002. Migration, local adaptation and the evolution of plasticity. *Trends Ecol. Evol.* 17:540–541.
- Knaus, B. J., and N. J. Grünwald. 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17:44–53.
- Knowles, L. L. 2000. Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. *Evolution*, 54:1337–1348.
- . 2001. Did the Pleistocene glaciations promote divergence? tests of explicit refugial models in montane grasshoppers. *Mol. Ecol.* 10:691–701.
- Koehler, P. A., R. S. Anderson, and W. G. Spaulding. 2005. Development of vegetation in the Central Mojave Desert of California during the late quaternary. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 215:297–311.
- Kuchta, S. R. 2007. Contact zones and species limits: hybridization between lineages of the California newt, *Taricha torosa*, in the southern Sierra Nevada. *Herpetologica* 63:332–350.
- Lamichhaney, S., F. Han, M. T. Webster, L. Andersson, B. R. Grant, and P. R. Grant. 2018. Rapid hybrid speciation in Darwin's finches. *Science* 359:224–228.
- Lapointe, F., and L. J. Rissler. 2005. Congruence, consensus, and the comparative phylogeography of codistributed species in California. *Am. Nat.* 166:290–299.
- Li, Y., F. Tada, T. Yamashiro, and M. Maki. 2016. Long-term persisting hybrid swarm and geographic difference in hybridization pattern: genetic consequences of secondary contact between two *Vincetoxicum* species (Apocynaceae–Asclepiadoideae). *BMC Evol. Biol.* 16:20.
- Liberal, I. M., M. Burrus, C. Suchet, C. Thébaud, and P. Vargas. 2014. The evolutionary history of *Antirrhinum* in the Pyrenees inferred from phylogeographic analyses. *BMC Evol. Biol.* 14:146.
- Lischer, H. E. L., L. Excoffier, and G. Heckel. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. *Mol. Biol. Evol.* 31:817–831.
- Loarie, S. R., B. E. Carter, K. Hayhoe, S. McMahon, R. Moe, C. A. Knight, and D. D. Ackerly. 2008. Climate change and future of California's endemic flora. *PLoS One* 3:e2502.
- Lubetkin, K. C., A. L. Westerling, and L. M. Kueppers. 2017. Climate and landscape drive the pace and pattern of conifer encroachment into sub-alpine meadows. *Ecol. Appl.* 27:1876–1887.
- Maier P. A., S. M. Ostoja, A. Aguilar, and A. J. Bohonak. 2016. Conservation genomics of the Yosemite toad: implications of deep divergence and limitations to meadow connectivity. USGS Cooperator's Report to the National Park Service. San Diego State University, San Diego, CA.
- Maier, P. A. 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, and San Diego State University, San Diego, CA.
- Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* 20:229–237.

- Mavárez, J., and M. Linares. 2008. Homoploid hybrid speciation in animals. *Mol. Ecol.* 17:4181–4185.
- Mayr, E. 1942. *Systematics and the origin of species: from the viewpoint of a zoologist*. Columbia Univ. Press, New York, NY.
- McCulloch, G. A., G. P. Wallis, and J. M. Waters. 2010. Onset of glaciation drove simultaneous vicariant isolation of alpine insects in New Zealand. *Evolution* 64:2033–2043.
- Menking, K., J. Bischoff, J. Fitzpatrick, J. W. Burdette, and R. O. Rye. 1997. Climatic/hydrologic oscillations since 155,000 yr B.P. at Owens Lake, California, reflected in abundance and stable isotope composition of sediment carbonate. *Quat. Res.* 68:58–68.
- Moritz, C. 1994. Defining “evolutionarily significant units” for conservation. *Trends Ecol. Evol.* 9:373–375.
- Mudelsee, M., and M. Schulz. 1997. The mid-Pleistocene climate transition: onset of 100 ka cycle lags ice volume build-up by 280 ka. *Earth Planet. Sci. Lett.* 151:117–123.
- Mulch, A., A. M. Sarna-Wojcicki, M. E. Perkins, and C. P. Chamberlain. 2008. A Miocene to Pleistocene climate and elevation record of the Sierra Nevada (California). *Proc. Natl. Acad. Sci. U. S. A.* 105:6819–6824.
- Mullally, D., and J. Cunningham. 1956. Aspects of the thermal ecology of the Yosemite toad. *Herpetologica* 12:57–67.
- Myers, N., R. Mittermeier, C. Mittermeier, G. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
- Noor, M. A. F., and S. M. Bennett. 2009. Islands of speciation or mirages in the desert examining the role of restricted recombination in maintaining species. *Heredity* 103:439–444.
- Nosil, P., L. J. Harmon, and O. Seehausen. 2009. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* 24:145–156.
- Ostoja, S. M., S. R. Lee, P. A. Maier, J. R. Matchett, R. A. Knapp, H. C. McKenny, M. L. Brooks, N. Danielle, and E. L. Berlow. Temporal shifts in Yosemite toad (*Anaxyrus [Bufo] canorus*) breeding meadow occupancy in Yosemite and Kings Canyon National Parks, California. Manuscript in preparation.
- Parris, M. J. 1999. Hybridization in leopard frogs (*Rana pipiens* complex): larval fitness components in single-genotype populations and mixtures. *Evolution* 53:1872–1883.
- Peter, B. M., and M. Slatkin. 2013. Detecting range expansions from genetic data. *Evolution* 67:3274–3289.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7:e37135.
- Petit, R. J., I. Aguinagalde, J.-L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R. Ennos, S. Fineschi, D. Grivet, M. Lascoux, et al. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300:1563–1565.
- Pfennig, K. S. 2007. Facultative mate choice drives hybridization. *Science* 318:965–967.
- Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Pp. 655–662 in *Proceedings of the 21st International Conference on Machine Learning*. ACM, New York, NY.
- Poorten, T. J., R. A. Knapp, and E. B. Rosenblum. 2017. Population genetic structure of the endangered Sierra Nevada yellow-legged frog (*Rana sierrae*) in Yosemite National Park based on multi-locus nuclear data from swab samples. *Conserv. Genet.* 18:731–744.
- Posada, D. 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25:1253–1256.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–59.
- Qiu, Y. X., C. X. Fu, and H. P. Comes. 2011. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world’s most diverse temperate flora. *Mol. Phylogenet. Evol.* 59:225–244.
- R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ratliff, R. D. 1982. A meadow site classification for the Sierra Nevada, California. Gen. Tech. Rep. PSW-60. Pacific Southwest Forest and Range Experiment Station, U.S. Dept. of Agriculture, Berkeley, CA.
- . 1985. Meadows in the Sierra Nevada of California: state of knowledge. Gen. Tech. Rep. PSW-84. Pacific Southwest Forest and Range Experiment Station, U.S. Dept. of Agriculture, Berkeley, CA.
- Remington, C. L. 1968. Suture-zones of hybrid interaction between recently joined biotas. Pp. 321–428 in T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. *Evolutionary biology*. Plenum, New York, NY.
- Reyer, H. U. 2008. Mating with the wrong species can be right. *Trends Ecol. Evol.* 23:289–292.
- Rissler, L. J., R. J. Hijmans, C. H. Graham, C. Moritz, and D. B. Wake. 2006. Phylogeographic lineages and species comparisons in conservation analyses: a case study of California herpetofauna. *Am. Nat.* 167:655–666.
- Ritz, C., and A. N. Spiess. 2008. qpcR: An R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 24:1549–1551.
- Rovito, S. M. 2010. Lineage divergence and speciation in the Web-toed salamanders (Plethodontidae: *Hydromantes*) of the Sierra Nevada, California. *Mol. Ecol.* 19:4554–71.
- Rubidge, E. M., J. L. Patton, and C. Moritz. 2014. Diversification of the alpine chipmunk, *Tamias alpinus*, an alpine endemic of the Sierra Nevada, California. *BMC Evol. Biol.* 14:34.
- Rudman, S. M., D. Schluter, S. M. Rudman, and D. Schluter. 2016. Ecological impacts of reverse speciation in threespine stickleback. *Curr. Biol.* 26:490–495.
- Rundel, P. W. 2011. The diversity and biogeography of the alpine flora of the Sierra Nevada, California. *Madroño* 58:153–184.
- Rundel, P. W., and C. I. Millar. 2016. Alpine ecosystems. Pp. 613–634 in H. Mooney and E. Zavaleta, eds. *Ecosystems of California*. Univ. of California Press, Berkeley, CA.
- Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecol. Lett.* 8:336–352.
- Ryan, M. E., J. R. Johnson, and B. M. Fitzpatrick. 2009. Invasive hybrid tiger salamander genotypes impact native amphibians. *Proc. Natl. Acad. Sci. U. S. A.* 106:11166–11171.
- Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol. Evol.* 16:372–380.
- Schoville, S. D., and G. K. Roderick. 2009. Alpine biogeography of Parnassian butterflies during Quaternary climate cycles in North America. *Mol. Ecol.* 18:3471–3485.
- Schoville, S. D., T. S. Tustall, V. T. Vredenburg, A. R. Backlin, E. Gallegos, D. A. Wood, and R. N. Fisher. 2011. Conservation genetics of evolutionary lineages of the endangered mountain yellow-legged frog, *Rana muscosa* (Amphibia: Ranidae), in southern California. *Biol. Conserv.* 144:2031–2040.
- Schumer, M., G. G. Rosenthal, and P. Andolfatto. 2014. How common is homoploid hybrid speciation? *Evolution* 68:1553–1560.
- Seehausen, O. 2006. Conservation: losing biodiversity by reverse speciation. *Curr. Biol.* 16:334–337.
- Seehausen, O., G. Takimoto, D. Roy, and J. Jokela. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* 17:30–44.

- Shaffer, H., and G. Fellers. 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.
- Shepard, D. B., and F. T. Burbrink. 2008. Lineage diversification and historical demography of a sky island salamander, *Plethodon ouachitae*, from the Interior Highlands. *Mol. Ecol.* 17:5315–5335.
- Shepard, D. B., and F. T. Burbrink. 2009. Phylogeographic and demographic effects of Pleistocene climatic fluctuations in a montane salamander, *Plethodon fourchensis*. *Mol. Ecol.* 18:2243–2262.
- Stebbins, G. L. 1950. Variation and evolution in plants. Columbia Univ. Press, New York, NY.
- Sota T., and A. P. Vogler. 2003. Reconstructing species phylogeny of the carabid beetles *Ohomopterus* using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses. *Mol. Phylogenet. Evol.* 26:139–154.
- Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- Stange, M., M. R. Sánchez-Villagra, W. Salzburger, and M. Matschiner. 2018. Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports miocene closure of the Panamanian Isthmus. *Syst. Biol.* 67:681–699.
- Stephens, M. 2001. Phylogeography of the *Bufo boreas* (Anura, Bufonidae) species complex and the biogeography of California. M.A. thesis. Sonoma State University, Rohnert Park, CA.
- Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *Am. Nat.* 166:581–591.
- Talavera, G., V. A. Lukhtanov, L. Rieppel, N. E. Pierce, and R. Vila. 2013. In the shadow of phylogenetic uncertainty: The recent diversification of *Lysandra* butterflies through chromosomal change. *Mol. Phylogenet. Evol.* 69:469–478.
- Tziperman, E., and H. Gildor. 2003. On the mid-Pleistocene transition to 100-kyr glacial cycles and the asymmetry between glaciation and deglaciation times. *Paleoceanography* 18:1–8.
- Unruh, J. R. 1991. The uplift of the Sierra Nevada and implications for late Cenozoic epeirogeny in the western Cordillera. *Geol. Soc. Am. Bull.* 103:1395–1404.
- US Fish & Wildlife Service [USFWS]. 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.* 79:24256–24310.
- Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. Springer, New York, NY.
- Vergilino, R., S. Markova, M. Ventura, M. Manca, and F. Dufresne. 2011. Reticulate evolution of the *Daphnia pulex* complex as revealed by nuclear markers. *Mol. Ecol.* 20:1191–1207.
- Viers, J. H., S. E. Purdy, R. A. Peek, A. Fryjoff-Hung, N. R. Santos, J. V. Katz, J. D. Emmons, D. V. Dolan, and S. M. Yarnell. 2013. Montane meadows in the Sierra Nevada: changing hydroclimatic conditions and concepts for vulnerability assessment. Center for Watershed Sciences Technical Report (CWS-2013-01), University of California Davis, Davis, CA.
- Vonlanthen, P., D. Bittner, A. G. Hudson, K. A. Young, R. Müller, B. Lundsgaard-Hansen, D. Roy, S. Di Piazza, C. R. Largiader, and O. Seehausen. 2012. Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* 482:357–362.
- Wallis, G. P., J. M. Waters, P. Upton, and D. Craw. 2016. Transverse alpine speciation driven by glaciation. *Trends Ecol. Evol.* 31:916–926.
- Waltari, E., R. J. Hijmans, A. T. Peterson, Á. S. Nyári, S. L. Perkins, and R. P. Guralnick. 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One* 2:e563.
- Wang, I. J. 2012. Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. *Divers. Distrib.* 18:1033–1041.
- Webb, W. C., J. M. Marzluff, and K. E. Olmland. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Mol. Ecol.* 20:2390–2402.
- Weber, E., and C. M. D'Antonio. 2000. Conservation implications of invasion by plant hybridization. *Biol. Invasions* 2:207–217.
- Weir, J. T. and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. *Proc. R. Soc. B Biol.* 271:1881–1887.
- Weisrock, D. W., S. D. Smith, L. M. Chan, K. Biebouw, P. M. Kappeler, and A. D. Yoder. 2012. Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. *Mol. Biol. Evol.* 29:1615–1630.
- Weixelman, D. A., B. Hill, D. J. Cooper, E. L. Berlow, J. H. Viers, S. E. Purdy, A. G. Merrill, and S. E. Gross. 2011. A field key to meadow hydrogeomorphic types for the Sierra Nevada and southern Cascade ranges in California. *Gen. Tech. Rep. R5-TP-034*. U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Region, Vallejo, CA.
- Went, F. W. 1948. Parallels between desert and alpine flora in California. *Madroño* 9:241–249.
- Wiens, J. J. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution* 58:193–197.
- Wiens, J. J., D. D. Ackerly, A. P. Allen, B. L. Anacker, L. B. Buckley, H. V. Cornell, E. I. Damschen, T. Jonathan Davies, J.-A. Grytnes, S. P. Harrison, et al. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* 13:1310–1324.
- Wiens, J. J., and C. H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Syst.* 36:519–539.
- Wood, S. H. 1975. Holocene stratigraphy and chronology of mountain meadows, Sierra Nevada, California. Ph.D. dissertation. California Institute of Technology, Pasadena, CA.
- Woolfenden, W. B. 1996. Quaternary vegetation history. Pp. 47–70 in *Sierra Nevada ecosystem project: final report to Congress*. Vol. 2, assessments and scientific basis for management options. Report No. 37. Centers for Water and Wildland Resources, University of California, Davis, CA.
- Zeisset, I., and T. J. C. Beebe. 2013. Donor population size rather than local adaptation can be a key determinant of amphibian translocation success. *Anim. Conserv.* 16:359–366.

Associate Editor: D. Weisrock
Handling Editor: D. W. Hall

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Average number of loci, SNPs, haplotypes, and locus coverage for each park and dataset.

Table S2. Summary of population genetic parameters calculated for each meadow and averaged at the lineage level.

Table S3. Summary of population genetic parameters for each meadow.

Table S4. Divergence date estimates for major lineages estimated in BEAST ultrametric tree, with clock rate prior set to confidence intervals of best available amphibian nuclear genetic clock.

Table S5. Environmental variables used for phylogeographic analyses, and their sources.

Table S6. Schoener's niche overlap metric D denoting pairwise overlap in climatic niche based on ordinated BioClim values for Yosemite NP.

Table S7. Summary of Hiest results for each meadow, along three spatial transects (East-North, East-West, East-South).

Figure S1. Spatial PCA of Lineage and Inter-Lineage Ancestry.

Figure S2. STRUCTURE and NewHybrids Evidence for Inter-Lineage Admixture.

Figure S3. BEAST Evidence of Inter-Lineage Admixture.

Figure S4. Maximum Likelihood Phylogram for Yosemite Toads.

Figure S5. Bayesian Chronogram for Yosemite Toads.

Figure S6. SNAPP tree of lineage history under the multispecies coalescent.

Figure S7. Maximum Likelihood Phylogram Balancing Yosemite/Western Toads.

Figure S8. Hypothesized Models of Admixed Lineage Origin.

Figure S9. Spatial Demographic Expansion from Glacial Refugia.

Figure S10. Simulated Values of S and H_1 After Admixture.