SAN DIEGO STATE UNIVERSITY

Conservation Genomics of the Yosemite Toad (*Anaxyrus canorus*): Implications of Deep Divergence and Limitations to Meadow Connectivity

By Paul A. Maier[1,2], Steven M. Ostoja[3], Andres Aguilar[4] and Andrew J. Bohonak[1]*

[1]Department of Biology, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182; [2]U.S. Geological Survey, Western Ecological Research Center, Sacramento, CA; [3]USDA California Climate Hub, John Muir Institute of the Environment, University of California, Davis, 1 Shields Ave., Davis, CA 95616; [4]Department of Biological Sciences, California State University, Los Angeles, Hertzberg-Davis Forensic Science Center, 5151 State University Dr., Los Angeles, CA 90032

October 11, 2016

*Contact Author: AJB, Phone 619-594-0414, Email abohonak@mail.sdsu.edu

Cover photos by Paul Maier, SDSU / USGS

Table of Contents

Figures and Tables

Conservation Genomics of the Yosemite Toad (*Anaxyrus canorus*): Implications of Deep Divergence and Limitations to Meadow Connectivity

By Paul A. Maier, Steven M. Ostoja, Andres Aguilar and Andrew J. Bohonak

Executive Summary

Primary Objectives:

1. Evaluate genetic diversity across the range of extant populations to identify genetically unique, isolated, or source populations to inform conservation priorities. Identify the spatial scale(s) at which populations are structured.
2. Determine which environmental, topographic, and climatic attributes most facilitate or limit population connectivity.
3. Characterize the extent of source-sink structure among toad meadows, and environmental correlates of sources and sinks.

Summary of Methods:

- Our sampling scheme is the first to include all known occupied regions of Yosemite toads (*Anaxyrus canorus*) in Yosemite and Kings Canyon National Parks.
- We use ddRADseq to produce thousands of unlinked nuclear SNPs and haplotypes
- Genetic structure, diversity, and isolation were analyzed using several approaches including hierarchical Bayesian clustering, multivariate ordination, tests for population differentiation, maximum likelihood phylogenetic reconstruction, and admixture analysis.
- We used hundreds of remotely-sensed and ground-collected habitat attributes to assess which ones had significant impacts inhibiting or facilitating toad gene flow.
- The toad occupies two distinct types of habitat, "meadow" and "dispersal." Therefore we employed a novel gravity modeling approach to ascertain which environmental features drive meadow connectivity at and between sites, and at multiple scales.
- By modeling asymmetrical rates of gene flow, we characterized the extent of source-sink dynamics for several genetic neighborhoods of toads. We used linear modeling and principal components analysis to pinpoint environmental correlates of sources and sinks.

Summary of Results:

- We identified hierarchical genetic structure. Meadow boundaries typically align with current gene pool boundaries. Prehistoric boundaries resulting from glacial refugia had a strong isolating effect on populations in the ancient past, and current populations retain this structure in allele frequencies.
- Intermediate to these two groupings we found a strong pattern of genetic isolation by distance, where the genetic similarity of any two meadows is a strong function of their spatial separation. We identified intermediate groupings that we call meadow "clusters," but found meadows were a better representation of gene pool boundaries than clusters.
- In several cases neighboring meadows were undifferentiated. For example, multiple meadows near Tioga Pass (Yosemite NP) and Martha Lake (Kings Canyon NP) function as homogeneous populations.

- Some meadows sampled in 2011–12 were temporally structured in Kings Canyon. This meadow substructure might exist due to assortative mating of family groups.
- Prehistoric boundaries resulted in three major rifts inside Yosemite National Park, the largest of which is likely from early in the species' history. We analyzed the resulting contact zone and found inter-lineage genetic admixture, and stable admixed populations.
- Certain meadows are more genetically distinct than others. For example, meadows near Isberg Pass (admixed), Slide Canyon (admixed), and Rancheria Creek are very genetically distinct in terms of their multilocus genotypes as identified by hierarchical Bayesian clustering.
- Gene flow was highest given lower slope and higher elevation. We found that roads and trails form barriers to gene flow. Hydrology, climate change, and vegetation at and between meadows are also drivers. Gene flow within each of four phylogeographic groups of toads was driven by unique weights of common predictors.
- We found that wetter montane meadows with longer hydroperiods produce larger tadpoles, however these meadows are less connected, possibly due to higher site fidelity.
- Our metapopulation analysis found that within each cluster of 5+ toad meadows, 1–3 meadows are genetic sinks. These highly admixed meadows have a net influx of genetic migration, and high diversity. Physically, these sink meadows tend to be larger and flatter.

Summary and Future Directions:

- Meadows are distinct biological units that are isolated by distance, and overlaid with additional complex structure. Meadows fall into clusters of genetically similar meadows, and clades resulting from population isolation early in the species' history. Assortative family group mating may additionally drive structure below the meadow level.
- Meadows are separate Management Units (MUs). If reintroductions or translocations are used, toads from nearby meadows, preferably ones that are admixed ("genetic sinks"), will likely be most successful.
- Toads from different lineages are separate Evolutionarily Significant Units (ESUs) until further evidence suggests absence of lineage-specific adaptations and inter-lineage outbreeding depression.
- Since toads in the Y-East and Y-North groups are much more diverse than toads in Y-South and Y-West, genetic rescue by translocation has the potential to mitigate future population declines. Any future translocations would benefit from first testing for local adaptation to hydrological and climatic conditions within each ESU.
- Admixture between two highly distinct groups of toads is occurring across a large area. It is likely that their gene copies are moving far beyond this contact zone. Our work suggests these individuals are viable and highly diverse. We suggest the possibility that such a zone is a crucible for adaptive diversity to combat stressors such as disease and drought. Utilizing this natural contact zone for future research on whether certain admixed toads have increased resistance to desiccation and disease would be beneficial.
- Tioga Road seems to increase mortality of migrant toads in a measureable way among toad metapopulations.
- Crossing hiking trails between meadows also seems to increase migrant mortality in Yosemite. The most genetically depauperate clade of toads (Evolution Canyon in Kings Canyon) is adjacent to the most highly used trail in our study area.

Glossary

adaptive – referring to a genotype or trait; conferring high relative fitness to individuals possessing it
adaptive diversity – diversity of genotypes in a population with potential for future trait evolution in new conditions
adaptive introgression – formation of adaptive recombinant genotypes when natural selection acts during admixture
admixture – mating between two individuals of different lineages or populations that produces offspring
allele – one particular form of a gene (diploid organisms have two of each)
backcross – the progeny of F1 and pure individuals, or any progeny with asymmetrical genetic ancestry
chytridiomycosis – infectious fungal disease in amphibians caused by *Batrachochytrium dendrobatidis* (Bd)
evolutionarily significant unit (ESU) – historically isolated lineage that may have adapted separately
F1 – the progeny of two pure individuals (i.e. the first filial generation after hybridization or admixture)
F2 – the progeny of two F1 individuals (i.e. the second filial generation after hybridization or admixture)
fitness – the net effect of viability, mating success, and fecundity that determines number of progeny
genotype – the combination of alleles at one or more DNA loci that one individual has for making proteins
hybridization – mating between two individuals of different species that produces offspring
inbreeding depression – lowered fitness caused by relatives breeding with relatives in small populations
lineage – a population having an independent evolutionary trajectory (i.e. branch on a phylogenetic tree)
management unit (MU) – genetically distinct population reflecting recent breeding isolation from other populations
metamorphosis – abrupt developmental change causing one species to occupy more than one niche
monophyletic – a group containing an ancestor and all its descendants (populations or species)
natural selection – differential survival and reproduction among individuals due to heritable trait differences
paraphyletic – a group containing an ancestor but only some of its descendants (populations or species)
population – a group of related individuals in the same geographic area that freely interbreed
recombinant genotype – a genotype with novel combinations of alleles arising from hybridization or admixture
secondary contact zone – region where two formerly isolated species or lineages are hybridizing or admixing
source-sink structure – populations with high habitat quality that recolonize neighboring low quality populations
species – separately evolving lineages showing reproductive isolation under most conditions (many definitions)

Introduction

The Yosemite toad (*Anaxyrus canorus*) is a species of high-elevation anuran that is endemic to the central Sierra Nevada of California. It exclusively breeds in the transient and highly productive water bodies of mountain meadows (Ratliff 1985) which make up <3% of the landscape. While tadpoles develop and metamorphose in meadow pools and flooded areas, adult toads forage, hibernate, and can disperse into intervening habitat over 1 km per season (Liang 2010). Yosemite toads are reputed to be declining in both distribution and abundance (Sherman & Morton 1993; Drost & Fellers 1994; Jennings & Hayes 1994; Brown & Olsen 2013; US Fish & Wildlife 2014; but see Ostoja et al. 2015). Despite extensive research into the causes of mortality and declines, such as UV radiation (Sadinski 2004), exotic predators (Grasso et al. 2010), meadow grazing (Roche et al. 2012a; b; Matchett et al. 2015), pesticide use (Sadinski 2004), and chytridiomycosis (Dodge et al., in prep), no clear patterns have emerged for the entire species (Brown et al. 2015). Climate change is ostensibly one of the greatest threats to Yosemite toad breeding ecology (Viers et al. 2013; US Fish & Wildlife Service 2014), but little research has been done to connect it with declines. A conservation genomic analysis is an important tool for highlighting the most effective mitigation actions against poorly understood threats; by describing the structure of populations, pinpointing sources of genetic diversity or uniqueness, and learning how to most effectively facilitate gene flow, managers will be better able to make informed judgments about future conservation actions.

Genetic connectivity is an essential force for fostering long-term persistence of a species with discrete populations by mitigating inbreeding depression and Allee effects, and promoting diversity. Pond-breeding amphibians with complex biphasic life histories present a special challenge to the study of landscape genetics, particularly when larvae are restricted to isolated patches. The process of emigration is at least partly influenced by habitat quality, larval recruitment, and density-dependence (Pulliam 1988; Matthysen 2005; Mathieu et al. 2010), while transient migration and immigration depend on adult behavior and limits to dispersal (Baguette & Van Dyck 2007; Clobert et al. 2009). The probability of Yosemite toad meadow occupancy and breeding is known to partly depend on the quality of other meadows in a local neighborhood (Berlow et al. 2013), and Wang (2012) found that 42.6% of genetic structure between meadows was attributable to slope and environment (including climate). However, virtually nothing is known about how meadows and the intervening landscape influence overall genetic network structure, or whether the change in specific climatic variables is important.

Predictive spatial models in landscape genetics and genomics are an important, burgeoning tool for summarizing evolutionary processes and streamlining decision-making (Vandergast et al. 2008; Sork et al. 2010; Fitzpatrick & Keller 2014). Gravity modeling (Anderson 1979; Murphy et al. 2010a) or linear mixed modeling (van Strien et al. 2012) can account for biphasic life history that spatially partitions habitat requirements. However, another important random effect implicit in models is phylogeographic history that might cause lineages to interact differently with their environments over large spatial areas. This is because these lineages have remained separate long enough to evolve separate adaptations to the same environment. While some previous studies have modeled local adaptive gradients that are countered by gene flow (e.g. Fitzpatrick & Keller 2014), this reverses the usual causal hypothesis in adaptive landscape genetics because we speculate that lineages are differentially adapted to environmentally-mediated gene flow.

Study Goals

Goal #1: Evaluate genetic diversity across the range of extant populations to identify genetically unique, isolated, or source populations to inform conservation priorities. Identify the spatial scale(s) at which populations are structured. A previous study (Shaffer & Fellers 2000) suggested Management Units are less than a few kilometers across, although that study had limited sampling coverage and evenness. More recently, the United States Geological Survey (USGS) used a newly delimited GIS layer of meadow polygons (Keeler-Wolf et al. 2012) to perform a systematic 6-year census of Yosemite toad occupancy at the meadow scale (Ostoja et al. 2015), which revealed that occupied toad meadows form naturally distinct and hierarchical clusters across the landscape. Brown et al. (2015) identified an important conservation priority for Yosemite toad: quantifying "the spatial scales over which populations function…and [their] variation." In addition, previous studies have suggested a large phylogenetic discontinuity residing somewhere between southern Yosemite and Sonora Pass (north of Yosemite). Using mitochondrial DNA, these studies have inferred mitochondrial DNA paraphyly at the level of the species (Graybeal 1993, 1997; Shaffer & Fellers 2000; Stephens 2001; Goebel 2005; Goebel & Ranker 2009). However, nuclear gene trees have subsequently suggested this latter pattern is an artifact of mitochondrial introgression, and supported a monophyletic *A. canorus* (Pauly, pers. comm.). We leveraged a high-density genomic dataset to resolve the ancestor-descendent relationships and population genetic structure to (1) identify the scale of connectivity in this species, (2) identify unique populations or groups, and (3) locate and characterize major genetic or phylogenetic rifts in our study area. We chose to use a highly robust sampling scheme to reduce bias associated with unsampled locations, and to sequence thousands of anonymous nuclear double-digest RADseq (ddRAD) markers to precisely identify phylogeographic structure (traces of population structure from ancient history), and identify gene pools (current population structure), as well as intermediate clusters.

Goal #2: Determine which environmental, topographic, and climatic attributes most facilitate or limit population connectivity. We used a novel combination of elements to answer this question and address general issues in landscape genetics. Topographic complexity and climate are known to be important predictors of connectivity (Wang 2012), and the network characteristics of meadow neighborhoods contribute to breeding probability (Berlow et al. 2013), so we hypothesized that each of these was important for genetic connectivity. Anthropogenetic impacts of packstock and livestock have been studied extensively with little evidence for an impact on toad breeding ecology (Roche et al. 2012a; b; Matchett et al. 2015), although other recreational activities might also be important, such as roads and trail crossings. Finally, we considered numerous measures of hydrology and climate change. Climate change is known to cause phenological and range shifts as well as genetic modulation in many organisms (reviewed in: Parmesan 2006), and specifically in amphibians (Alexander & Eischeid 2001; Pounds 2001; Pounds et al. 2006). Because Yosemite toads are obligate meadow pond breeders, tadpole survival in this species depends critically upon snowpack runoff and spring recharge. Significant larval mortality is often observed even during years with above-average snowpack (Sherman 1980; Sherman & Morton 1993; Brown et al. 2015). We examined the effect of climate and climate change variables by using the 2014 California Basin Characterization Model (BCM; Flint & Flint 2012; Flint et al. 2013), which consists of downscaled 270m rasters derived from 800m PRISM data (temperature and precipitation) as well as elevation, geology, and soils data. If we detected any influence of present climate on gene flow, we asked whether the mean change in those climate variables over 60 years has also impacted genetic patterns.

Because choice of scale can significantly improve the fit of landscape genetic models or emphasize different spatiotemporal processes (Anderson et al. 2010; Galpern et al. 2012), we separately modeled alternative dispersal transect cutoffs. We also incorporated phylogenetic signal (discussed below) as a random factor in the analyses to test the hypothesis that toad dispersal in each clade has adapted to separate environmental limitations. Finally, we assessed the relative contributions of meadow versus intervening landscape habitat attributes to genetic connectivity among toad populations. Ideally, these two habitat permeabilities should be independently modeled because (a) this will improve accuracy of predictions, and (b) patch-specific influences on gene flow may be more tractable for land managers. To incorporate all of these considerations, we incorporated aspects of causal modeling (Cushman et al. 2006), least cost transect analysis (van Strien et al. 2012), and gravity modeling (Murphy et al. 2010a).

Goal #3: Characterize the extent of source-sink structure among toad meadows, and environmental correlates of sources and sinks. We asked whether genetic sources were also meadows with high ecological value (probability of breeding occupancy). High quality meadows could be genetic sources if these meadows recruit a larger proportion of adult dispersers through density-dependent effects, or if toads in lower quality "sink" meadows are locally adapted and have lower reproductive success at higher quality "source" meadows. However, high quality meadows might be genetic sinks if toads in higher quality meadows show higher rates of philopatry, or if these toads are locally adapted to the ecology, either due to physiological constraints on survival or premating behavioral isolation. Identifying source-sink dynamics will help identify robust, unique, and declining meadows that can be earmarked for specific conservation actions, such as relocations. We used estimates of asymmetrical migration and graph theory to examine network patterns of source-sink dynamics. Because Yosemite toads occupy essentially two different biotic zones (Fig. 1), it is reasonable to expect these dynamics might look different in montane metapopulations versus subalpine ones. If there are underlying ecological drivers for source-sink dynamics, this will help management to extrapolate these patterns to Yosemite toad in other national parks and forests.

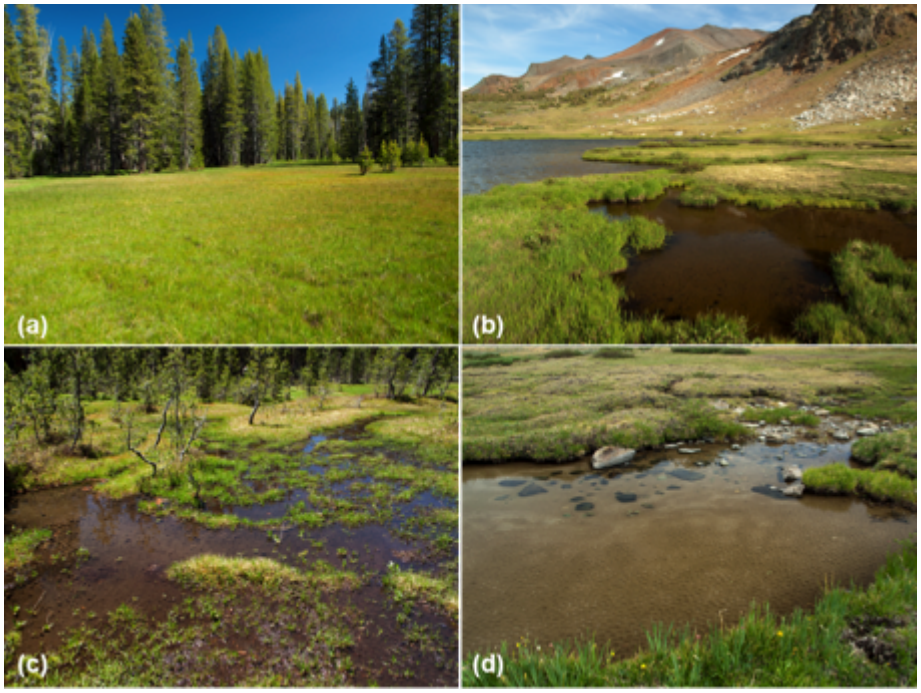


Figure 1. Yosemite Toad Breeding Habitat
Examples of Yosemite toad meadows (a, b) and corresponding natal pools (c, d) for developing larvae. Yosemite toads occupy meadows that can be broadly categorized as montane (a, c – Chilnualna Creek, Yosemite NP) and subalpine (b, d – Ireland Lake, Yosemite NP). Photo credit: Paul Maier.

Methods

Study Area

Yosemite toads are meadow-breeding specialists restricted to the central Sierra Nevada of California, and found from approximately 6,400' to 11,500' (Mullally & Cunningham 1956; Karlstrom 1962; USGS unpublished data). 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*), while 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. 1). Yosemite National Park (3,027 km^2) was chosen as the primary study area because it overlaps with previous studies (Shaffer & Fellers 2000; Wang 2012; Berlow et al. 2013) and offers a representative snapshot of ecological conditions experienced by the species; however, we also included samples from Kings Canyon National Park for reference and future work (Fig. 2). We used Yosemite to parameterize landscape genetic models, but analyzed genetic structure in both parks. Yosemite is heavily forested in the southwest and pocked with glacially carved cirques, hanging valleys, moraines, and canyon systems at higher elevations. The Merced and Tuolumne River canyons present major biogeographical barriers along east-west axes, and numerous canyons feed into the Tuolumne from the north. The snow-free season ranges from 3 – 6 months depending on elevation, latitude, and year (pers. obs.).

Genetic Sampling

Tail tissue was collected from larval toads using a sterilized syringe during summers of 2011-2013. All samples were preserved in 95% EtOH and stored at -20°C within one week. An

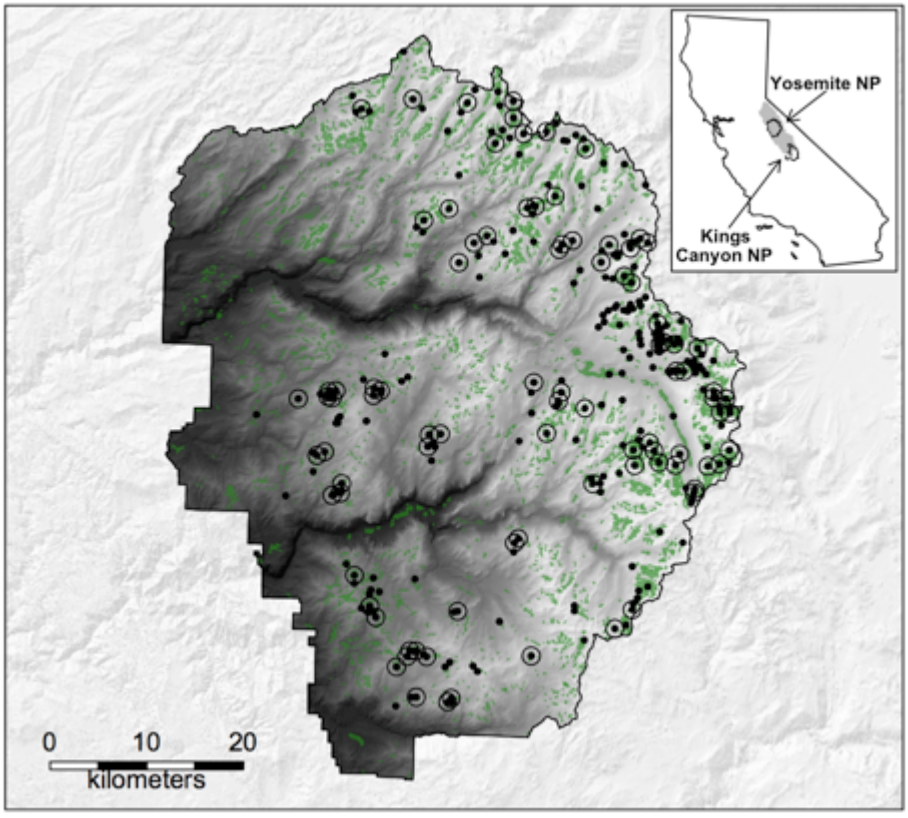


Figure 2. Study Area Study area in Yosemite National Park, CA includes approximately 30% of the range of Yosemite toads. Inset shows the range of Yosemite toads in gray, and the boundaries of Yosemite and Kings Canyon National Parks in black. Green polygons are all 2928 meadows within Yosemite. Solid black circles indicate all known Yosemite toad meadows identified between 1924 and the present. Open black circles indicate the 90 meadows sampled and sequenced in the present study. 8 samples of *A. boreas halophilus*, 1 sample of *A. punctatus* and 109 *A. canorus* samples sequenced from Kings Canyon are not shown.

existing USGS meadow layer based on a park-wide vegetation map with 0.5 ha resolution was used to delimit sampling sites (Keeler-Wolf et al. 2012; Berlow et al. 2013). Sites were chosen to maximize representation across all known breeding locations from a recent 6-year survey effort (Ostoja et al. 2015, in review) and overlap with previous studies. Tadpoles were sampled across multiple clutches, pools, and years to maximize inclusion of available genetic diversity. Additionally, tadpoles were measured for tail length (TL) to the nearest mm, and assigned Gosner (1960) field stages as follows: 1 (Gosner 1-25), 2 (Gosner 26-30), 3 (Gosner 31-35), 4 (Gosner 36-39), 5 (Gosner 41), 6 (Gosner 42-44), 7 (Gosner 45), 8 (Gosner 46). 8 *Anaxyrus boreas* and one *A. punctatus* toe clips were obtained as outgroups for phylogenetic analysis from the Museum of Vertebrate Zoology (University of California, Berkeley) and USGS (San Diego Field Station).

Molecular Methods

We chose double-digest RADseq (ddRAD) markers over traditional markers such as microsatellites for two reasons: (1) thousands of markers afford a high precision on parameter estimates such as F_{ST} (population divergence) or N_e (effective population size), and (2) a genomewide analysis can reveal subtle genetic patterns that are restricted to particular parts of the genome, such as inter-lineage admixture or selective sweeps.

A total of 653 samples were chosen for sequencing (535 samples from Yosemite, 109 from Kings Canyon, and 9 outgroup samples). A minimum of 5 samples was used per meadow grouping (defined as having 1+ additional meadows sequenced within 1 km), and 10 per meadow for remaining meadows, unless insufficient samples were available. Methods for DNA extraction, ddRADseq library preparation, and Illumina sequencing are described in Appendix A.

Bioinformatic Data Processing

Raw DNA sequence data were processed into either single nucleotide polymorphisms (SNPs, biallelic loci with two of either A, T, C, or G) or haplotypes (alleles encoded by integers describing each unique combination of SNP alleles in a DNA sequence). The full bioinformatic pipeline is described in Appendix A. In brief, a genotype matrix of SNPs and haplotypes was created for analyses of genetic structure, phylogenetic history, and population differentiation. SNPs are best suited for unbiased estimates of F_{ST} (population differentiation) and for the reconstruction of phylogenetic history, which models the mutational process for each nucleotide separately. (Phylogenies based on only SNPs must be appropriately corrected, because SNPs are a biased sample of DNA sequences that only includes variable nucleotides.) Haplotypes can be more informative about genetic structure since each locus contains on average more diversity.

Genetic Structure

We performed multiple analyses to discover and characterize spatial patterns of genome-wide structure across Yosemite and Kings Canyon (complete methods in Appendix A). Our goal was to delineate the boundaries of current populations and rank them by uniqueness. In contrast to current population connectivity, historical isolation from geological and climatic events during the Pleistocene is important to characterize because this genetic structure can be indicative of incompatible local adaptations or even early reproductive isolation between alternative regions.

First we checked all loci for deviations from Hardy-Weinberg and linkage disequilibrium assumptions. We then used a spatial principal components analysis (sPCA) to look for major genetic discontinuities that correspond to their geographic arrangement in Yosemite. We

followed this up with maximum likelihood estimation of population phylogeny, using meadows as operational taxonomic units. We included all Yosemite toad samples from Yosemite and Kings Canyon, as well as *Anaxyrus boreas* samples from across California, and one *Anaxyrus punctatus* outgroup sample. Based on the correspondence between these last two tests, we estimated the dates of historical divergence among major clades using full sequence alignments from 2 – 4 individuals per clade, and then running relaxed-clock, Bayesian estimation of the coalescent times. This method can estimate relative dates of divergence, but a known rate of evolution (in substitutions per year) must be used to calibrate these dates into actual times. Thus, we focused more on the relative divergence dates than absolute dates.

We found that two historically isolated lineages of toads had come into secondary contact in northern Yosemite (see "Results"), and so we estimated the proportion of genetic ancestry that admixed toads have inherited from these two lineages. We also inferred how viable these admixed toads are by estimating the length of time (in number of generations) that admixture has been occurring. Any reproductive incompatibilities resulting from admixture, if they existed, would curtail the extent of admixture to one or a few generations.

We delineated current population boundaries in four steps. First, we noted that meadows are differentiated from each other strongly as a function of their geographic separation. This isolation-by-distance (IBD) pattern can confound Bayesian clustering programs, which try to assign individuals into a set of clearly differentiated ancestral populations. Hence, we used two such programs to delineate population boundaries: (1) a hierarchical STRUCTURE analysis which accounts for this pattern of IBD by re-running each analysis recursively until no further sub-structure is found, and (2) an ADMIXTURE analysis that does not account for IBD, although is better able to partition large datasets into one set of populations. Finally, we used a hierarchical analysis of molecular variance (AMOVA) to assess which of these two results is a better reflection of gene pool boundaries.

Landscape Genetic Analysis

We implemented a three-step landscape genetic procedure to determine the most important environmental features impacting toad gene flow between populations. We used meadows as units, since we found meadows were the best population unit in most cases (see results). First, we constructed plausible dispersal paths between each pair of meadows based on environmental resistance. Previous research and expert opinion suggested that higher slope, ridgelines, and drier vegetation might limit dispersal. Hence, we built 11 hypothesized cost rasters that combined these variables in various combinations, constructed a model of least cost

Table 1. Environmental data sources. Res. = spatial resolution (the length represented by one pixel).

Source	Year	Res.	Type	Comments
California BCM 2014	1920-2010	270m	Hydrology, temperature	Downscaled from 800m PRISM
SRTM	2000	10m	Topography	National elevation dataset
Yosemite Vegetation Map	2012	10m	Vegetation, hydrology	Aerial imagery
MODIS	2002-2007	500m	Snow meltoff date	Daily satellite albedo
LANDSAT 5	1984-2011	30m	Veg., moisture, burn severity	Satellite images, CDR processed
Cal Fire	1900-2014	10m	Fire frequency	Rasterized fire perimeters
Daymet	1980-1997	1000m	Moisture	Alternative to BCM
Cal Water	2004	10m	Watershed attributes	CalWater 2.2.1 boundaries
Yosemite National Park	2016	10m	Trails, Roads	Rasterized, weighted by traffic
USGS survey data	2009-2014	—	Meadow network attributes	Details in Berlow et al. (2013)

paths for each raster, and used partial mantel tests to choose the least cost path model that maximized correlation with pairwise meadow F_{ST} after accounting for pairwise meadow distance. This step, similar to causal modeling, selected the most likely path that toads use for dispersal between each pair of meadows.

Second, we constructed 100-meter buffers around these least cost paths as well as the breeding locations themselves, and extracted the average or standard deviation of remotely sensed and USGS survey data values between and at sites (Table 1). We considered a total of 22 variables (2 variables from each of 11 categories), which we selected by considering variable importance in a random forests analysis (Table S1).

Third, we used linear mixed modeling to find likely and parsimonious predictors of meadow connectivity. This method (similar to gravity modeling, see Appendix A) accounts for the non-independence of at- and between-site data points by using site (source meadow) as a random intercept. In addition, we suspected that toads in separate clades might have evolved different sensitivities to environmental effects, so we included random slopes by clade for any environmental variable, if doing so significantly improved the model without overfitting. We built and evaluated models using a combination of stepwise addition, AICc, and likelihood ratio tests. In total, we optimized 6 models: those with transect cutoffs at three distances (10km, 20km, 30km) in order to infer whether short-term and long-term dispersal are driven by different environmental processes, and two transect types (least cost path, straight-line) to assess whether our buffered least cost path procedure significantly improved models. A detailed description of this procedure can be found in Appendix A.

Metapopulation Dynamics

We tested the hypothesis that clusters of nearby meadows included strongly "source" and "sink" meadows using the method of Sundqvist et al. (2016). Specifically, we hypothesized that a few meadows would be sources of migrants for nearby satellite meadows. In essence, this method calculates geometric means of allele frequencies for a theoretical pool of migrants between each pair of meadows using alleles present in both meadows. Then asymmetrical G_{ST} values are calculated by comparing each meadow to the migrant pool. We identified environmental correlates of source and sink meadows using mantel correlations, multiple regression of distance matrices, and principal components analysis. For each metapopulation we examined, estimates were mapped onto igraph (Csárdi & Nepusz 2006) objects to display the extent of source-sink dynamics in a network.

Table 2. Average coverage and number of loci, SNPs, and haplotypes for each park and dataset. For a complete description of "combined" and "concatenated" datasets, see supplemental methods in Appendix A.

	Yosemite		Kings Canyon	
	Combined	Concatenated	Combined	Concatenated
Avg. Coverage	87.13	43.99	73.58	41.27
No. SNPs	12986	3132	10448	4025
No. Haps	10299	2252	9272	3247
No. SNP Loci	4350	1044	3574	1341
No. Hap Loci	3887	679	3202	1021
Avg. SNPs/Locus	2.99	3.00	2.92	3.00
Avg. Haps/Locus	2.65	3.32	2.90	3.18
Max. Haps/Locus	6	9	7	6

Results

Illumina Data Quality and Quantity

In total, 1.88 billion reads were returned from 7 lanes of an Illumina HiSeq 2500 run, of which 1.60 billion were identifiably barcoded. This included 161.68 gigabases (Gb) of useable sequence data with a mean quality score of 35.57. Quality scores are logarithmic measures of the proportion of bases incorrectly called during sequencing. For example, a quality score of 10 means that 1 in 10 bases is incorrectly called, a quality score of 20 means that 1 in 100 bases is incorrectly called, and a quality score of 30 means that 1 in 1000 bases is incorrectly called. A total of 535 individuals from 90 meadows were sequenced at 4350 loci for Yosemite, and 109 individuals from 12 meadows were sequenced at 3574 loci for Kings Canyon (Table 2). Concatenating the paired-end reads greatly reduced the number of loci passing coverage and quality thresholds, as only 1044 and 1341 concatenated loci were retained for Yosemite and Kings Canyon, respectively. Similarly, coverage was lower for concatenated data (average of 43.99 and 41.27) versus combined data (average of 87.13 and 73.58). Despite this reduction in data quantity, concatenating loci increased the average number of haplotypes that each contained, thus potentially increasing the genetic diversity contained within any particular locus. However, SNPs across approximately 200bp of sequence are in tight linkage, and hence only one SNP should be used per locus. Therefore, using haplotypic data effectively increased the number of alleles per locus in all cases from 2 SNP alleles to an average of 2.65 – 3.32 haplotypic alleles (depending on dataset), and a maximum of 9 haplotype alleles (Table 2).

Spatial Genetic Patterns

Genetic diversity showed a dramatic increase from lower (western/southern Yosemite) to higher (eastern/northern Yosemite, Kings Canyon) elevation meadows (Tables 3, S2, S3). Meadow elevation and observed heterozygosity showed a strong positive correlation, ($\rho = 0.68$, $df = 100$, $p < 0.001$), as did meadow elevation and mean gene diversity (π) ($\rho = 0.61$, $df = 100$, $p < 0.001$). Elevation is a covariate of many ecological patterns and processes, including a transition from montane red fir meadow habitats to subalpine meadow habitats, so this observation by itself is not explanatory. The number of private alleles was also positively

Table 3. Summary of population genetic parameters for each meadow and averaged at the clade level. Results summarized at the meadow and cluster levels are available in Appendix B. YOSE = Yosemite, KICA = Kings Canyon, PA = private alleles, P = frequency of most frequent allele, H_o = observed heterozygosity, H_e = expected heterozygosity, π = average gene diversity, F_{IS} = fixation index, % P Loci = proportion of variable sites that are locally polymorphic, Clade N_e = effective population size. estimated for entire clade.

Clade	Park	PA	P	H_o	H_e	π	F_{IS}	% P Loci	Clade N_e
Evolution	KICA	28	0.966	0.055	0.045	0.050	-0.011	0.291	3.300
Goddard	KICA	24	0.953	0.068	0.063	0.067	-0.001	0.438	24.300
South	YOSE	5	0.969	0.051	0.040	0.045	-0.012	0.243	23.600
East	YOSE	7	0.962	0.060	0.050	0.056	-0.007	0.307	51.700
West	YOSE	6	0.967	0.052	0.042	0.047	-0.009	0.256	33.100
North	YOSE	18	0.960	0.062	0.053	0.058	-0.006	0.348	18.600
	KICA	**26**	**0.960**	**0.062**	**0.054**	**0.058**	**-0.006**	**0.365**	**13.800**
	YOSE	**9**	**0.965**	**0.056**	**0.046**	**0.052**	**-0.008**	**0.288**	**31.750**
	All	**15**	**0.963**	**0.058**	**0.049**	**0.054**	**-0.008**	**0.314**	**25.767**

correlated with elevation, ($\rho = 0.27$, df = 100 $p < 0.01$), although this pattern was not significant when excluding Kings Canyon, ($p = 0.28$), indicating that Kings Canyon toads might have higher meadow fidelity. Kings Canyon meadows not only had more private alleles than Yosemite, ($t = -3.50$, df = 12.11, $p < 0.01$), but also higher observed heterozygosity ($t = -3.62$, df = 12.67, $p < 0.01$), π ($t = -3.48$, df = 12.72, $p < 0.01$), and average number of polymorphic sites, ($t = -2.85$, df = 12.48, $p < 0.05$), in spite of having far fewer SNPs ($t = 6.75$, df = 34.51, $p < 0.001$). Overall, Kings Canyon showed a trend of higher genetic diversity than Yosemite. This is possibly because the sites in Kings Canyon have a higher mean ($t = -5.64$, df = 16.57, $p < 0.001$), and lower (albeit insignificant) variance in elevation ($F = 1.73$, df = 11, $p = 0.16$).

Spatial PCA analysis revealed 47 positive factors identifying spatially explicit discontinuities in the Yosemite genetic data. The largest ranked PC score (46.72) was over 3x the second largest (14.09), and represented northern and eastern regions flanking the Tuolumne River gorge (Fig. 3). The first three PCs (scores: 46.72, 14.09, 7.94) coincided with three major barriers: Tuolumne River gorge, Merced River gorge, and an elevational cline between the two river gorges. These three PCs divided the park into four groups, which we also found by performing DAPC analysis (Fig. S1). In addition, phylogenetic analysis using RAxML recovered a topology supporting three major clades within Yosemite ("Y-North", "Y-South", and "Y-East" + "Y-West"). This third clade consists of a monophyletic "Y-West" and paraphyletic "Y-East" (Fig. 4). The oldest divergence was between Yosemite and Kings Canyon, and was estimated to be 1.9 – 4.6 Ma (Fig. S2). The second oldest divergence between Y-North and Y-East + Y-West + Y-South was estimated at 0.95 – 2.14 Ma, and the median for this age (1.51 Ma) was 47.6% that of the oldest divergence (3.17 Ma). Interestingly, the youngest clade (Y-West clade) is estimated 305 – 671 ka, which is much older than three major glaciations that

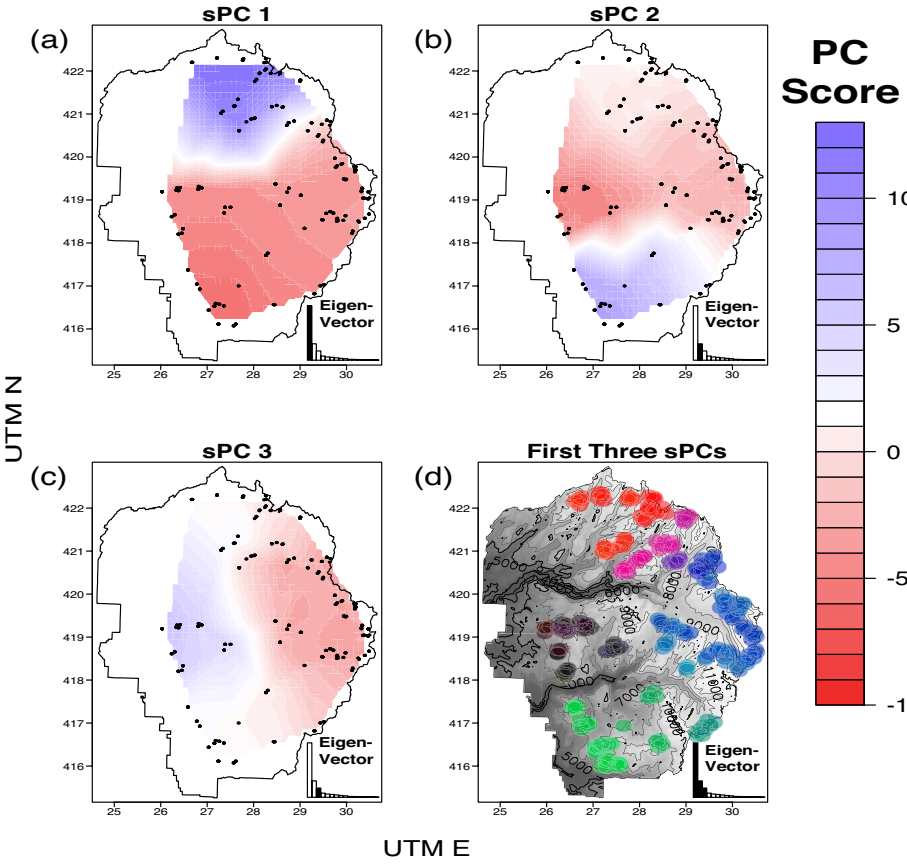


Figure 3. Spatial PCA Plot of Yosemite Toad Genetic Structure in Yosemite NP
First three panels show the first (a), second (b), and third (c) spatial components contributing to global structure (positive Moran's I values). PC scores are the product of principal components and Moran's I coefficients, thus they resemble gradients of genetic variance that follow strong spatial patterns. Panel (d) shows all three axes of multivariate spatial genetic structure (red: Y-North, blue: Y-East, black: Y-West, green: Y-South).

covered most of the Sierra Nevada: the Tahoe (maximum at 70 ka), Tenaya, and Tioga (maximum at 21 ka) glaciations. Additionally, all 8 *Anaxyrus boreas halophilus* samples from both northern and southern California formed one monophyletic lineage, sister to all *A. canorus* samples. All published phylogenetic studies of the *A. boreas* group have thus far recovered *A. canorus* as paraphyletic or polyphyletic with respect to *A. boreas* using mtDNA markers

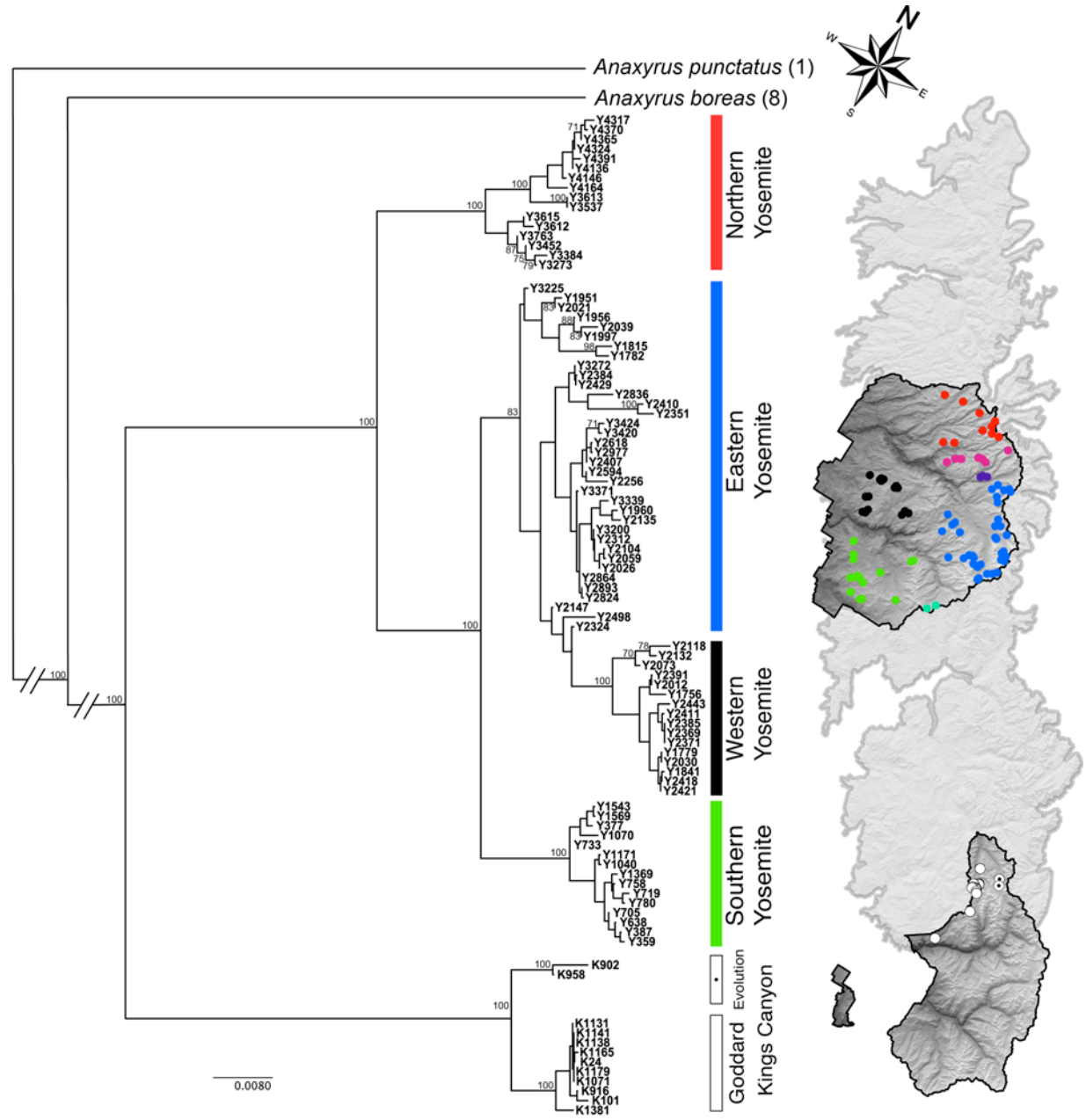


Figure 4. Maximum Likelihood Phylogenetic Reconstruction of Yosemite Toad Ancestry
Maximum likelihood (RAxML) tree with highest likelihood using 1000 bootstrap replicates and 10 ML heuristic searches given a GTR + Γ model. Bootstrap support values > 70 are shown for each node. Outgroup branches for *Anaxyrus punctatus* and *A. boreas* individuals are broken for clarity. Meadows with suspected admixture among clades based on sPCA results (Y3400, Y3414, Y3342, Y1856, Y942, Y1097; represented as intermediate purple and cyan circles) were excluded from the analysis to bolster node confidence. Maps of both national parks are overlaved on the distribution of Yosemite toads.

(Graybeal 1993, 1997; Shaffer & Fellers 2000; Stephens 2001; Goebel 2005; Goebel & Ranker 2009), although one unpublished study recovered *A. canorus* as monophyletic using nuclear genes (Pauly, pers. comm.). A more inclusive species-wide sampling scheme using ddRAD or similarly high-coverage nuclear markers will fully elucidate the ancestral relationships between these two species (Stephens, pers. comm.).

The "Goddard" clade along the S. San Joaquin River in Kings Canyon was found to have the highest genetic diversity of any clade in either park (Table 3). Compared to the second-most diverse clade, Goddard has an average meadow H_o of 0.068 (11% > Y-North), H_e of 0.063 (18% > Y-North), and 0.44% of loci are polymorphic (26% > Y-North). This was expected given that more adult toads and tadpoles are seen in Goddard canyon than anywhere else (pers. obs.), but surprising given this is the latitudinal and elevational range limit for the species. Its sister clade "Evolution," in contrast, has third lowest H_e, H_o, π, % polymorphic loci, and the lowest N_e (3.3; all others have N_e > 18). The Y-South and Y-West clades are in forested low-elevation regions,

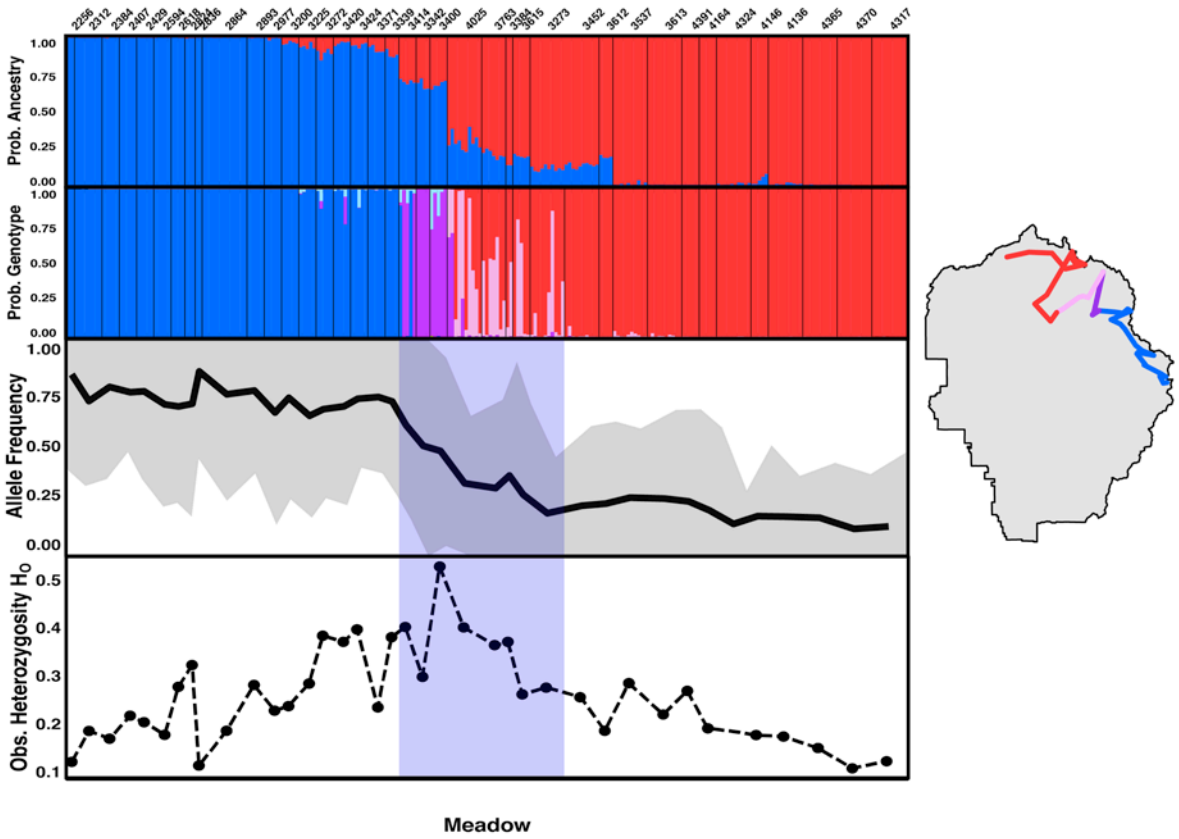


Figure 5. Contact Zone Dynamics in Northern Yosemite NP
Gradient in genotypes, admixture classes, allele frequency, and observed heterozygosity following a transect across two clades: Y-East (blue) and Y-North (red), shown on the map. All 5278 haplotypic alleles at 2945 loci are included in (a), whereas only 53 alleles from 27 loci with the highest 1% contribution to sPC 1 are used in (b-d). (a) STRUCTURE barplot showing probability of ancestry from Y-East and Y-North. K=2 was tested 10 times for 100,000 generations and 10,000 burn-in each, then combined using CLUMPP. (b) NEWHYBRIDS barplot showing probability of recent admixture across a putative contact zone. Each color represents an admixed genotype class (P1: blue, P2: red, F2: purple, F1-P1 backcross: light blue, F1-P2 backcross: pink). (c) Allele frequency gradient across the contact zone. Solid line indicates mean, gray shading indicates 95% confidence interval among loci. (d) Mean observed heterozygosity. Purple box in (c-d) highlights meadows with identified admixed genotypes from (b).

whereas Evolution is in a high-elevation subalpine cirque similar to Goddard. Overall, Y-East has the highest N_e of 51.7, and in Yosemite it has only slightly lower H_e, H_o, and π than Y-North. Y-North likely has artificially inflated values of H_e, H_o, and π because some of its meadows have admixture from both Y-North and Y-East (discussed below).

Contact Zone Dynamics

The deep divergence between Y-North and the rest of Yosemite shows an allelic, genotypic, and diversity gradient suggestive of isolation followed by recent contact, and genetic admixture between ancient clades (Fig. 5). Several lines of evidence suggest this. The estimated divergence time of these clades (1.51 Ma, 47.6% the estimated age of the Yosemite – Kings Canyon divergence, which is 3.1 Ma; Fig. S2) is very old. Northern Yosemite is filled with potentially isolating features (Fig. 2), but F_{ST} values specifically increase when crossing this and other contact zones ($F = 210.99$, $df = 2,1711$, $p < 0.001$). Using the ranked contributions of all 5278 alleles to the first spatial PC, the highest 1% showed a sharp cline in frequency along the zone, while observed heterozygosity H_o showed a spike, reflecting admixture within the same region. STRUCTURE results that included meadows from Y-East and Y-North, which were constrained by K=2, showed a similar cline in admixture between two Hardy-Weinberg gene pools. NEWHYBRIDS binned admixed individuals into three general categories, consisting of F2

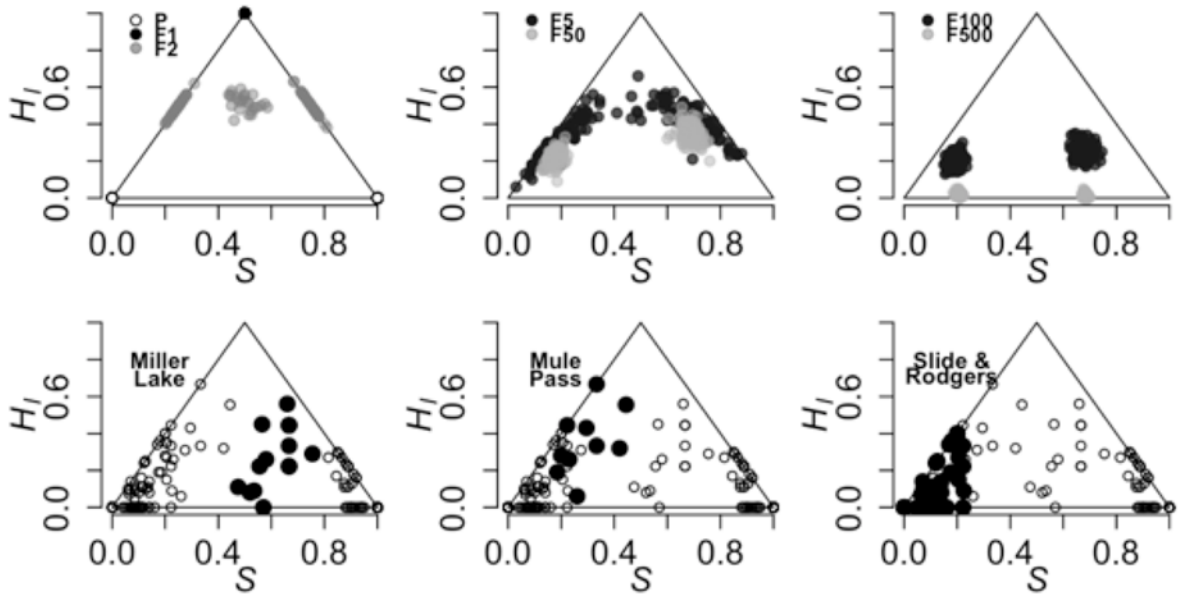


Figure 6. Admixed Ancestry and Heterozygosity in Northern Yosemite NP
Two-dimensional status of northern admixed meadows following Fitzpatrick (2012). In each panel, the horizontal axis represents percent ancestry, and the vertical axis represents heterozygosity of admixture-informative markers. Points represent individuals in one population. The top row of panels shows simulated populations with 100 individuals, 100 markers, no gene flow, and initial admixture proportion of 25% / 75%. The other panels represent observed admixture scores at Miller Lake, Mule Pass, Slide Canyon, and Rodgers Lake. From left to right indicates increasing ancestry from Y-East, and decreasing ancestry from Y-North. Pure types (P) are located in the bottom corners, F1 genotypes are in the top corner. Heterozygosity declines from 100% (F1s, black dot in top left panel) as admixed individuals breed in successive generations after initial admixture. Advanced-stage admixed individuals approach the bottom side (e.g. F500, top right), while pure backcrosses are on lateral sides. For comparison, the program NEWHYBRIDS found Miller Lake individuals to be F2s, and Mule Pass, Slide, and Rodgers to be F1-backcrosses.

and each F1-parental backcross (Fig. 5). However, the quantitative HIest approach suggested these individuals have been admixing for more than two generations (Fig. 6). Backcrosses between F1s and Y-North parental genotypes were the most abundant category across the zone, which might indicate directional introgression due to either demographic or adaptive reasons. Clearly, phylogeographic boundaries and contact zones are prominent features shaping the distribution of gene copies on the landscape. For this reason, we incorporated the three clades (Y-North, Y-South, Y-West) and one paraphyletic group (Y-East; hereafter "clade") into subsequent environmental modeling.

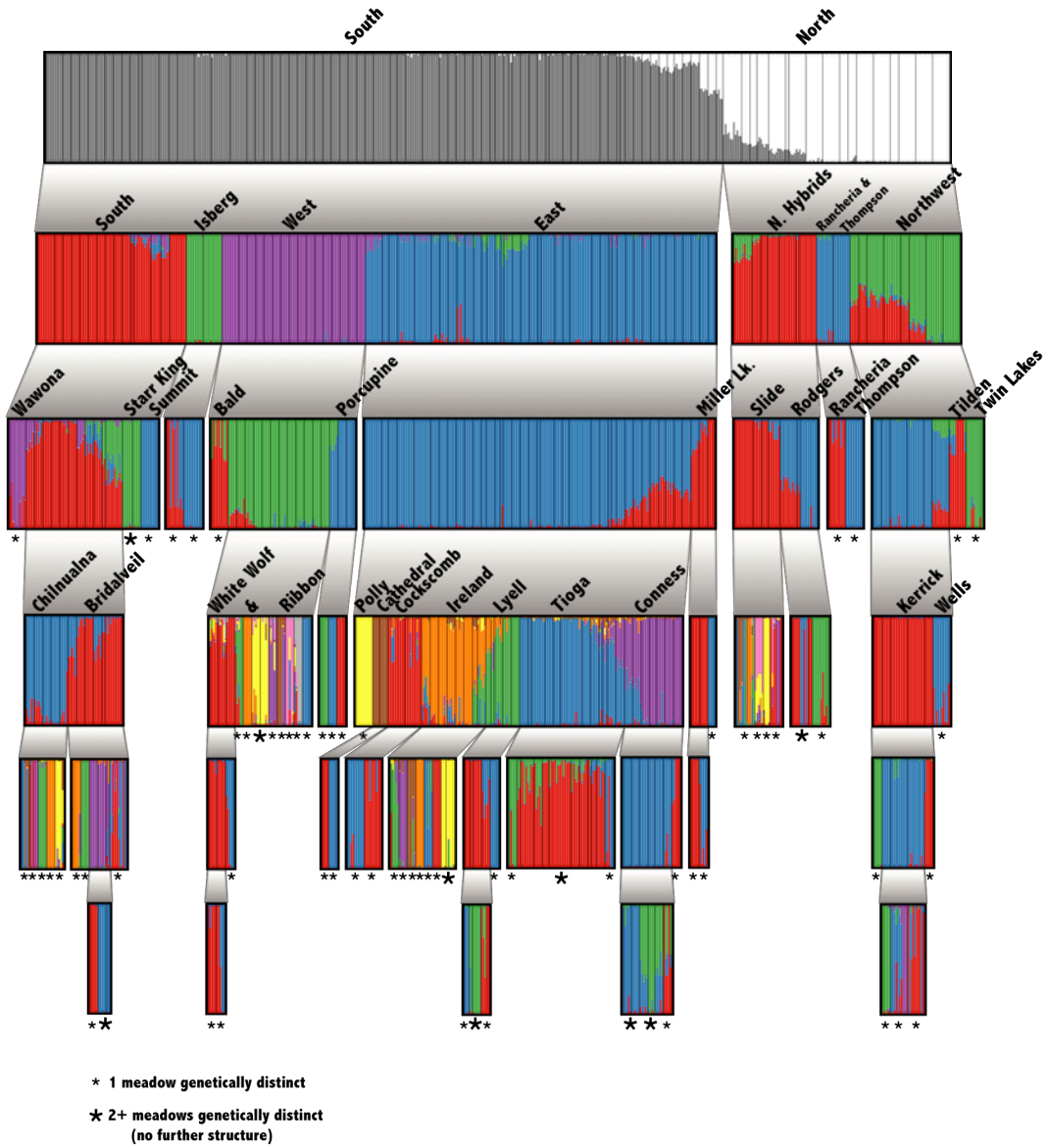


Figure 7. Hierarchical Genetic Structure in Yosemite NP
Barplots show hierarchical genetic structure, revealing that most meadows are distinct gene pools, but some are more distinct than others. Thin, black, vertical bars indicate meadow boundaries. All 90 meadows from Yosemite are included, and those found to be distinct are labeled with a small (*), while groups of meadows with no further structure are labeled with a large (*). The program structure was run for K = 1-10, and the optimal K was chosen following Evanno (2005). The hypothesis that K=1 was tested using a paired Wilcoxon test. Each cluster was rerun until reaching the meadow level, or finding no further structure.

What Unit is a Population?

The hierarchical structure analysis revealed significant structure below the level of clades, with most meadows being genetically distinct populations (Fig. 7). Neighboring meadows that were not distinct later showed strong metapopulation structure (discussed below), including the Tioga Pass, Conness, Ireland Lake, and White Wolf areas. Some meadows and regions were found to be more genetically distinct than others (Fig. 7). Two regions with admixture between clades (contact zones) are highly distinct: Isberg Pass (admixture between Y-South and Y-East), and the northern admixture area from Miller Lake, to Rodgers Lake, extending along Slide Canyon to Mule Pass. Rancheria Creek, a region without admixture, is also notably distinct from other regions in northwestern Yosemite. Below this tier, other regions are distinct to lesser degrees: Wawona Dome, Starr King Meadows, and Summit Meadow (Y-South); Bald Mountain and Porcupine Flat (Y-West); and Twin Lakes, Tilden Lake, and Wells Peak (Y-North).

A strong pattern of isolation by distance (IBD) in our dataset up to at least 30km distances (Fig. 8) suggested the possibility that groups of nearby meadows, in addition to meadows themselves, might be biologically distinct. Our pattern of IBD was consistent with stepping-stone migration and hierarchical genetic structure using allozyme and mitochondrial DNA previously found by Shaffer & Fellers (2000). Based on the recent USGS census of Yosemite toad breeding locations, that study contained large sampling gaps, and our spatially intensive sampling allowed us to analyze hierarchical structure at a finer scale. The 90 meadows sampled in the present study represent all watersheds where extant Yosemite toads are found, with multiple meadows from each watershed. In order to assess hierarchical structure at a fine scale, we used the results of ADMIXTURE analysis to delineate units of genetic cohesion intermediate between meadow (ecologically delineated units) and clade (phylogenetically delineated units). The lowest ADMIXTURE cross validation (CV) error was found at K=28 clusters, although subsequent K values elicited only gradual increases in CV error. Hence, meadows were assigned to clusters for the purpose of testing intermediate genetic structure, but

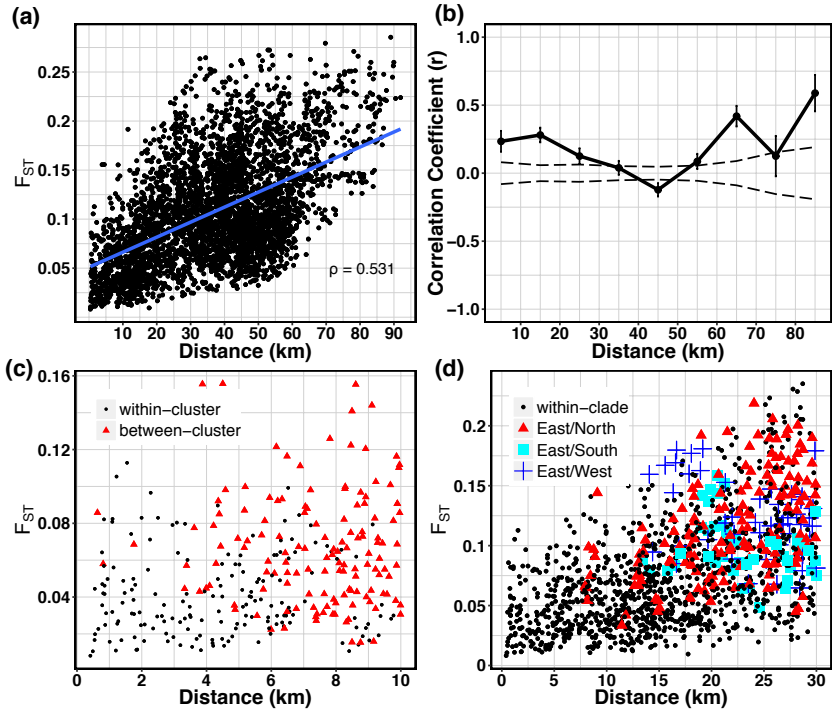


Figure 8. Isolation By Distance Isolation by distance (IBD) plots using least cost path-derived, topographically corrected distances. (a) Scatterplot of all pairwise comparisons with mantel correlation of ρ = 0.531 with p < 0.01. (b) Mantel correlogram showing mantel correlation coefficients as significant up to at least 30km. Solid lines are actual correlations, stippled lines represent confidence intervals for null distribution, and error bars represent bootstrapped confidence intervals for correlation. (c) IBD plot <10km categorized by within- and between-cluster comparisons. Clusters refer to ADMIXTURE clusters. (d) IBD plot <30km categorized by within- and between-clade comparisons.

the 28 ADMIXTURE clusters were not deemed exact biological units (Fig. 9). Table S2 details names of clusters, and average summary statistics at this scale. Most meadows (>50%) were contained in clusters with 2-6 meadows, but some (e.g. Tioga) had as many as 12. Out of 34 clusters, 16 (47%) contained only one meadow, and the mean cluster size was 3.00 meadows. Many singular clusters were in terrain separated by canyons, such as Rancheria, Thompson, Wells, Tilden, and Twin Lakes in Y-North (Fig. 9). In other cases, including Wawona and Summit in Y-South, and Bald and Ribbon 2 in Y-West, and to an extent Cockscomb, Polly, and Tressider in Y-East, topography appears to be a minor barrier.

We tested relative distinctiveness of clades, clusters, meadows, and individuals using a hierarchical AMOVA, which hierarchically partitions variation in population structure (Table 4). All Φ-statistics were significant based on 100 permutations with a two-tailed randomization test. As expected, a large (83.15%) proportion of variance was within individuals ($\Phi_{HI} = 0.17$), due to heterozygosity of haplotypes at many loci. Meadows within clusters showed the next highest variance of 15.33% ($\Phi_{PS} = 0.16$), supporting the idea that meadows structure current populations more than clusters or clades. Patterns of conformity to Hardy Weinberg expectations at the meadow and cluster scales further supported meadow distinctiveness, where clusters showed only 7.9% conformity across all loci and groups, but meadows showed 77.8% conformity.

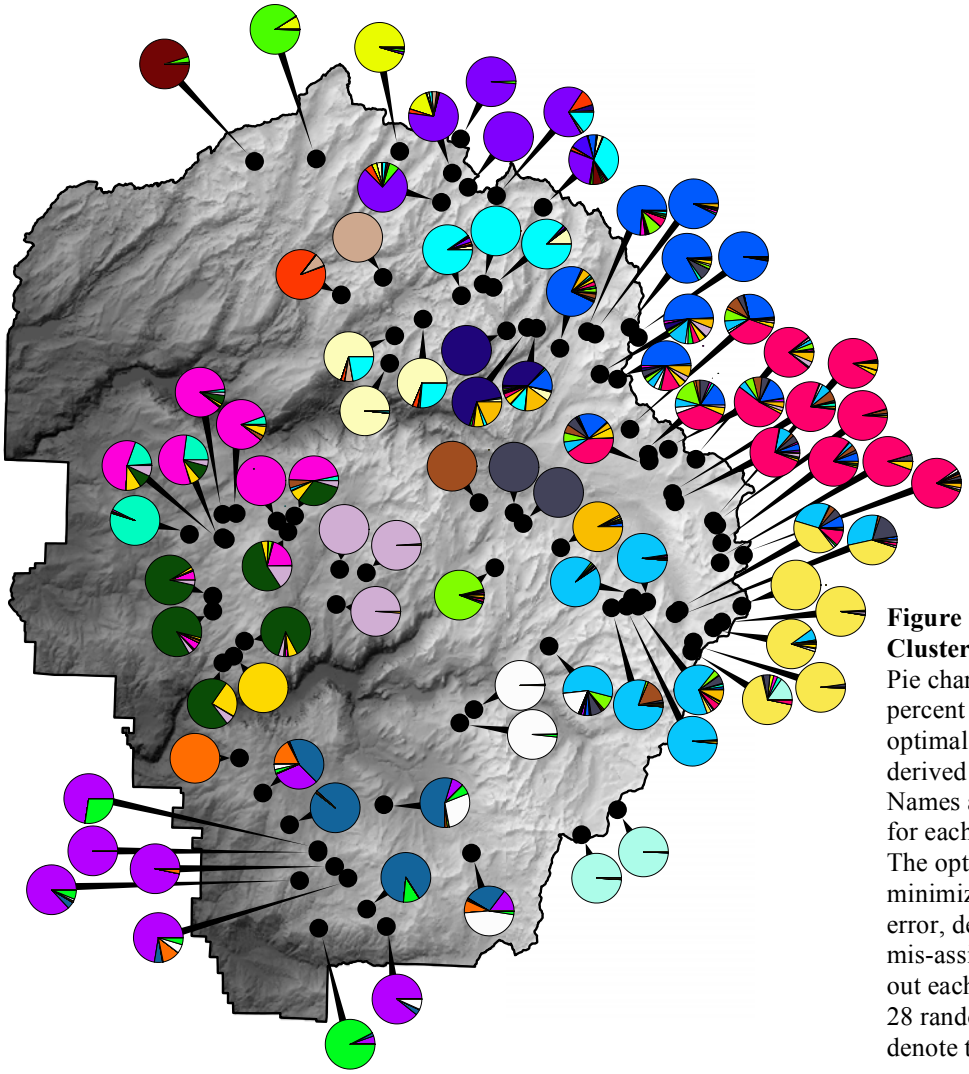


Figure 9. ADMIXTURE Genetic Clustering in Yosemite NP
Pie charts denote meadows with percent ancestry from an optimal K=28 gene pools derived from ADMIXTURE. Names and summary statistics for each cluster are in Table S2. The optimal K was chosen by minimizing cross-validation error, defined as the percent mis-assignment when holding out each 10% of the genotypes. 28 randomly chosen colors denote the 28 clusters.

Table 4. Analysis of Molecular Variance (AMOVA) results based on four hierarchical levels: individuals, meadows, ADMIXTURE clusters, and clades. Significance based on 100 random permutations with two-tailed test. B/T = between, W/I = within, df = degrees of freedom, SS = sums of squares, MS = mean squares, σ = variance, % Total = proportion of total of variance, Φ-Statistic = ratio of variances, and p = p-value for statistic.

Variance Component	df	SS	MS	σ	% Total	Φ-Statistic	P
B/T clades	3	11115.190	3705.064	9.936	3.379	Φ_{RT} = 0.0338	0.01
B/T clusters W/I clades	26	30046.330	1155.628	7.792	2.650	Φ_{SR} = 0.0274	0.01
B/T meadows W/I clusters	60	41222.250	687.038	45.063	15.326	Φ_{PS} = 0.1630	0.01
B/T individuals W/I meadows	445	97012.260	218.005	-13.234	-4.501	Φ_{IP} = -0.0572	<0.01
W/I individuals	535	130793.500	244.474	244.474	83.146	Φ_{HI} = 0.1690	0.01
Total	1069	310189.540	290.168	294.030	100.000		

Interestingly, individuals within meadows showed significantly higher heterozygosity than expected (-4.50% variance, Φ_{IP} = -0.06), and this pattern was further supported by high meadow-wide F_{IS} (Tables 3, S2, S3). Elevated meadow heterozygosity could be caused by sampling bias, due to family groups mating non-randomly over space and time within meadows. If family group structure and IBD (Fig. 8) are strong, it is conceivable that no single unit (i.e. meadows, clusters, or clades) perfectly describes current gene pools. Therefore, we chose to model meadows as populations because (a) they most closely approximated conformity to Hardy-Weinberg, and (b) this facilitated comparison to previous genetic and ecological studies conducted at this scale (Sherman & Morton 1993; Drost & Fellers 1994; Liang 2010; Wang 2012; Berlow et al. 2013; Brown & Olsen 2013; Ostoja et al. 2015).

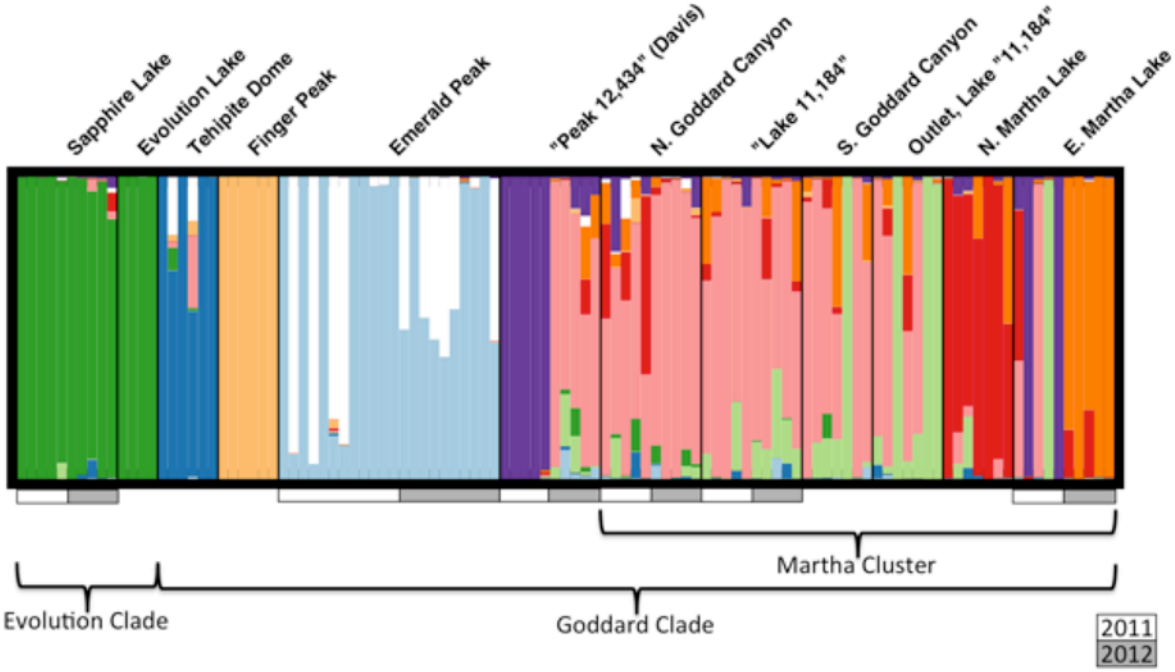


Figure 10. Spatial and Temporal Structure in Kings Canyon NP
STRUCTURE analysis with optimal K = 10 for Kings Canyon National Park from 12 sampled meadows, 6 of which were sampled during both 2011 and 2012. These were extremely wet and dry years, respectively. The white and grey boxes highlight those 3/6 meadows that show strong temporal structure. This suggests that these meadows contain 2+ populations or family groups that breed on alternating years, and do not freely interbreed.

Finally, Kings Canyon showed strong differentiation between Evolution and Goddard Canyons (Figs. 4, 10) with stark differences in diversity as described above. Within the Goddard clade, some meadows in the Martha Lake cluster were highly admixed. Structure within meadows and between years (2011 and 2012) was found in 3/6 meadows examined (Fig. 10). This suggests that in some meadows, gene flow between toads is limited. This could be due to assortative mating of family groups that breed in different years or conditions, or because toads have recently colonized these meadows from elsewhere. The pattern at "Peak 12,434" suggests the former, while East Martha Lake and Emerald Peak suggest the latter is possible (Fig. 10).

Meadow, Landscape, and Climate Effects on Connectivity

The present study combined elements of causal modeling, least cost transect analysis (LCTA), and a modified version of gravity modeling into a single workflow that significantly improved the performance compared to gravity modeling alone. The three-step analysis first chose a plausible set of adult toad dispersal corridors between meadows by assuming that slope, vegetation, and ridgeline barriers limit dispersal. Next, it created 100-meter buffers around these dispersal corridors as well as the breeding locations within meadows, and extracted the average and standard deviation of environmental variables hypothesized to influence gene flow for adult toads. Finally, it modeled gene flow among meadows using a linear mixed-effects framework (similar to gravity modeling). This last step accommodated the idea that both the environment between breeding meadows experienced by adult toads, along with the meadow environment experienced by all life stages, might have some impact on overall gene flow. In addition, since certain pairs of meadows reside in different phylogenetic lineages, they likely have different evolutionary histories. Our models accommodated the idea that the exact relationship between environment and gene flow will sometimes differ among lineages. We did this by including random slopes by clade whenever that inclusion significantly improved model fit.

We found that the least cost path model with 3:1 weight of slope/vegetation and ridgeline barriers had the highest significant partial mantel correlation with F_{ST} after accounting for Euclidean distance, and this model was buffered for variable extraction (Table 5). Landscape genetic models were constructed using 30km, 20km, and 10km cutoffs for least cost paths (Table 6), in order to represent processes occurring over different spatial scales (and temporal scales,

Table 5. Results of causal modeling using 11 resistance rasters and F_{ST}. Significant p-values from partial mantel tests after accounting for Euclidean distance are in bold. R=Ridges, S=Slope, V=Vegetation, E=Euclidean Distance. Slope and vegetation cost values were added together by weighting them either 1:1, 3:1, or 1:3, then adding impenetrable ridgelines. Mantel and partial mantel tests were performed with 1000 random permutations.

Model	Mantel ρ	p-value	Partial Mantel ρ	p-value
E	0.311	0.001	—	—
S	0.324	0.001	0.164	**0.021**
V	-0.070	0.723	-0.081	0.753
S+V (1:1)	0.314	0.001	0.042	0.306
S+V (3:1)	0.326	0.001	0.134	0.057
S+V (1:3)	0.317	0.001	0.061	0.230
R, S	0.332	0.001	0.138	0.081
R, V	0.309	0.001	0.041	0.310
R, S+V (1:1)	0.322	0.001	0.086	0.195
R, S+V (3:1)	0.341	0.001	0.176	**0.040**
R, S+V (1:3)	0.313	0.001	0.051	0.303

Table 6. Results of gravity modeling. Each of 6 models has the lowest AICc score for its category, where adding each effect gave a significant LR test score. The best model for each transect cutoff length (in km) is highlighted in bold. All variables fitted with a clade-specific random slope are in bold. All variables are listed in order of their inclusion during the forward-selection, and signs represent coefficient directions with respect to meadow connectivity ($1 - \sin^{-1}\sqrt{F_{ST}}$). BT variables are taken from buffered transects between sites, whereas AT variables are taken from buffers at sites. Variable abbreviations are as follows: **slope** = slope, **dem** = elevation, **runch** = 60-year change in runoff, **ndmi** = normalized moisture index, **trail** = count of trail pixels, **roadw** = count of road pixels weighted by traffic, **pcon** = % conifer pixels, **pwat** = % water pixels, **nbrm** = normalized burn ratio, **pflood** = % seasonally or intermittently flooded meadow, **petch** = 60-year change in potential evapotranspiration, **pptch** = 60-year change in precipitation, **breed** = network-boosted breeding occupancy of Berlow et al. (2013), **slope2** = standard deviation of 2nd derivative of slope. See Supplementary Table 1 for details about variables.

Dist.	Formula	Type	Rand. Slopes	-lnL	R^2_c	R^2_m	AICc	BIC
30	D - **BT(slope)** + **BT(dem)** - BT(runch) - **BT(ndmi)** - BT(trail) - BT(roadw) - **BT(pcon)** - BT(pwat) - BT(nbrm) - AT(pflood)	LCP		-2166.73	0.75	0.59	4361.62	4444.14
			✔	**-2096.76**	**0.81**	**0.58**	**4252.18**	**4422.79**
	D - **BT(slope)** + **BT(dem)** - **BT(runch)** - BT(nbrm) + BT(petch) - BT(trail) - BT(pcon) - BT(ndmi) - BT(roadw) - AT(pflood) - BT(pwat)	Euc.		-2334.41	0.72	0.55	4699.00	4787.40
			✔	-2198.96	0.86	0.47	4448.41	4595.56
20	D - BT(slope) + **BT(dem)** - BT(runch) - **BT(roadw)** - **BT(ndmi)** - BT(trail) - **BT(pwat)** - **BT(pcon)** - BT(nbrm) - AT(pflood) + AT(pptch)	LCP		-1627.01	0.67	0.43	3284.31	3365.64
			✔	**-1571.47**	**0.76**	**0.37**	**3216.53**	**3410.82**
	D - **BT(runch)** - BT(ndmi) - BT(nbrm) + **BT(dem)** - **BT(slope)** - **BT(trail)** - **BT(pcon)** + **BT(petch)** - BT(roadw) - AT(pflood) + AT(pptch)	Euc.		-1733.51	0.62	0.37	3497.30	3578.63
			✔	-1635.87	0.87	0.18	3360.03	3591.73
10	D - **BT(slope)** - AT(precip) - BT(runch) - BT(pwat) - BT(trail) - BT(roadw) - **BT(dem)** - AT(breed)	LCP		-567.11	0.63	0.41	1158.80	1209.94
			✔	**-549.52**	**0.78**	**0.41**	**1136.34**	**1212.63**
	D - **BT(ndmi)** - BT(runch) - BT(slope) - AT(pflood) - AT(breed) - AT(slope2)	Euc.		-598.44	0.58	0.31	1217.30	1259.99
			✔	-577.44	0.73	0.33	1181.56	1236.91

given that toad dispersal over space is limited by time). In all cases the best least cost path model was significantly preferred over the best Euclidean model, suggesting the validity of our method (Table 6). Slope between sites was the most important predictor of connectivity at all spatial scales, with lower slope causing higher connectivity (Table 6). Between-site elevation was the second-most important predictor at 30km and 20km, with higher elevation facilitating connectivity at these scales. However at 10km, elevation inhibited connectivity for Y-South. At other scales, elevation was least important for Y-South, and most important for Y-North. The 60-year increase in mean runoff and the proportion of water between sites inhibited connectivity at all scales, a counterintuitive result that is discussed below. In contrast, higher spatial variability in summer moisture index and May burn index decreased connectivity. Both roads and trails were found to be barriers to connectivity at all scales and across all clades, a result with important implications discussed below. Roads had clade-specific effects at 20km, where the two clades bisected by Highway 120 (Y-West and Y-East) were most and least impacted, respectively, while the other two clades with fewer road crossings had intermediate slopes. Finally, proportion of conifer cover between sites facilitated connectivity for the two low-elevation clades (Y-South, Y-West), and was inhibitory for the higher-elevation clades (Y-North, Y-East) at 20km and 30km.

Meadow-specific effects were all related to moisture in the best models. Moisture had the effect of reducing gene flow away from meadows, or genetic emigration. At 10km, higher inter-annual meadow precipitation (Daymet) was the second largest inhibitor of connectivity, after slope between sites. Additionally, meadows with a higher percent intermittently to permanently flooded area reduced connectivity at 20km and 30km. These meadows tended to be smaller, with smaller N_e, surrounded by coniferous forest, and at lower elevation. Therefore we hypothesized that these meadows were more isolated by intervening landscape, but with higher larval growth rate and recruitment from better hydrological conditions. A two-way ANOVA showed that tadpoles from these meadows were larger at every stage of development compared to meadows with lower-than-median precipitation (Fig. S4, $F = 9.098$, $df = 1,1651$, $p < 0.01$).

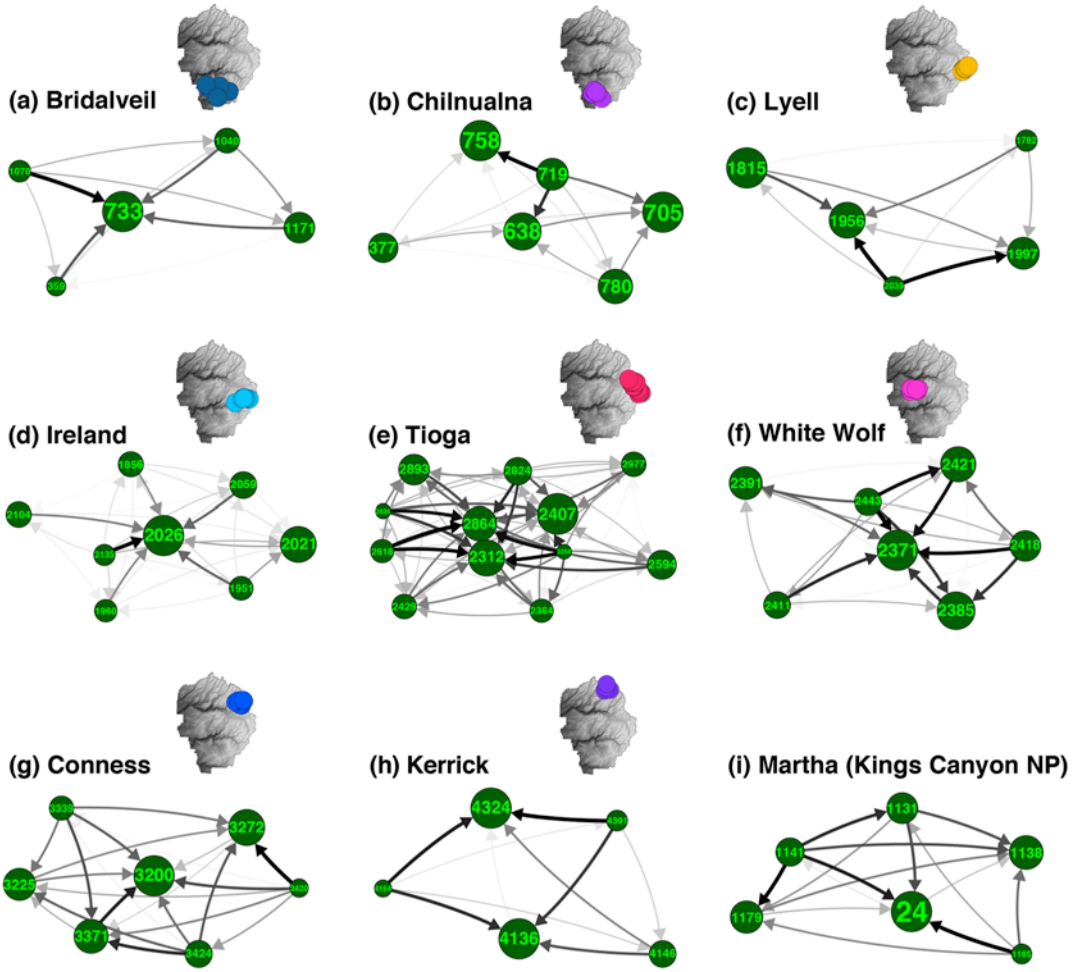


Figure 11. Metapopulation Dynamics
ADMIXTURE clusters are groups of closely related meadows with one or a few highly admixed "sinks" surrounded by more genetically distinct "sources". Graphs summarizing asymmetrical migration rates and directions for 8 meadow clusters in Yosemite NP and 1 in Kings Canyon NP are shown. Numbers are Meadow IDs. Clusters with at least 5 meadows were chosen for analysis. Arrow direction and darkness indicate the amount of difference between genetic emigration and immigration from a given meadow, calculated using the divMigrate tool following the method of Sundqvist (2016). For example, a dark arrow indicates strong asymmetrical migration in that direction, whereas a light arrow indicates slight asymmetrical migration in that direction. No arrow between two meadows indicates symmetrical migration. Node sizes are weighted by relative observed heterozygosity. Each cluster is indicated with inset map, with identical colors to those used in Fig. 9.

Metapopulation Dynamics

At the meadow scale, we calculated a "source index" by averaging all values of δG_{ST} within one ADMIXTURE cluster, where an overall negative value implies more genetic emigration than immigration. Fig. 11 shows source index for meadows in all 9 clusters containing >5 meadows to illustrate the strength of source-sink dynamics. Source-sink dynamics are strong in many clusters, which often have one or a few highly diverse meadows (high H_o) with much stronger genetic immigration than emigration. These can be thought of as genetic sinks, where incoming gene flow (and therefore admixture) is higher compared to neighboring meadows. Meadows that are spatially central appear to be larger and less rugged, so we hypothesized that larger, flatter meadows disproportionately act as sinks. Fig. 12 shows that sink meadows are indeed characterized by larger size and less topographic complexity around breeding sites. Log meadow area showed a significant negative Mantel correlation with δG_{ST} ($\rho = -0.540$, df = 201, $p < 0.001$), and source index was significantly negatively correlated with log meadow area ($\rho = -0.417$, df = 77, $p < 0.001$). Average slope within 500m of breeding area showed a significant

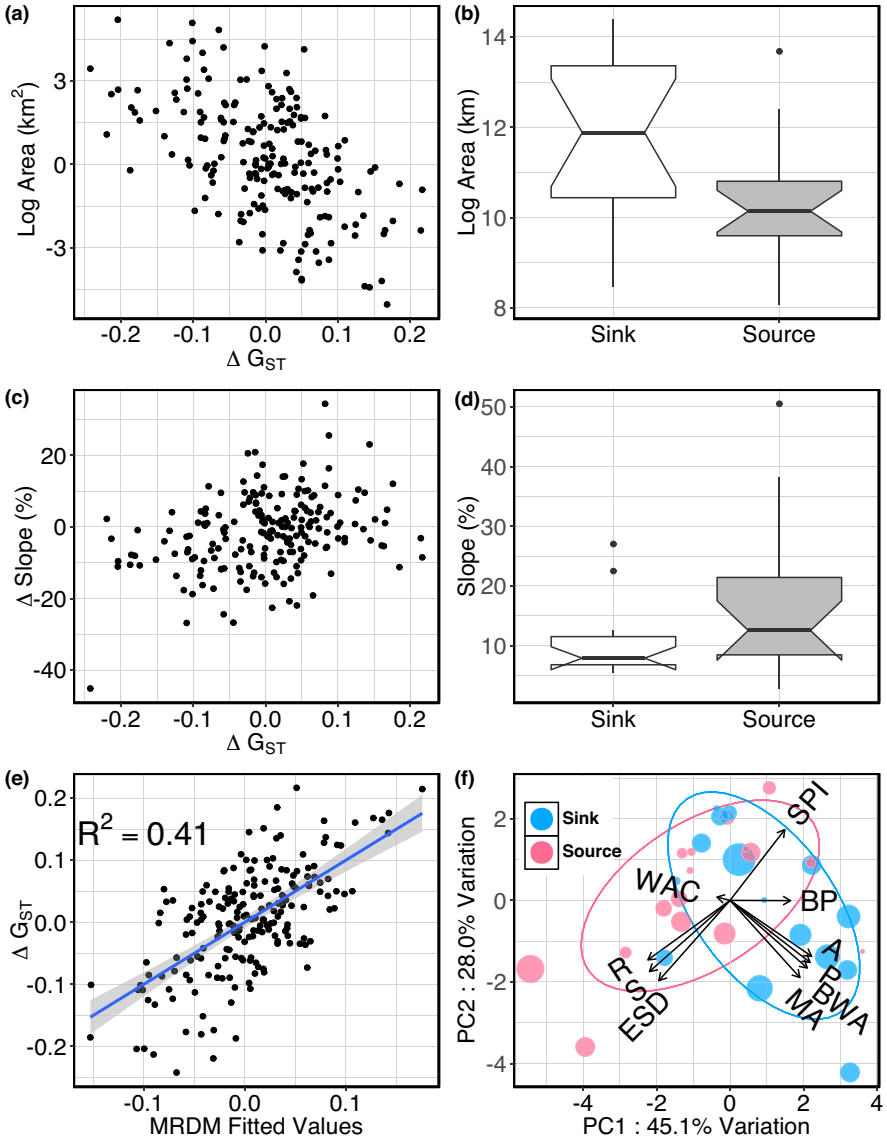


Figure 12. Meadow Attributes of Source and Sink Meadows Relationship between relative direction of gene flow (δG_{ST}) measured by the difference between "emigration" and "immigration" rates (following Sundqvist et al. 2016) and relative difference in pairwise meadow area (a) and slope (c). These pairwise differences are averaged per meadow (b, d) to show average meadow attributes for sources ($\delta G_{ST} > 0.05$) and sinks ($\delta G_{ST} < -0.05$). In (e) and (f) a multiple regression of distance matrices and principal component analysis show that sources and sinks are generally divided along two orthogonal axes: "meadow area" and "topographic roughness". A = log-area, P = perimeter length, BWA = breeding wet area within 1km^2 radius, S = % slope in 500m radius, R = roughness index in 500m radius, WAC = wet area of closest meadow, ESD = elevation standard deviation in 500m radius, MA = meadow area within 1km radius, BP = network-boosted breeding probability from (Berlow et al. 2013), SPI = slope position index in 500m radius.

positive Mantel correlation with δG_{ST} ($\rho = 0.319$, df = 201, $p < 0.001$), and source index was significantly positively correlated with slope ($\rho = 0.355$, df = 77, $p < 0.001$). A principal components analysis using 10 other variables related to meadow size and topographic complexity showed these two categories generally define source and sink meadows, albeit with regression of distance matrices ($F = 23.09$, df = 6,196, $p < 0.001$), with R^2 of 0.414, further supporting the hypothesis that the size and topographic smoothness of meadows relates to their status as sources and sinks.

Discussion

Are Meadows an Effective Unit for Population Management?

Identifying the unit at which a population is ecologically and evolutionarily independent is crucial to properly allocating conservation resources. Management units (MUs) are usually defined as demographically independent populations with sufficiently high growth rates and low connectivity to neighboring units (Palsboll et al. 2007). In practice, MUs are delineated based on genetic discontinuity or local adaptation to unique environmental conditions (Dizon et al. 1992). Moritz (1994) defined MUs as having significant nuclear or mitochondrial allele frequency divergence, representing present-day genetic or demographic cohesion. The category of MUs is used to reflect present-day population boundaries, and should be contrasted with Evolutionarily Significant Units (ESUs), which are lineages reflecting long-term or ancient isolation sufficient to create fixed genetic or morphological differences, and perhaps even reproductive isolation. There are many technological challenges involved in accurately delineating MU boundaries. For example, it has been acknowledged that different processes underlie demographic and genetic connectivity (Slatkin 1987; Bohonak 1999; Lowe & Allendorf 2010). The number of migrants ($N_e m$) is solely responsible for genetic connectivity, while demographic connectivity is influenced by m relative to birth and death rates. Additionally, the flow of gene copies and individuals can occur at dramatically different rates, a phenomenon known as Slatkin's Paradox (Bohonak 1999; Yu et al. 2010). Finally, many tests for population structure or differentiation implicitly assume equilibrium between the forces of genetic drift and gene flow. Many conditions likely violate this assumption in Yosemite toad populations, such as temporally fluctuating gene flow in metapopulations, recent colonization from a source population, or stepping stone migration preventing the formation of Hardy-Weinberg gene pools. For these reasons, we first identified 6 ESUs (4 in Yosemite, 2 in Kings Canyon), and then used a consilience of methods to pinpoint meadows as the likely scale of MU boundaries. We encourage future studies to complement our results with demographic (i.e. dispersal and growth rates) and adaptive (i.e. experimental translocations) evidence of MU boundaries.

A priori, we assumed that Yosemite toad meadows constitute the minimum possible scale for gene pools given their high vagility. Adult toads are known to disperse an average of 275m (maximum 1136m; Martin 2008; Liang 2010) outside breeding meadows in only weeks, and live 12-15 years (Sherman & Morton 1984). In contrast, the mean diameter of meadows occupied by toads in Yosemite is 265.6m. However, in Kings Canyon we encountered some evidence that complex temporal structure divides family groups within a meadow, perhaps due to assortative mating or inbreeding (Fig. 10). Since we measured this temporal structure across two consecutive years with extreme low and high levels of precipitation, it is difficult to speculate the degree to which climate influenced this structure. Additionally, we did not have adequate sample

sizes to thoroughly investigate sub-meadow structure in both parks; denser sampling across each meadow and multiple years is needed to pursue these hypotheses.

In this study, we first used a hierarchical STRUCTURE analysis to ascertain that meadows are typically distinct, with notable exceptions such as the Tioga and Goddard clusters (Fig. 7). Next we used ADMIXTURE to show that clusters of meadows form neighborhoods (Fig. 9). In order to assess which scale best reflects MU boundaries we performed a hierarchical AMOVA, which showed that meadows globally cause the largest reduction in heterozygosity and are the best designation for MUs (Table 4). Tests for Hardy-Weinberg at each scale also corroborated our conclusion that meadows account for the vast majority of population structure, with clusters playing a minor role. Finally, we showed that historical events have isolated Yosemite into 4 phylogeographic gene pools (ESUs) that we speculate have had limited gene flow since their origin (Fig. 3, 4, S1).

Genetic clustering algorithms can be sensitive to fine-scale divergence because they optimize Hardy-Weinberg simultaneously across many loci; however, small deviations from linkage or migration-drift equilibrium (as are expected in Yosemite toads due to isolation by distance) can result in substructure not being detected (Schwartz & McKelvey 2009; Lowe & Allendorf 2010. Our hierarchical clustering analysis has two major advantages to a more general Bayesian clustering method: (1) it overcomes the technical issue mentioned above, whereby stepping stone migration between meadows can prevent typical clustering analyses from detecting substructure, and (2) the results (Fig. 7) display the order of meadow distinctiveness, and thus allow unique meadows to be identified. Interestingly, many of the most distinct groups of meadows have mixed ancestry (i.e. from two ESUs), including meadows from the two major contact zones: 7 meadows near Slide Canyon, Rodgers Lake, and Mule Pass (Y-North / Y-East contact zone), 3 meadows near Miller Lake (Y-North / Y-East contact zone), and 2 meadows near Isberg Pass (Y-South / Y-East contact zone). The HIest results (Fig. 6) suggest these admixed areas have been admixing for multiple generations and are therefore stable, producing viable toads. As we discuss in more detail below, inter-lineage admixture is likely an important source of genetic novelty and unique haplotypes that can accelerate the pace of natural selection. Recombination between these lineages shuffles the deck of genetic information, which has two effects: (1) it creates a wider range of genetic (and hence trait) variation. Due to the recurrent nature of recombination over space and time, it (2) boosts the initial frequency of novel haplotypes. More extreme traits are possible (for example very high or low tadpole growth rates), and natural selection can more efficiently replace old traits with new, successful ones because they are already at high frequencies.

Admixed meadows are genetically unique because they are populations that exhibit unique and recombinant sets of genotypes. Other meadow clusters without any apparent admixture are also among the most distinct meadows in Yosemite. The following meadow clusters (containing 1 – 3 meadows each) should be considered substantially unique: Wawona, Starr King, and Summit (Y-South); Bald Mountain and Porcupine Flat (Y-West); and Rancheria, Thompson, Twin Lakes, Tilden, and Wells (Y-North) (see Table S2 and Fig. 9 for meadow IDs and locations). In contrast, recent colonization or high gene flow is likely obscuring significant meadow differentiation near Tioga, Conness, Ireland, and White Wolf. Two out of four of these undifferentiated clusters (Tioga and White Wolf) are adjacent to a major highway (CA-120), which has been identified as a barrier to gene flow (discussed below), and hence it is important for future work to determine whether gene flow and dispersal are high enough to make these areas single MUs. This could be approached in three possible ways: (1) these analyses can be

repeated with larger sample sizes using markers with higher mutation rates (e.g. microsatellites), (2) simulation, Bayesian, or approximate Bayesian methods can be used to estimate migration rates, or (3) a tagged genetic recapture study can be performed to estimate both dispersal and migration rates.

Phylogeography and Implications of Contact Zones for Inbreeding Rescue

We detected substantial phylogeographic structure based on RAxML, sPCA, and DAPC analyses, meaning that ancient divergence of gene pools continues to strongly structure the genetic data today into ESUs (Figs. 3, 4, S1). Using a published molecular clock for anuran nuclear DNA, we estimated that Y-North and Y-East + Y-West + Y-South originally diverged 1.51 years ago (Fig. S2). This fact has great importance for management of the species within Yosemite, because it means two ESUs (Y-North and Y-East) have been isolated for a substantial portion of the species' history, and subsequently have come back into contact. We discuss the implications of this fact below. It is important to note that using a rate calibration from other species assumes similar rates of evolution, and Yosemite toads might evolve at faster rates than anurans analyzed by Crawford (2003) if small population sizes interact with weak natural selection to accelerate DNA substitutions. Therefore we can only definitively interpret the relative (rather than absolute) divergence dates provided for the ultrametric tree in Fig. S2. With that caveat in mind, we briefly discuss the possible role of prehistoric climate in isolating Yosemite toads into the ESUs we observe today.

All divergence estimates are much older than the Tahoe (maximum at 70 ka), Tenaya, and Tioga (maximum at 21 ka) glaciations. Assuming accuracy of rate calibrations, this means these clades have either survived through glacial maxima in high-elevation nunatak or peripheral refugia (Holderegger & Thiel-Egenter 2009), or remained reproductively isolated while in lowland refugia. This latter scenario seems less likely if populations were in sympatry during glacial maxima, and given the apparent ability of clades to interbreed and produce admixed populations (Figs. 5, 6), and even hybrid populations with *A. boreas* (Mullally & Powell 1958). Also, the two largest divergences in Yosemite correspond to the Tuolumne and Merced River gorges (Fig. 3), a pattern that is more consistent with fragmentation that originated and continued to exist in Yosemite. In contrast to refugia theory (Haffer 1969), which predicts that cold-adapted species should expand their ranges during ice ages, some amphibians have remained inside their current distributions. For example, phylogeographic and IBD analyses showed that *Desmognathus wrighti* remained fragmented at high elevation during the Pleistocene, and this restriction was likely driven by ecological interactions (Crespi et al. 2003). It remains unclear where and why Yosemite toads retained distinct populations during the Pleistocene, and this is a question better answered with phylogeography at a broader spatial scale. However it is interesting to note that *Rana sierrae* show exactly the same spatial extent of their three major gene pools in Yosemite (Knapp, pers. comm.), and hence common prehistoric barriers might have structured each of the two species. One possible hypothesis for the formation of Y-South and Y-West is shown in Fig. 13, where these two lower elevation clades colonized glacial refugia during a glacial maximum, in this case the Tioga glaciation from 19 ka. The Y-East and Y-North clades might have remained in high-elevation nunatak or peripheral refugia as we have discussed above, or sought lowland refugia as displayed in Fig. 13.

Our data strongly supports a monophyletic *Anaxyrus canorus* with all 8 *A. boreas halophilus* specimens from around California forming a monophyletic outgroup to the species. Since all published studies to date have found para- or polyphyly using mtDNA, this study is

unique in that it reconstructed *A. canorus* ancestral relationships using nuDNA, and supports a monophyletic *A. canorus* (Graybeal 1993, 1997; Shaffer & Fellers 2000; Stephens 2001; Goebel 2005; Goebel & Ranker 2009); however, previous unpublished work found the same pattern using three nuclear genes (Pauly, pers. comm.). Therefore, mitochondrial introgression by female-biased dispersal, asymmetrical population sizes, or environmental selection on sex-linked traits during rare hybridization events probably caused the previously observed topologies (McGuire et al. 2007; Fontenot et al. 2011; Pavlova et al. 2013). However it is important to fully elucidate ancestor-descendent relationships for the boreas group using high-coverage nuclear markers and full spatial coverage of *A. boreas, A. canorus, A. exsul,* and *A. nelsoni*.

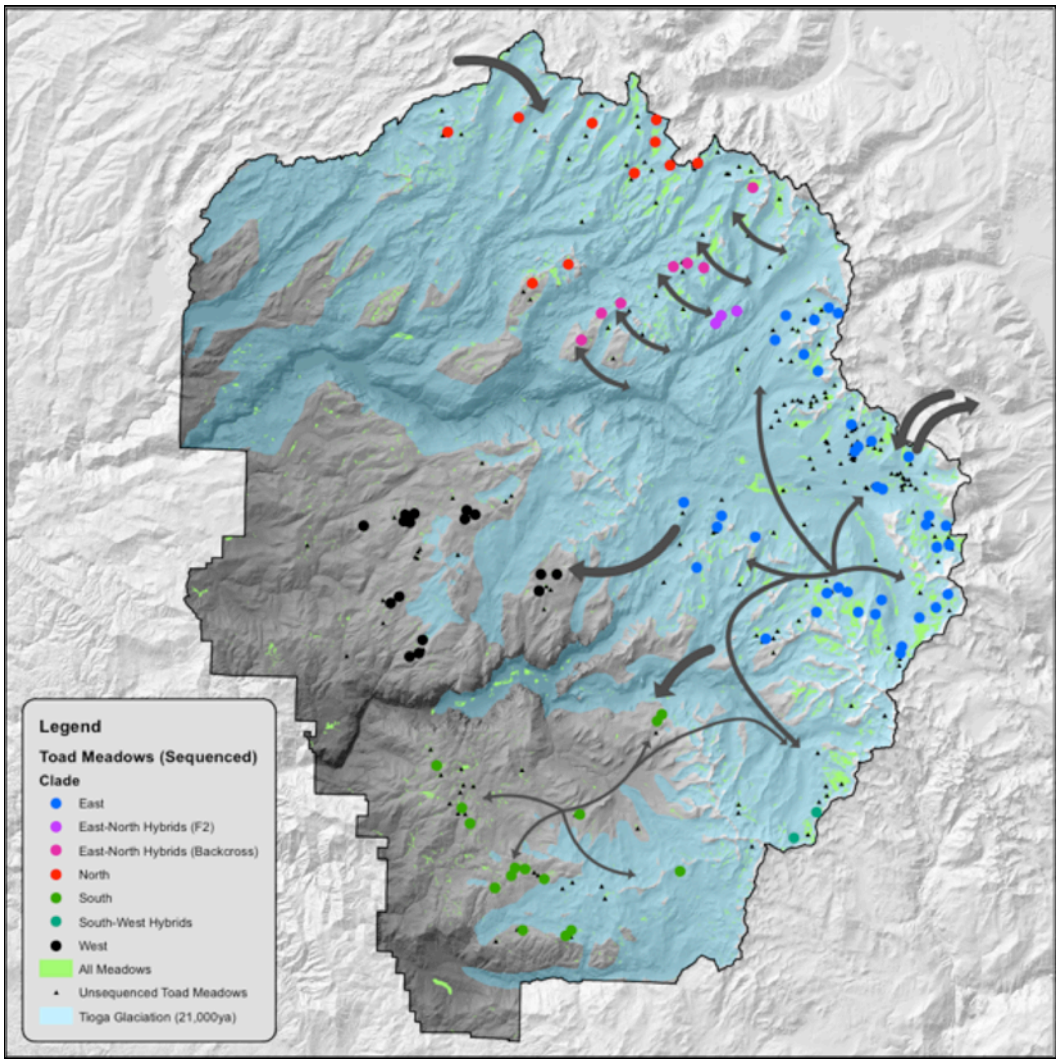


Figure 13. Historical Phylogeography in Yosemite NP
Glacial refugia at lower elevations during the last ice age may have sparked the origin of at least two distinct clades in Yosemite, Y-West and Y-South. Bold arrows indicate hypothesized colonization of the four major clades historically, including divergence of Y-West from Y-East, and the earlier divergence of Y-South from (Y-East + Y-West). Thin arrows represent recent movement, including range expansion supported by the program rangeExpansion (Peter & Slatkin 2013) and admixture between Y-East and Y-North. In spite of occurring in glacial refugia, Y-West and Y-South are more genetically depauperate than the other clades, possibly due to persisting founder effects. The North-East divergence is not easily explained by any feature within Yosemite and may have occurred elsewhere. These hypotheses need to be tested in a formal phylogeographic framework to be assessed. The Tioga glaciation layer is provided courtesy of Yosemite NP.

The ancient Y-East/Y-North interlineage admixture zone of secondary contact presents an intriguing research opportunity for admixed fitness and conservation (Figs. 5, 6). 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 & Elam 1993; Weber & 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 different view of natural hybridization posits it is common (Mallet 2005), and this new source of genetic variation might be important irrespective of fitness in F1 and other early filial generations (Arnold et al. 1999). Antipathy toward hybridization probably arose when the Biological Species Concepts downplayed its importance in the modern synthesis, and was crystallized by the "Hybrid Policy" of the US Endangered Species Act of 1973 (Mayr 1966; O'Brien & Mayr 1991; Haig & Allendorf 2006). In fact, adaptive introgression could be an essential source of novelty in species such as the Yosemite toad, which has an average N_e of 19.52 (Table S2). Preliminary analysis of Yosemite toad evolutionary rates have shown a strong negative relationship with census population size, indicating genetic drift might overpower adaptation from standing variation and new mutations (Maier, unpublished data). The high amount of recombination in introgressed individuals imposes a constraint upon polygenic traits, however the waiting time to fixation for recurrently arising (n) adaptive hybrid alleles is potentially far lower ($1/n\mu$) than for new mutations ($1/\mu$) (Hedrick 2013). In the present study, three meadows near Miller Lake contained F2+ individuals across two sampling years (Figs. 5, 6). We observed parental genotypes there but no F1s, possibly indicating assortative mating between F1 admixed individuals. In addition, backcrosses between F1 individuals and Y-North parental genotypes were abundantly found northwest of Miller Lake, while virtually no Y-East backcrosses were sampled, possibly indicating directional introgression. This could either be due to asymmetrical admixed fitness, or differences in population sizes. Thus, phenotypic differences across the Y-East/Y-North admixture zone can be studied for phenotypic differences and transgressive segregation that could make certain introgressed individuals superior stock for translocation efforts (Hamilton & Miller 2015).

Natural and Anthropogenic Limitations on Yosemite Toad Connectivity

Slope was ostensibly the most important landscape predictor of genetic connectivity regardless of spatial scale, and at transect cutoffs > 10km, transects with low slope at high elevation were least genetically structured (Table 6). Wang (2012) found a similar pattern using variance partitioning on microsatellite markers. Interestingly, clade-specific coefficients suggested high elevation dispersal increases connectivity most in Y-North, where isolation by distance is highest, and meadows are most isolated from each other. This region of Yosemite is defined by a series of glacially carved canyons with steep headwalls that toads likely traverse only after extensive horizontal and vertical dispersal (Fig. 2). Possibly, Yosemite toads in Y-North are behaviorally biased toward upland dispersal that maximizes the likelihood of encountering suitably wet, subalpine breeding meadows. Alternatively, during longer excursions between canyons that inevitably involve lower-elevation dispersal, toads might be more likely to encounter and breed in marginal non-meadow habitat (e.g. springs, stream channels, lakes, or

bogs). Breeding in such inter-meadow habitat would result in lower meadow-meadow migration rates, and decrease the probability of toads reaching any particular meadow destination. This hypothesis is supported by the fact that wetter conditions between meadows (60-year increase in mean runoff, and proportion of water between sites) inhibited connectivity across all models, which means that dispersal routes containing non-meadow breeding habitat might be important for toad "layovers." Several such observations of Yosemite toad larvae or subadults in streams and lakes have previously been made (Maier, Knapp, Lee, pers. obs.). We also found that higher variability in summer moisture (NDMI) and May burn-index (NBR2) decreased connectivity. Dispersal routes with unpredictable desiccation or fire regimes might increase the mortality of migrants.

Conifers increased connectivity among the low-elevation clades (Y-West and Y-South) which are dominated by red fir and lodgepole pine matrix habitat, and had the opposite effect for both high-elevation clades (Y-East and Y-North) whose meadows are primarily bordered by talus or sparsely-vegetated whitebark pine forest. This pattern is consistent with divergent environmental selection causing neutral genetic divergence, or "isolation by adaptation" (IBA; Nosil et al. 2008). In this case, the specific type of IBA is unique in that low- and high-elevation adult toads are presumably isolated not by different mating preferences or phenologies, but rather dispersal preferences. Theoretical models predict that divergent or heterogeneous environments should select for higher dispersal rates within the patch type that locally has a higher carrying capacity (McPeek & Holt 1992). If local adaptation of dispersal strategies occurs across a large enough region (e.g. Y-East), selection should overpower gene flow and prevent dispersal across an environmental gradient such as proportion of conifer forest (Gros et al. 2006). Another possible explanation for opposite signs in conifer random coefficients is that each clade resides almost entirely in conifer or subalpine habitat, and inter-clade phylogenetic signal is erroneously measured as conifer-mediated isolation by adaptation. This latter possibility seems unlikely given the patchwork of environmental heterogeneity within and among regions. For all of our non-stationary results that imply possible adaptations in dispersal by lineage (i.e. IBA for slope, elevation, conifers, etc.; Table 6), other explanations are possible. The most effective way to disentangle scenarios of isolation by distance, isolation by adaptation, and stochastic divergence such as "isolation by colonization" (IBC) is to compare genomic signals of divergence at neutral and non-neutral markers with morphological or behavioral biases in dispersal capacities (Orsini et al. 2013; Manthey & Moyle 2015).

Roads and trails were significant predictors of reduced Yosemite toad connectivity in all models. Tioga Road (CA-120) bisects two clades, Y-West and Y-East, and both regions include meadow metapopulations that cross this highway (Figs. 9, 11). Interestingly, the road apparently has a very large negative influence on Y-West connectivity, and a comparatively minor impact in Y-East. However, larger intrinsic growth rates, population sizes, and $N_e m$ in Y-East might mask considerable road mortality rates. Indeed, many Yosemite toad museum specimens were Y-East individuals collected on or near Tioga Road (VertNet 2016). The impact of roads and trails is unsurprising given the magnitude of recreation Yosemite experiences each year. In 2014, 3,882,642 people drove into Yosemite National Park, 190,339 of whom hiked on trails and camped in the backcountry (National Park Service 2015). This user load has gradually accumulated over more than a century. Automobiles began driving over Tioga Road in 1900, and the route became popular for businessmen beginning in the 1920s (Trexler 1975). The trail system was slowly built up from existing Native American trails during the 1860s by the U.S. Geological Survey of California, and later in the 1920s and 1930s by the Sierra Club and

National Park Service (Bingaman 1968). Future mark-recapture studies might reveal that these linear barriers also impact demographic connectivity.

Meadow Attributes Impact Network Connectivity

This landscape genetics study is among the first to isolate source-specific environmental effects on genetic connectivity (Murphy et al. 2010a; Dileo et al. 2014), and the first to accomplish this goal for Yosemite toads. We found two clear environmental patterns: one based on gravity modeling results, and the second based on metapopulation modeling.

First, we found that wetter (i.e. higher precipitation, percent semi- to permanently flooded area) meadows at lower elevation have reduced genetic emigration (Table 6) and, perhaps as a consequence, these meadows have lower heterozygosity and N_e. This effect was present at all scales, but especially strong for local (10km) transects in gravity models. At the same time, tadpoles in these montane meadows (shown in Fig. 1a, c) grow more slowly, allowing them to metamorphose at greater size (Fig. S4). Possibly, these tadpoles are less threatened by early desiccation, and have longer growing periods, better nutrition, or higher insolation, leading to larger body size at metamorphosis. This hypothesis is corroborated by the result that network-boosted meadow breeding probability (Berlow et al. 2013) has a slightly negative effect on connectivity at 10km. Additionally, more resources for fewer tadpoles could reduce larval resource competition, effectively increasing larval carrying capacity (K). Populations under carrying capacity generally produce fewer migrants (McPeek & Holt 1992), and possibly foster higher site-fidelity. The rate of emigration is primarily determined by the ratio of population size (N) to K (Dasmann 1981; Pflüger & Balkenhol 2014). Body size at metamorphosis is an important predictor of successful recruitment (Smith 1987; Altwegg & Reyer 2003), and more consistently high quality meadows might have selected for higher site-fidelity. The precise relationship between hydrology, elevation, and clade remains to be described. What is clear, Yosemite toads are much more isolated in montane meadows (Fig. 1a, c) than in subalpine meadows (Fig. 1b, d), and the overall effect is to lower genetic diversity (Fig. S4), and increase genetic structuring (Figure 7). This is in spite of any putative hydrological benefit to larval development at lower elevation.

Second, we found that larger and topographically flatter meadows tend to act as regional genetic sinks, and often these meadows are centrally positioned in the spatial arrangement of these metapopulations (Fig. 11, 12). This means that one or several meadows within each cluster (shown in Fig. 9) tend to receive a disproportionately large amount of genetic input, and send a disproportionately small amount of genetic output. They tended to be not only bigger and flatter, but have more wet area nearby, and be positioned within valleys rather than near ridgelines. At the smallest scale, it would seem these metapopulations are fueled by small, rugged, satellite meadows that create genetic admixture in a central area. The reason for this is difficult to speculate: perhaps these genetic sinks periodically reach carrying capacity and send out migrants to suboptimal meadows, or perhaps these central meadows have temporally fluctuating habitat quality. Since we have previously identified the pervasiveness of admixture on a larger scale (i.e. between clades), it is unsurprising to find admixture plays an important role at a smaller scale. This result comes with the caveat that many clusters have few meadows, and we only analyzed those with >5 meadows. Therefore there seem to be many examples of isolated meadows existing without a large, admixed meadow nearby.

Conclusions

In this study, we produced the first genome-wide Yosemite toad marker set of 2945 haplotypic loci sequenced for 535 individuals in Yosemite, and revealed that at least three clades exist in the park, the oldest of which represents a deep divergence from early in the species history. We also revealed a secondary contact zone of admixture between these two monophyletic clades with apparently stable F2 and backcrossed individuals. It is imperative to better understand the implications of interlineage admixture before attempting translocations, and we suggest the possibility that such a zone is a crucible for valuable adaptive diversity to combat stressors such as disease and drought. We found that most meadows are highly genetically divergent in accord with previous studies, but several regions are connected by either historical colonization or recent gene flow. We find that meadows are MUs, and the clades Y-North, Y-East, Y-West, Y-South, Goddard, and Evolution are ESUs, unless future evidence demonsrates genetic substructure within meadows, and/or absence of reproductive incompatibility between clades, respectively. Our environmental analysis affirmed the impact of slope on Yosemite toad connectivity, and found that other topographic (elevation), vegetation (conifers), hydrological (precipitation and meadow wetness), and disturbance (roads and trails) variables are important too. Tioga Road intersects two clades with highly interconnected meadows, and hence road mortality might be substantial enough to cause this observed pattern of genetic differentiation. Many of these relationships significantly vary by lineage implying a pattern of isolation by adaptation. We found some evidence that wetter, low-elevation meadows are less connected to each other in spite of larger tadpole growth, and we hypothesize that these toads are less driven by density-dependent dispersal. Finally, we found that metapopulation dynamics are strong on a local scale and characterized by larger, flatter meadows receiving more genetic input and sending less genetic output.

Summary and Future Directions:

- Meadows are distinct biological units that are isolated by distance, and overlaid with additional complex structure. Meadows fall into clusters of genetically similar meadows, and clades resulting from population isolation early in the species' history. Assortative family group mating may additionally drive structure below the meadow level.
- Meadows are separate Management Units (MUs). If reintroductions or translocations are used, toads from nearby meadows, preferably ones that are admixed ("genetic sinks"), will likely be most successful.
- Toads from different lineages are separate Evolutionarily Significant Units (ESUs) until further evidence suggests absence of lineage-specific adaptations and inter-lineage outbreeding depression.
- Since toads in the Y-East and Y-North groups are much more diverse than toads in Y-South and Y-West, genetic rescue by translocation has the potential to mitigate future population declines. Any future translocations would benefit from first testing for local adaptation to hydrological and climatic conditions within each ESU.
- Admixture between two highly distinct groups of toads is occurring across a large area. It is likely that their gene copies are moving far beyond this contact zone. Our work suggests these individuals are viable and highly diverse. We suggest the possibility that such a zone is a crucible for adaptive diversity to combat stressors such as disease and drought. Utilizing this natural contact zone for future research on whether certain

admixed toads have increased resistance to desiccation and disease would be beneficial.
- Tioga Road seems to increase mortality of migrant toads in a measureable way among toad metapopulations.
- Trail crossings between meadows also seem to increase migrant mortality in Yosemite. The most genetically depauperate clade of toads (Evolution Canyon in Kings Canyon) is adjacent to the most highly used trail in our study area.

Acknowledgements

This work was funded by a grant to Steven M. Ostoja, Eric L. Berlow, and Matthew L. Brooks. In addition the first author received funding from the Harold & June Memorial and Jordan D. Covin scholarships. The UC Merced Sierra Nevada Institute in Wawona, CA provided housing and accommodations for fieldwork. The Institute for Integrative Genome Biology (University of California, Riverside) provided access to sequencing and bioinformatics resources. The Museum of Vertebrate Zoology (University of California, Berkeley) and USGS San Diego Field Station donated specimens for outgroup analysis. Jared Grummer provided invaluable technical advice that improved library preparation for sequencing. Melanie Murphy and Jeffrey Evans provided R code that assisted in developing the methodology. 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 assisted in field specimen collection and wilderness companionship. All animal handling was performed in accordance with SDSU animal care and use protocol #13-03-001B. All data and sample collection was only completed after obtaining and adhering to the regulations of NPS research permits. Comments and suggestions by Jonathan Q. Richmond and Brian J. Halstead significantly improved the writing of this report. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data Accessibility

R code, Python code, and ddRADseq data can be made available upon request.

References

Alexander MA, Eischeid JK (2001) Climate variability in regions of amphibian declines. *Conservation Biology*, **15**, 930–942.

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **19**, 1655–1664.

Altwegg R, Reyer H-U (2003) Patterns of natural selection on size at metamorphosis in water frogs. *Evolution*, **57**, 872–882.

Anderson J (1979) A theoretical foundation for the gravity equation. *The American Economic Review*, **69**, 106–116.

Anderson CD, Epperson BK, Fortin M-J *et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, **19**, 3565–75.

Arnold ML, Bulger MR, Burke JM, Hempel AL (1999) Natural hybridization: how low can you go and still be important? *Ecology*, **80**, 371–281.

Baguette M, Van Dyck H (2007) Landscape connectivity and animal behavior: Functional grain as a key determinant for dispersal. *Landscape Ecology*, **22**, 1117–1129.

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Berlow EL, Knapp R, Ostoja SM *et al.* (2013) A network extension of species occupancy models in a patchy environment applied to the Yosemite toad (*Anaxyrus canorus*). *PLoS One*, **8**, e72200.

Bingaman JW (1968) *Pathways: A Story of Trails and Men*. End-Kian Publishing Company, Lodi, CA.

Bohonak A (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, **74**, 21–45.

Breiman L (2001) Random forests. *Machine Learning*, **45**, 5–32.

Brown C, Hayes MP, Green GA, Macfarlane DC, Lind AJ (2015) *Yosmemite toad conservation assessment*.

Brown C, Olsen AR (2013) Bioregional monitoring design and occupancy estimation for two Sierra Nevadan amphibian taxa. *Freshwater Science*, **32**, 675–691.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171–82.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–40.

Clobert J, Le Galliard JF, Cote J, Meylan S, Massot M (2009) Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters*, **12**, 197–209.

Crawford AJ (2003) Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, **57**, 636–641.

Crespi EJ, Rissler LJ, Browne RA (2003) Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Molecular Ecology*, **12**, 969–984.

Csárdi G, Nepusz T (2006) The igraph software package for complex network research. *InterJournal Complex Systems*, **1695**, 1695.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *American Naturalist*, **168**, 486–99.

Dasmann RF (1981) *Wildlife Biology*. John Wiley and Sons, New York, NY.

Dileo MF, Siu JC, Rhodes MK *et al.* (2014) The gravity of pollination: integrating at-site features into spatial analysis of contemporary pollen movement. *Molecular Ecology*, **23**, 3973–3982.

Dizon AE, Lockyer C, Perrin WF, Demaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. *Conservation Biology*, **6**, 24–36.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Drost C, Fellers G (1994) *Decline of frog species in the Yosemite section of the Sierra Nevada*. Davis, CA.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology Evolution and Systematics*, **24**, 217–242.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–20.

Fitzpatrick BM (2012) Estimating ancestry and heterozygosity of hybrids using molecular markers. *BMC Evolutionary Biology*, **12**, 131.

Fitzpatrick BM, Johnson JR, Kump DK *et al.* (2010) Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences of the USA*, **107**, 3606–10.

Fitzpatrick MC, Keller SR (2014) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, **18**, 1–16.

Flint LE, Flint AL (2012) Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, **1**, 1–15.

Flint LE, Flint AL, Thorne JH, Boynton R (2013) Fine-scale hydrologic modeling for regional landscape applications: the California basin characterization model development and performance. *Ecological Processes*, **2**, 1–21.

Fontenot BE, Makowsky R, Chippindale PT (2011) Nuclear-mitochondrial discordance and gene flow in a recent radiation of toads. *Molecular Phylogenetics and Evolution*, **59**, 66–80.

Funk WC, Blouin MS, Corn PS *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, **14**, 483–96.

Funk W, Tallmon D, Allendorf F (1999) Small effective population size in the long-toed salamander. *Molecular Ecology*, **8**, 1633–40.

Galpern P, Manseau M, Wilson P (2012) Grains of connectivity: analysis at multiple spatial scales in landscape genetics. *Molecular Ecology*, **21**, 3996–4009.

Goebel A (2005) Conservation systematics: the Bufo boreas species group. In: *Amphibian Declines: The conservation Status of United States species* , pp. 210–221. Univ of California Press.

Goebel A, Ranker T (2009) Mitochondrial DNA evolution in the *Anaxyrus boreas* species group. *Molecular Phylogenetics and Evolution*, **50**, 209–225.

Gosner K (1960) A simplified table for staging anuran embryos and larvae with notes on identification. *Herpetologica*, **16**, 183–190.

Grasso RL, Coleman RM, Davidson C (2010) Palatability and antipredator response of Yosemite toads (*Anaxyrus canorus*) to nonnative brook trout (*Salvelinus fontinalis*) in the Sierra Nevada Mountains of California. *Copeia*, **2010**, 457–462.

Graybeal A (1993) The phylogenetic utility of cytochrome b: lessons from Bufonid frogs. *Molecular Phylogenetics and Evolution*, **2**, 256–269.

Graybeal A (1997) Phylogenetic relationships of Bufonid frogs and tests of alternate macroevolutionary hypotheses characterizing their radiation. *Zoological Journal of the Linnean Society*, **119**, 297–338.

Gros A, Joachim Poethke H, Hovestadt T (2006) Evolution of local adaptations in dispersal strategies. *Oikos*, **114**, 544–552.

Haffer J (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.

Haig SM, Allendorf FW (2006) Hybrids and policy. In: *The Endangered Species Act at Thirty: Conserving Biodiversity in Human-Dominated Landscapes* (eds Scott JM, Goble DD, Davis FW), pp. 150–163.

Hamilton JA, Miller JMJ (2015) Adaptive introgression: a resource for management and genetic conservation in a changing climate. *Conservation Biology*, **00**, 1–29.

Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, **22**, 4606–4618.

Hohenlohe PA, Bassham S, Etter PD *et al.* (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Holderegger R, Thiel-Egenter C (2009) A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. *Journal of Biogeography*, **36**, 476–480.

Ivanova N V, Dewaard JR, Hebert P (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, **6**, 998–1002.

Jennings M, Hayes M (1994) *Amphibian and reptile species of special concern in California*.

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Dufour A, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.

Karlstrom E (1962) The toad genus *Bufo* in the Sierra Nevada of California: ecological and systematic relationships. , **38**, 38–39.

Keeler-Wolf T, Reyes ET, Menke JM, Johnson DN, Karavidas. DL (2012) *Yosemite National Park vegetation classification and mapping project report*.

Lemmon AR, Lemmon EM (2012) High-throughput identification of informative nuclear loci for shallow-scale phylogenetics and phylogeography. *Systematic Biology*, **61**, 745–61.

Liang CT (2010) Habitat modeling and movements of the Yosemite toad (*Anaxyrus* (=*Bufo*) *canorus*) in the Sierra Nevada, California. University of California, Davis.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, **19**, 3038–51.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

Manthey JD, Moyle RG (2015) Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: a landscape genomics approach. *Molecular Ecology*, **24**, 3628–3638.

Martin DL (2008) Decline, movement and habitat utilization of the Yosemite toad (*Bufo canorus*): an endangered anuran endemic to the Sierra Nevada of California. University of California, Santa Barbara.

Masek JG, Vermote EF, Saleous NE *et al.* (2006) A landsat surface reflectance dataset, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, **3**, 68–72.

Matchett JR, Stark PB, Ostoja SM *et al.* (2015) Detecting the influence of rare stressors on rare species in Yosemite National Park using a novel stratified permutation test. *Scientific Reports*, **5**, 1–12.

Mathieu J, Barot S, Blouin M *et al.* (2010) Habitat quality, conspecific density, and habitat pre-use affect the dispersal behaviour of two earthworm species, *Aporrectodea icterica* and *Dendrobaena veneta*, in a mesocosm experiment. *Soil Biology and Biochemistry*, **42**, 203–209.

Matthysen E (2005) Density-dependent dispersal in birds and mammals. *Ecography*, **28**, 403–416.

Mayr E (1966) *Animal Species and Evolution*. The Belknap Press of Harvard University Press, Cambridge, MA.

McGuire JA, Linkem CW, Koo MS *et al.* (2007) Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of Crotaphytid lizards. *Evolution*, **61**, 2879–97.

McPeek MA, Holt RD (1992) The evolution of dispersal in spatially and temporally varying environments. *American Naturalist*, **140**, 1010–1027.

Moritz C (1994) Defining "evolutionarily significant units" for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Mullally D, Cunningham J (1956) Aspects of the thermal ecology of the Yosemite toad. *Herpetologica*, **12**, 57–67.

Mullally D, Powell D (1958) The Yosemite toad: northern range extension and possible hybridization with the Western toad. *Herpetologica*, **14**, 31–33.

Murphy M, Dezzani R, Pilliod D, Storfer A (2010a) Landscape genetics of high mountain frog metapopulations. *Molecular Ecology*, **19**, 3634–49.

Murphy MA, Evans JS, Storfer A (2010b) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology*, **91**, 252–61.

National Park Service (2015) National park service visitor use statistics.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution*, **62**, 316–36.

O'Brien SJ, Mayr E (1991) Bureaucratic mischief: recognizing endangered species and subspecies. *Science*, **251**, 1187–1188.

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–99.

Ostoja SM, Matchett JR, McKenny H *et al.* (2015) Yosemite toad (*Anaxyrus* [*Bufo*] *canorus*) breeding occupancy of meadows in Yosemite and Kings Canyon National Parks, California. *Herpetological Review*.

Palsboll P, Berube M, Allendorf F (2007) Identification of management units using population genetic data. *Trends in Ecology & Evolution*, **22**, 11–16.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics*, **37**, 637–669.

Pavlova A, Amos JN, Joseph L *et al.* (2013) Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution*, **67**, 3412–28.

Peter BM, Slatkin M (2013) Detecting range expansions from genetic data. *Evolution*, **67**, 3274–3289.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model

species. *PLoS One*, **7**, e37135.

Pflüger FJ, Balkenhol N (2014) A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Molecular Ecology*, **23**, 2146–56.

Pounds J (2001) Climate and amphibian declines. *Nature*, **410**, 639–640.

Pounds JA, Bustamante MR, Coloma LA *et al.* (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, **439**, 161–167.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–59.

Pulliam HR (1988) Sources, sinks, and population regulation. *American Naturalist*, **132**, 652–661.

R Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ratliff R (1985) *Meadows in the Sierra Nevada of California: state of knowledge*. US Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station, Berkeley, CA.

Roche LM, Allen-Diaz B, Eastburn DJ, Tate KW (2012a) Cattle grazing and Yosemite toad (*Bufo canorus*, Camp) breeding habitat in Sierra Nevada meadows. *Rangeland Ecology & Management*, **65**, 56–65.

Roche LM, Latimer AM, Eastburn DJ, Tate KW (2012b) Cattle grazing and conservation of a meadow-dependent amphibian species in the Sierra Nevada. *PLoS One*, **7**, e35734.

Ryan ME, Johnson JR, Fitzpatrick BM (2009) Invasive hybrid tiger salamander genotypes impact native amphibians. *Proceedings of the National Academy of Sciences of the USA*, **106**, 11166–11171.

Sadinski W (2004) *Final report to the Yosemite fund*.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Shaffer H, Fellers G (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. *Molecular Ecology*, **9**, 245–257.

Sherman CK (1980) A comparison of the natural history and mating system of two anurans: Yosemite toads (*Bufo canorus*) and Black toads (*Bufo exsul*). University of Michigan.

Sherman CK, Morton ML (1984) The toad that stays on its toes. *Natural History*, **93**, 72–78.

Sherman C, Morton M (1993) Population declines of Yosemite toads in the eastern Sierra Nevada of California. *Journal of Herpetology*, **27**, 186–198.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Smith DC (1987) Adult recruitment in chorus frogs: effects of size and date at metamorphosis. *Ecology*, **68**, 344–350.

Sork VL, Davis FW, Westfall R *et al.* (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Molecular Ecology*, **19**, 3806–23.

Stephens M (2001) Phylogeography of the *Bufo boreas* (Anura, Bufonidae) species complex and the biogeography of California. Sonoma State University.

van Strien M (2013) Advances in landscape genetic methods and theory: lessons leart from insects in agricultural landscapes. Wageningen University.

van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, **21**, 4010–23.

Sundqvist L, Keenan K, Zackrisson M, Prodöhl P, Kleinhans D (2016) Directional genetic differentiation and relative migration. *Ecology and Evolution*, **6**, 3461–3475.

Trexler KA (1975) *The Tioga Road: A History, 1883-1961.* Yosemite Natural History Association.

US Fish & Wildlife Service (2014) *Endangered and threatened wildlife and plants; endangered status for the Sierra Nevada Yellow-legged frog and the northern distinct population segment of the Mountain Yellow-legged frog, and threatened status for the Yosemite toad: final rule*.

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

VertNet (2016) http://portal.vertnet.org/search?q=specificepithet:canorus+genus:(anaxyrus+OR+bufo)+mappable:1 (accessed on 2016-03-01)

Viers JH, Purdy SE, Peek RA *et al.* (2013) *Montane meadows in the Sierra Nevada: Changing hydroclimatic conditions and concepts for vulnerability assessment*. Davis, CA.

Wang IJ (2012) Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. *Diversity and Distributions*, **18**, 1033–1041.

Weber E, D'Antonio CM (2000) Conservation implications of invasion by plant hybridization. *Biological Invasions*, **2**, 207–217.

Weir BS (1996) *Genetic data analysis II: methods for discrete population genetic data*. Sinauer Associates, Inc., Sunderland, MA.

Yu H, Nason JD, Ge X, Zeng J (2010) Slatkin's Paradox: When direct observation and realized gene flow disagree. A case study in *Ficus*. *Molecular Ecology*, **19**, 4441–4453.

Appendix A: Supplementary Methods

Molecular Methods

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. Extracted DNA was eluted into 10 mM Tris-Cl, pH 8.0. We constructed ddRAD libraries following the protocol of Peterson et al. (2012; Protocol S1). Briefly, 200-500 ng DNA was digested with 5U SbfI-HF and MspI (New England Biolabs), for 3 hrs at 37°C and cleaned with 1.5x Ampure XP beads. Digested DNA was quantified by Qubit 2.0 Fluorometer (Life Technologies) and ligated to oligo-nucleotide adapters with one of 8 unique 5bp MID barcode sequences at 25°C for 30 minutes, followed by a 10 minute heat kill at 65°C. Ligated DNA was cleaned with Ampure XP and pooled by adapter, then size selected by Pippin Prep between 424 and 525bp using a 1.5% gel cassette (Sage Science). Optimal fragment size was chosen by in silico digestion of the *Xenopus tropicalis* v4.1 genome following the method of Lemmon & Lemmon (2012). Size-selected DNA was amplified with Illumina primers containing one of 12 unique indices using a Phusion PCR kit (New England Biolabs). The following cycle profile was used: 98°C 30 secs, [98°C 10 secs, 72°C 20 secs @ 16% ramp], 72°C 10 mins, 4°C hold. Finally, amplicons were bead-cleaned, quantified by BioAnalyzer (Agilent Technologies), and pooled in equimolar amounts for sequencing. This combinatorial approach allowed 8x12 unique samples to be sequenced in parallel on a single Illumina flowcell, and use of double-restricted fragments dramatically increased locus recovery across samples. All ddRAD libraries were 2x100bp sequenced on 7 lanes of an Illumina HiSeq 2500 at the Institute for Integrative Genome Biology, UC Riverside, CA.

Bioinformatic Data Processing

Raw data were filtered and processed using STACKS v1.19 (Catchen et al. 2011, 2013). Sequences were demultiplexed using process_radtags with a threshold of 1nt error in barcodes. Reads with an average phred score of <10 across a sliding window of 20% sequence length were discarded. Paired-end (PE) reads were non-overlapping by design, meaning that both ends of each DNA fragment were sequenced, but the intervening DNA (of variable length) was not sequenced. Therefore, two datasets were subsequently possible: (1) "combined" (PE sequence reads treated independently) and (2) "concatenated" (PE sequence reads concatenated). Concatenated reads are less ideal than contigs because the reads are nonoverlapping, but they still include information about allelic phase, and thus can increase the number of haplotypes at each locus. These two datasets were used for different downstream purposes owing to a tradeoff between haplotype diversity and number of loci; "combined" loci were mined for biallelic F_{ST} estimates and phylogenetic markers, while "concatenated" and "combined" loci were jointly used to analyze population structure and admixture. ustacks was used to identify alleles ("stacks") and subsequently call SNPs using a multinomial likelihood algorithm (Hohenlohe et al. 2010). Combined loci (100nt) were given a maximum stack distance of 3nt both within and between individuals, and this value was doubled for concatenated (200nt) loci. Each stack had a minimum coverage of 3 reads. Secondary stacks with 1-2 sequencing errors were used to increase power for SNP likelihood ratio tests. Loci with >3 stacks were either removed or deleveraged into multiple loci using a minimum spanning tree algorithm; this allowed paralogs and highly repetitive sequences to be either separated or discarded. Catalogs of consensus loci were constructed by cstacks using a subset (270/653 for combined, 126/653 for concatenated) of

individuals due to limited computing power. The program sstacks was used to match ustacks loci against the cstacks catalog to call genotypes. After applying a locus genotype coverage threshold of 10, loci were removed from meadows if absent from >25% individuals, and removed from the dataset if absent from >25% individuals overall. Finally, we built and implemented a custom python script (fasta2genotype.py, available upon request) to identify unique haplotypes. Population genetic summary statistics were computed using populations.pl in STACKS, along with Weir & Cockerham's F_{ST} estimates (Weir 1996). Mean pairwise F_{ST} values were calculated using bi-allelic SNPs with significant F_{ST} (based on Fisher's exact test $p<0.05$ on contingency table test of allele count differences). All processing was performed on a high-performance biocluster at the Institute for Integrative Genome Biology, UC Riverside, CA.

Genetic Structure

All markers in the haplotype dataset were initially tested for Hardy-Weinberg frequencies using the adegenet package (Jombart 2008) in R 3.2.1 (R Core Team 2014). NEESTIMATOR 2.01 was used to estimate N_e values using the linkage disequilibrium method (Do et al. 2014). A spatial PCA (sPCA; Jombart et al. 2008) was performed to reveal any cryptic phylogeographic discontinuities in the data by creating orthogonal synthetic variables that optimize the product of their variance and spatial autocorrelation (measured by Moran's I). Phylogenetic structure among an alignment of SNPs was assessed using the GTR + Γ nucleotide model in RAXML 8.1.11. This evolutionary model was chosen using JMODELTEST 2.1.4 to select the model with lowest AIC score with a significant likelihood ratio. We chose to use meadows as units to increase computational efficiency, and because the resulting ambiguity codes should only decrease signal/noise without introducing any systematic bias. A heuristic search was performed with RAXML for 10 independent reps 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. Putative admixed meadows (Y942, Y1097, Y1856, Y3342, Y3400, Y3414) based on the sPCA and maximum likelihood trees were removed and RAXML was rerun to bolster node confidence. We estimated dates of divergence using BEAST 2.3.2 on an alignment of sequences from 2–4 individuals in each major clade, with a coalescent tree model, relaxed log-normal clock, and calibration from an amphibian nuclear DNA clock of 9.24^{-10} to 1.53^{-9} substitutions/site/year, following Crawford (2003).

We used the ade4 package to check for significant meadow isolation by distance with a mantel test, and to perform a hierarchical AMOVA with 100 permutations for significance. Levels for the AMOVA included haplotypes, individuals, meadows, meadow clusters, and clades. Meadow clusters were derived by choosing one SNP per locus and running ADMIXTURE (Alexander et al. 2009) with 10-fold cross validation, then choosing the K with lowest cross-validation error. Each meadow was assigned to the cluster that contained the plurality of its ancestry. Clade assignments were designated using the find.clusters function DAPC in adegenet, which uses K-means clustering to find the K with lowest BIC score after 1e9 iterations at each possible value of K. Meadow differentiation was also tested using a hierarchical STRUCTURE analysis, whereby STRUCTURE 2.3.4 (Pritchard et al. 2000) was run 3x each at K=1-10 using 400,000 reps and 100,000 burn-in. If K was greater than 1 based on a paired Wilcoxon test, groupings were further analyzed in the same fashion until no further substructure was seen.

A large interlineage contact zone was detected within *A. canorus* and analyzed further. STRUCTURE was run 10x each at K=2 for individuals within the two clades, using 100,000 reps and 10,000 burn-in. Results were combined using CLUMPP 1.1.2 and visualized with DISTRUCT

1.1. IBD values across the putative zone were also assessed for a significant increase compared to background values. NEWHYBRIDS was used to estimate genotype categories for admixed individuals (i.e. those with genetic backgrounds from 2+ lineages). A maximum likelihood estimator in the R package HIest (Fitzpatrick 2012) was used to estimate whether admixed individuals were admixed beyond one generation (i.e. F3 – FN, instead of F2).

Least Cost Path Transects

Landscape genetic studies are designed to model environmentally mediated gene flow. Most use a statistical framework that falls into one of two categories: (1) resistance surfaces are parameterized with one or more environmental variable so that each pixel represents a hypothesized cost to dispersal, and subsequently least cost paths, least cost corridors, or circuit-theoretical currents are calculated. The total cost or length of least cost paths is subsequently regressed against some measure of genetic structure, and hypotheses can be evaluated by ranking model fit (e.g. causal modeling). (2) Rectilinear transects based on Euclidean lines between sites are used to collect the raw values of environmental variables. The influence of raw variables can then be individually modeled, e.g. using multiple regression of distance matrices, random forests, or gravity modeling. This latter approach can accommodate transect data between sites as well as source- or destination-specific data. Both approaches have merits and limitations: Cost surfaces directly model the spatial flow of gene copies and presumably model an organism's dispersal corridor better than rectilinear transects, however surface values must be subjectively defined based on the perceived impact of each environmental variable. Rectilinear transect methods avoid the subjectivity of cost surfaces, but might fail to capture relevant landscape attributes if dispersal corridors deviate drastically from straight-line transects. For example, Yosemite includes extremely rugged terrain and amphibians are unlikely to cross major ridgelines (Funk et al. 1999, 2005; Murphy et al. 2010b). Therefore Euclidean transects might poorly reflect environmental drivers of connectivity. We implemented the benefits of both methods in a two-step procedure: (1) simple least cost path models were constructed and buffered as described elsewhere (van Strien et al. 2012; 2013), then (2) environmental variables were individually extracted from least cost path transects for downstream variable selection and modeling

Simple least cost path models were constructed to reflect hypothesized dispersal routes based on 11 resistance rasters: slope, vegetation, 50% slope/vegetation, 75/25% slope/vegetation, 25/75% slope/vegetation, ridges + 50% slope/vegetation, ridges + 75/25% slope/vegetation, ridges + 25/75% slope/vegetation, ridges + slope, ridges + vegetation, and no resistance. Slope was modeled using a flat linear function where each 10% increment corresponded to values 1-10. Vegetation was modeled using a categorical function based on the Yosemite Vegetation Map, where each vegetation class was assigned a moisture index from 1-10. Ridges were given pixel values of 1^6 to encode them as impenetrable. Cost rasters were resampled to 30m for computational efficiency. Least cost path models and topographically corrected distances were calculated using the gdistance package and customized functions for multicore processing. The least cost path model with highest significant partial mantel correlation between least cost path distance and F_{ST} after accounting for Euclidean distance was buffered and used to extract all environmental data between all sampled meadows. Mantel tests were performed using the vegan package. All data were extracted from a least cost path buffer distance of 100m to minimize the influence of terrain outlying the hypothesized dispersal corridors, except for BCM climate data, which were extracted from a 500m buffer. BCM climate data were only available in >100m (270m) resolution and expected to have broader spatial effect. In order to assess the

improvement of least cost path transects, all between-meadow data were also collected from Euclidean buffers of identical bandwidth. In previous transect-based landscape genetic studies, bandwidth choice had minimal impact on overall model fit (Murphy et al. 2010a; van Strien et al. 2012), and therefore we chose intermediate bandwidths most likely to encapsulate dispersal without testing multiple bandwidths directly.

At-site environmental data were extracted from a 100m (most data) or 500m (BCM climate data) buffer around the mean sampling location for each USGS meadow sampled, except MODIS snowmelt and Daymet precipitation data which were extracted directly from meadow polygons. Source meadow data were used to model the possible influence of environmentally mediated larval recruitment on connectivity using gravity models as described below. All spatial data extractions were performed using the gdal module in a custom python script (ZonalStatistics.py, available upon request).

Environmental Data

Remotely sensed and USGS survey data were used to parameterize landscape genetic models. Between-site (267) and at-site (319) variables were initially chosen from 10 sources: BCM 2014 (climate and climate change), Shuttle Radar Topography Mission (topography), Yosemite Vegetation Map (vegetation and hydrology), MODIS (snowmelt timing), LANDSAT 5 (vegetation, moisture, and burn severity), Cal Fire (fire frequency), Daymet (precipitation), Cal Water (watershed attributes), Yosemite NPS (trails and roads), and USGS (meadow network attributes) (Table 1). LANDSAT surface reflectance images were corrected using the USGS Landsat climate data record (CDR), which corrects raw images for water vapor, ozone, shadows, and other atmospheric aberrations (Masek et al. 2006). Only images taken between the months of May and September were used from LANDSAT. Metrics of climate change were calculated as the difference between BCM mean or coefficient of variation (CV) during 1980-2010, and 1920-1950. Linear barriers such as streams, roads, and trails were rasterized, with each pixel representing presence or absence of the barrier. Additionally, we tried weighting barriers (roads were weighted by level of traffic, trails by amount of use, and streams by order). All data were sampled in their original resolutions. Resampling is necessary for creating combined resistance surfaces from multiple variables, although transect sampling extracts each variable independently without the need to combine rasters.

Two variables from each of 11 categories were chosen for the environmental models (Table S1), using random forests to rank variable importance. Random forests is a classification and regression tree method that estimate the importance and accuracy of an ensemble of decision trees (Breiman 2001). It can handle datasets containing many variables, their higher-order interactions, and multicollinearity when modeling a response variable. Variable importance was assessed using the "importance" parameter of the "randomForest" function from the randomForest package, which evaluates the mean decrease in accuracy and MSE when permuting each variable. Initially, phylogenetic clade (discussed below) was found to be the most important variable, because between-clade F_{ST} values were higher than within-clade F_{ST} values. To avoid phylogenetic bias, variable selection was rerun without transects that cross clade boundaries. Variables in each category were ranked by importance, and the best two with low multicollinearity were included, for a total of 22 variables. Multicollinearity was assessed using recursive backwards selection until variance inflation factor (VIF) was less than 10 for each one, using the fmsb package. During VIF variable selection, elevation at and between sites was included to reduce overall multicollinearity.

Gravity Modeling Analysis

Gravity models were initially constructed following the general form of Anderson (1979) and implemented following Murphy et al. (2010a) and Dileo et al. (2014). Models assumed the formula:

$$\boldsymbol{F_{ij} = m_i D_{ij}^{\alpha} R_{ij}^{\beta} S_i^{\gamma}} \quad \textbf{(1)}$$

where F_{ij} represents the flow of gene copies between sites and m_i is a site-specific random intercept. D_{ij}, R_{ij}, and S_i represent the effects of distance, environmental resistance between sites, and production effects of sites respectively, and α, β, and γ are parameters to be estimated. However, we chose not to log-transform the data for three reasons: (1) models fit the data poorly using this transformation and caused the residuals to strongly deviate from normality, (2) climate change variables included negative values, and (3) we wanted to fit random slopes as well as intercepts. Instead we fitted linear mixed models with the general form:

$$\boldsymbol{F_{ij} = m_i + \alpha D_{ij} + \beta R_{ij} + \gamma S_i} \quad \textbf{(2)}$$

An arcsine transformation for proportional data was applied to normalize the dependent variable, and the resultant value was subtracted from 1.0 to represent connectivity instead of differentiation ($1 - \sin^{-1}\sqrt{F_{ST}}$). Sites with <5 observations were removed to decrease the standard deviations associated with each random grouping level. All variables were centered and scaled by subtracting the mean and dividing by the standard deviation.

Linear mixed effects models were constructed with the lme4 package (Bates et al. 2015) using both the full environmental and climate-only datasets using three transect distance cutoffs for each: 10km, 20km, and 30km. These three cutoffs were chosen to represent processes occurring over different spatiotemporal scales. In addition, both Euclidean and least cost path transects were modeled to assess whether least cost path transect modeling substantially improves or changes models of dispersal-mediated gene flow. Stepwise addition of each variable was performed using AICc to rank models, and a variable was retained only if a likelihood ratio test between it and its nested model was significant. Finally, we tested the hypothesis that phylogenetic clades interact differentially with their environments by adding random slopes with the form:

$$\boldsymbol{F_{ij} = m_{ik} + \alpha D_{ij} + (\beta + \beta_k) R_{ijk} + (\gamma + \gamma_k) S_{ik}} \quad \textbf{(3)}$$

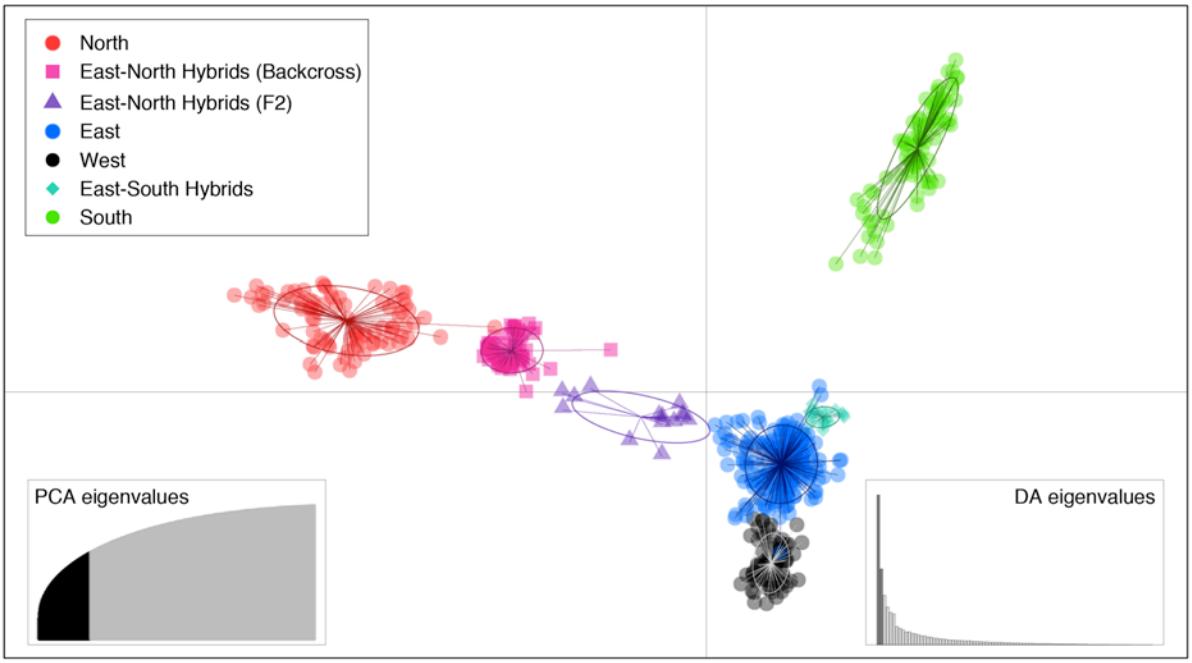
Variable random slopes by clade were added using the stepwise process described above. Marginal and conditional R^2 values were calculated using the MuMIn package.

Metapopulation Dynamics

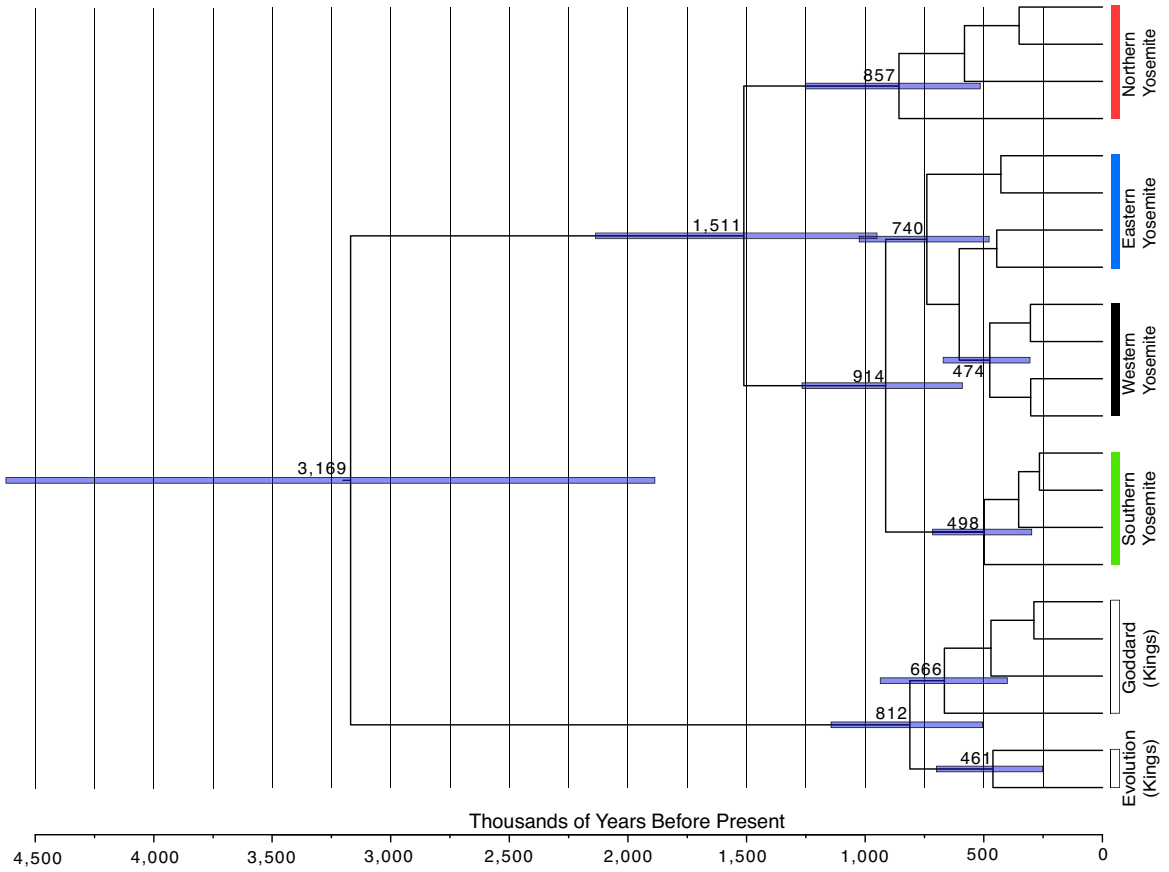
We tested the hypothesis that clusters of nearby meadows included strongly "source" and "sink" meadows using the method of Sundqvist et al. (2016). Specifically, we hypothesized that a few meadows would be sources of migrants for nearby satellite meadows. In essence, this method calculates geometric means of allele frequencies for a theoretical pool of migrants between each pair of meadows using alleles present in both meadows. Then asymmetrical G_{ST}

values are calculated by comparing each meadow to the migrant pool. We identified environmental correlates of source and sink meadows using mantel correlations, multiple regression of distance matrices, and principal components analysis. For each metapopulation we examined, estimates were mapped onto igraph (Csárdi & Nepusz 2006) objects to display the extent of source-sink dynamics in a network.

Appendix B: Supplementary Figures and Tables

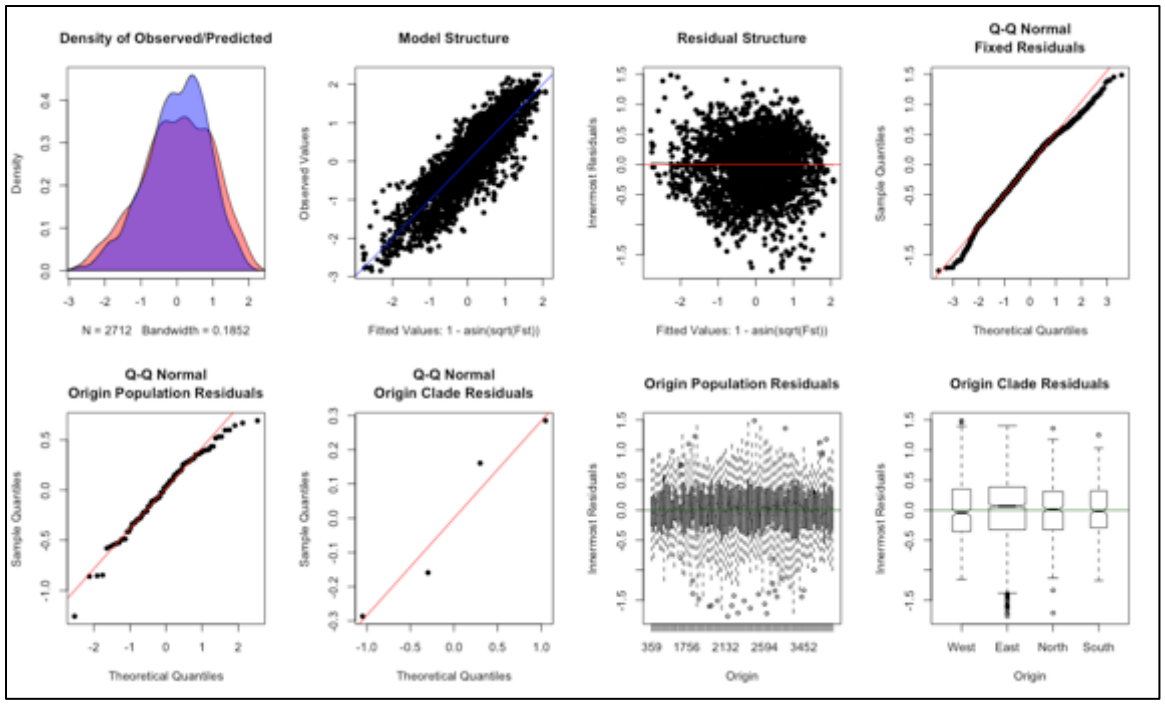


Supplementary Figure 1. Discriminant Analysis of Principal Components Biplot
The first two DA eigenvalues are shown, based on the first 100/535 PCs. Clade membership is determined by k-means clustering with 10^9 iterations using the find.clusters function in the adegenet package. Y-North, Y-East, Y-West, and Y-South clades based on previous figures are shown, with northern admixed meadows Y3400, Y3414, Y3342, Y4025, Y3273, Y3384, Y3452, Y3612, Y3615, Y3763 shown in purple.

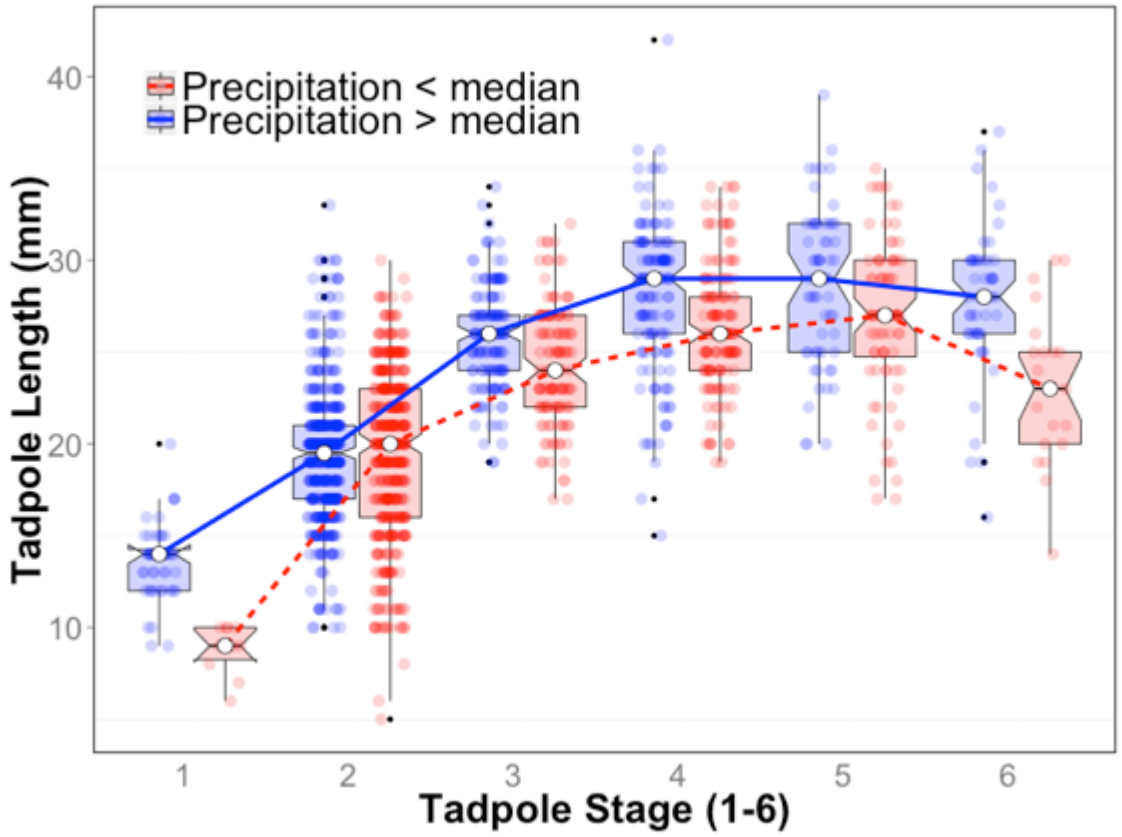


Supplementary Figure 2. Divergence Dating of Major Clades

Divergence date estimates for major clades in Yosemite and Kings Canyon, and the resulting BEAST ultrametric tree. 2-4 individuals from each RAxML clade are included to represent the range of genetic variation found within each lineage. Individuals from areas with known admixture were excluded. Ages for nodes are calibrated to an amphibian nuclear DNA clock of 9.24^{-10} to 1.53^{-9} substitutions/site/year, following Crawford (2003). Ages shown are median ages from posterior distribution. Bars represent 95% CI of node ages. All dates are listed in thousands of years before present.



Supplementary Figure 3. Model Performance
Model performance for best 30km model. (a) Distribution of observed (red) and predicted (blue) values. (b) Fit of observed values to predicted values on regression line. (c) Deviation of residuals from central predicted values. (d) Q-Q plot of fixed residual quantiles against normally distributed quantiles. (e) Q-Q plot of random intercept residual quantiles against normally distributed quantiles. (f) Q-Q plot of random slope residual quantiles against normally distributed quantiles. (g) Deviation of population residuals from central predicted values, by population. (h) Deviation of clade residuals from central predicted values, by clade.



Supplementary Figure 4. Yosemite Toad Growth Curves Differ by Precipitation Level
Comparison of growth curves among meadow-scaled Daymet precipitation levels, > and < median values. Tadpole field stages correspond to Gosner (1960) stages as follows: 1 (Gosner 1-25), 2 (Gosner 26-30), 3 (Gosner 31-35), 4 (Gosner 36-39), 5 (Gosner 41), 6 (Gosner 42-44). Tadpoles spend the majority of time as Field Stage 2.

Supplementary Table 1. Environmental variables chosen for gravity models. Two variables encompassing 2 locations (at or between sites) and 7 types were used, for a total of 22 variables from 11 categories. These included: topography (at/between), disturbance (at/between), vegetation/hydrology (at/between), hydrology/temperature/climate/climate change (at/between), LANDSAT vegetation/moisture/burn severity (between), meadow network attributes (at), and snow meltoff date (at).

Table Abbreviation	Location	Source	Type	Defintion
B(dem)	Between	SRTM	Topography	Mean of elevation values between sites
B(slope)	Between	SRTM	Topography	Sum of slope values between sites
B(trail)	Between	NPS	Disturbance	Sum of per-pixel trail crossings between sites
B(roadw)	Between	NPS	Disturbance	Sum of per-pixel road crossings between sites, weighted by traffic
B(pcon)	Between	Yosemite Vegetation Map	Vegetation, hydrology	Percent coniferous vegetation between sites
B(pwat)	Between	Yosemite Vegetation Map	Vegetation, hydrology	Percent standing water, permanent pool, or stream between sites
B(petch)	Between	BCM 2014	Hydrology, temperature, climate (change)	60-year change in CV of potential evapotranspiration, mean between sites
B(runch)	Between	BCM 2014	Hydrology, temperature, climate (change)	60-year change in mean of runoff, mean between sites
B(nbrm)	Between	LANDSAT 5	Vegetation, moisture, burn severity	Normalized Burn Ratio 2: interannual mean of May mean, standard deviation between sites
B(ndmi)	Between	LANDSAT 5	Vegetation, moisture, burn severity	ND Moisture Index: interannual mean of May-Sept mean, standard deviation between sites
A(netsim)	At	(Berlow et al. 2013)	Meadow network attributes	Network-boosted model probabililty of breeding at meadows
A(breed)	At	Yosemite Vegetation Map	Meadow network attributes	Distance to nearest breeding meadow at sites
A(roadd)	At	NPS	Disturbance	Distance to nearest road at sites
A(fired)	At	Cal Fire	Disturbance	Distance to nearest fire perimeter at sites
A(spi)	At	SRTM	Topography	Standard deviation of slope position index values at sites
A(slope2)	At	SRTM	Topography	Standard deviation of second derivative of slope values at sites
A(pcon)	At	Yosemite Vegetation Map	Vegetation, hydrology	Percent coniferous vegetation at sites
A(pflood)	At	Yosemite Vegetation Map	Vegetation, hydrology	Percent semi- to permanently flooded meadow vegetation at sites
A(p50sd)	At	MODIS	Snow meltoff date	Standard deviation of percent days (2002-2007) with >50% meadow snow cover at meadows
A(meltsd)	At	MODIS	Snow meltoff date	Standard deviation of meadow meltoff date (2002-2007) at meadows
A(pptch)	At	BCM 2014	Hydrology, temperature, climate (change)	60-year change in mean precipitation, mean at sites
A(precip)	At	Daymet	Hydrology, temperature, climate (change)	Inter-annual (1980-1997) mean of precipitation at meadows

Supplementary Table 2. Summary of population genetics by meadow. YOSE = Yosemite, KICA = Kings Canyon, PA = private alleles, N = sample size, P = percent polymorphism, H_o = observed heterozygosity, H_e = expected heterozygosity, Pi = nucleotide diversity, F_{IS} = inbreeding coefficient, Sites = number of sites, Var Sites = # of variable sites, Pol Sites = # of polymorphic sites, % Pol Loci = x. Ne is shown for each clade, cluster and meadow where available.

Clade	Cluster	MeadowID	Park	PA	N	P	Ho	He	Pi	Fis	Sites	Var Sites	Pol Sites	% Pol Loci	Clade Ne	Cluster Ne	Meadow Ne
Evolution	Sapphire	958	KICA	24	9.604	0.959	0.065	0.055	0.058	-0.014	602145	13796	2256	0.375	3.3	3.3	8.2
		902	KICA	32	3.659	0.972	0.046	0.035	0.041	-0.009	608741	14105	1259	0.207			-
Goddard	Tehipite	101	KICA	16	5.847	0.961	0.061	0.052	0.057	-0.006	596020	13740	2038	0.342	24.3	20.6	20.6
	Emerald	916	KICA	47	21.270	0.957	0.062	0.059	0.060	-0.002	640827	14686	2766	0.432		4.2	4.2
	Finger	1381	KICA	34	5.838	0.962	0.061	0.049	0.054	-0.015	622232	14348	1770	0.285		60.2	60.2
	Martha	1131	KICA	10	6.748	0.950	0.071	0.067	0.073	0.006	592882	13523	2789	0.470		53.7	-
		1138	KICA	26	9.435	0.949	0.072	0.069	0.073	0.005	619643	14208	3131	0.505			160.7
		1141	KICA	15	6.731	0.952	0.070	0.065	0.070	0.002	613355	13998	2723	0.444			47.9
		1165	KICA	12	6.492	0.957	0.068	0.057	0.062	-0.010	572522	13101	2208	0.386			-
		1179	KICA	12	9.650	0.950	0.071	0.068	0.072	0.004	635015	14626	3140	0.495			19.9
		24	KICA	56	9.639	0.947	0.076	0.073	0.077	0.006	629723	14339	3409	0.541			152
	Peak 12,434	1071	KICA	9	9.474	0.950	0.072	0.068	0.072	0.003	598925	13766	2902	0.485		9.3	9.3
South	Bridalveil	1040	YOSE	6	4.782	0.969	0.050	0.040	0.045	-0.009	619634	14045	1510	0.244	23.6	13.2	-
		1070	YOSE	3	1.996	0.968	0.054	0.038	0.052	-0.003	699008	15678	1417	0.203			-
		1171	YOSE	16	4.882	0.967	0.050	0.043	0.048	-0.004	707461	15935	1907	0.270			-
		359	YOSE	8	4.812	0.973	0.047	0.036	0.040	-0.014	699497	15798	1519	0.217			-
		733	YOSE	11	9.700	0.965	0.053	0.047	0.050	-0.007	715290	16212	2262	0.316			5.9
	Chilnualna	377	YOSE	3	4.926	0.971	0.051	0.037	0.042	-0.018	666530	15000	1500	0.225		15.2	-
		638	YOSE	1	4.896	0.969	0.048	0.040	0.045	-0.006	699848	15783	1753	0.251			-
		705	YOSE	5	4.921	0.968	0.051	0.041	0.046	-0.011	706184	15972	1721	0.244			27.3
		719	YOSE	2	4.870	0.970	0.054	0.038	0.042	-0.023	678108	15302	1430	0.211			-
		758	YOSE	1	4.880	0.969	0.052	0.041	0.046	-0.012	678000	15294	1673	0.247			-
		780	YOSE	4	4.849	0.970	0.053	0.039	0.044	-0.017	688960	15530	1655	0.240			-
	Wawona	387	YOSE	5	9.726	0.973	0.045	0.036	0.038	-0.014	697445	15691	1621	0.232		15.7	15.7
	Summit	1369	YOSE	5	9.739	0.971	0.053	0.037	0.039	-0.028	715850	16155	1551	0.217		-	-
	Starr King	1543	YOSE	10	5.850	0.968	0.052	0.042	0.046	-0.011	705168	15935	1892	0.268		29.7	-
		1569	YOSE	1	4.705	0.967	0.051	0.042	0.048	-0.006	624616	14080	1597	0.256			-
East	Isberg	1097	YOSE	6	10.582	0.956	0.062	0.059	0.062	0.002	651247	14508	2614	0.401	51.7	17.3	18.2
		942	YOSE	7	9.690	0.958	0.063	0.057	0.060	-0.004	691777	15649	2723	0.394			14.4
	Ireland	1856	YOSE	15	4.843	0.964	0.059	0.047	0.052	-0.013	693483	15557	2005	0.289		19.3	-
		1951	YOSE	2	2.994	0.964	0.065	0.046	0.055	-0.017	677487	15191	1743	0.257			-
		1960	YOSE	5	4.912	0.966	0.056	0.045	0.050	-0.011	713278	16118	1985	0.278			-
		2021	YOSE	4	4.826	0.960	0.061	0.053	0.059	-0.002	670631	14972	2248	0.335			-
		2026	YOSE	9	4.881	0.959	0.064	0.055	0.062	-0.004	715689	16091	2556	0.357			-
		2059	YOSE	0	4.941	0.963	0.065	0.048	0.053	-0.022	722809	16301	2084	0.288			-
		2104	YOSE	0	4.906	0.964	0.063	0.047	0.053	-0.021	711277	16061	2012	0.283			-
		2135	YOSE	2	4.867	0.965	0.062	0.044	0.049	-0.025	694557	15629	1749	0.252			-
	Lyell	1782	YOSE	2	2.991	0.963	0.068	0.045	0.054	-0.024	562059	12615	1384	0.246		11.9	-
		1815	YOSE	7	1.991	0.955	0.070	0.056	0.076	0.008	315225	7212	964	0.306			-
		1956	YOSE	13	4.847	0.959	0.062	0.054	0.061	-0.001	715446	16078	2454	0.343			-
		1997	YOSE	8	4.869	0.961	0.059	0.052	0.058	-0.001	713281	16087	2327	0.326			-
		2039	YOSE	2	4.924	0.965	0.062	0.044	0.049	-0.025	657231	14718	1631	0.248			-
	Tresidder	2147	YOSE	2	9.669	0.962	0.061	0.051	0.054	-0.015	664369	14919	2244	0.338		11.4	11.4
	Cockscomb	2324	YOSE	6	9.685	0.963	0.055	0.049	0.051	-0.007	669010	14904	2121	0.317		12	12
	Cathedral	2351	YOSE	2	3.844	0.972	0.049	0.034	0.040	-0.017	703625	15860	1319	0.188		3	-
		2410	YOSE	1	4.861	0.969	0.055	0.039	0.043	-0.021	697857	15726	1525	0.219			-
	Polly	2498	YOSE	3	9.244	0.970	0.053	0.039	0.042	-0.024	637011	14315	1426	0.224		-	-

	Tioga	2256	YOSE	4	1.997	0.967	0.051	0.040	0.054	0.004	682252	15347	1454	0.213		-
		2312	YOSE	19	7.378	0.958	0.059	0.057	0.061	0.006	690450	15556	2656	0.385		-
		2384	YOSE	6	4.855	0.963	0.058	0.049	0.055	-0.006	666673	15040	2048	0.307		-
		2407	YOSE	6	5.664	0.957	0.060	0.059	0.065	0.014	653070	14637	2633	0.403		-
		2429	YOSE	10	4.860	0.962	0.057	0.050	0.056	-0.001	715263	16049	2265	0.317		-
		2594	YOSE	3	4.641	0.960	0.063	0.053	0.060	-0.007	659929	14844	2185	0.331	181.2	-
		2618	YOSE	5	3.885	0.962	0.055	0.049	0.056	0.004	713070	16125	2115	0.297		-
		2824	YOSE	10	2.993	0.960	0.064	0.052	0.062	-0.003	663649	14909	2010	0.303		-
		2836	YOSE	2	1.997	0.967	0.056	0.040	0.054	-0.004	717203	16223	1539	0.215		-
		2864	YOSE	30	12.449	0.958	0.058	0.056	0.059	0.006	675260	15138	2779	0.412		-
		2893	YOSE	13	4.901	0.959	0.063	0.054	0.061	-0.003	718752	16213	2511	0.349		-
		2977	YOSE	3	4.881	0.962	0.060	0.050	0.056	-0.007	710096	16007	2246	0.316		-
	Conness	3200	YOSE	17	4.900	0.957	0.062	0.057	0.063	0.004	728048	16476	2641	0.363		-
		3225	YOSE	10	4.867	0.960	0.061	0.053	0.059	-0.002	721912	16279	2446	0.339		-
		3272	YOSE	7	4.589	0.958	0.066	0.055	0.062	-0.007	651654	14543	2207	0.339		-
		3339	YOSE	0	3.865	0.963	0.065	0.048	0.055	-0.018	711896	16096	2029	0.285	29.8	-
		3371	YOSE	12	4.736	0.959	0.060	0.054	0.060	0.002	535599	11989	1777	0.332		82.4
		3420	YOSE	6	4.572	0.967	0.054	0.043	0.049	-0.010	680824	15355	1815	0.267		-
		3424	YOSE	6	4.829	0.962	0.062	0.050	0.056	-0.012	635710	14230	1925	0.303		17.8
	Miller	3342	YOSE	3	3.831	0.966	0.057	0.043	0.050	-0.014	669341	15108	1572	0.235		-
		3400	YOSE	7	4.781	0.959	0.065	0.054	0.060	-0.009	670636	15071	2252	0.336	9.2	-
		3414	YOSE	12	4.843	0.957	0.062	0.058	0.065	0.007	697514	15640	2627	0.377		-
West	Bald	2369	YOSE	4	9.725	0.968	0.048	0.043	0.045	-0.006	694699	15542	1888	0.272	12	12
	Ribbon 1	1779	YOSE	6	4.832	0.965	0.049	0.046	0.052	0.006	688962	15407	1953	0.284		-
		1841	YOSE	3	4.822	0.964	0.056	0.047	0.053	-0.004	651141	14545	1913	0.294	19.3	-
		2012	YOSE	11	4.796	0.965	0.053	0.046	0.051	-0.004	650050	14696	1879	0.289		-
		2030	YOSE	11	4.865	0.967	0.051	0.043	0.048	-0.005	710236	15968	1827	0.257		8.2
	Ribbon 2	1756	YOSE	1	4.907	0.972	0.049	0.035	0.040	-0.019	707100	15924	1368	0.194	-	-
	Porcupine	2073	YOSE	5	4.931	0.971	0.051	0.038	0.042	-0.016	723989	16300	1646	0.227		-
		2118	YOSE	3	4.879	0.971	0.050	0.036	0.040	-0.020	654372	14665	1253	0.192	2.8	-
		2132	YOSE	1	4.859	0.971	0.053	0.036	0.041	-0.022	709566	15951	1455	0.205		-
	White Wolf	2371	YOSE	9	4.820	0.965	0.054	0.047	0.053	-0.002	688384	15369	2043	0.297		-
		2385	YOSE	7	4.805	0.965	0.053	0.046	0.051	-0.003	685841	15403	1963	0.286		-
		2391	YOSE	8	4.853	0.966	0.054	0.044	0.049	-0.008	684791	15383	1848	0.270		-
		2411	YOSE	19	4.864	0.969	0.049	0.041	0.046	-0.005	717748	16178	1823	0.254	22.8	7
		2418	YOSE	1	4.877	0.967	0.054	0.043	0.048	-0.012	709080	16083	1843	0.260		-
		2421	YOSE	2	4.795	0.966	0.055	0.045	0.050	-0.008	688506	15488	1909	0.277		-
		2443	YOSE	4	4.833	0.968	0.058	0.041	0.046	-0.022	667581	15044	1585	0.237		-
North	Slide	3612	YOSE	2	3.733	0.965	0.058	0.046	0.053	-0.009	686773	15518	1896	0.276		-
		3615	YOSE	19	4.817	0.956	0.064	0.058	0.065	0.004	704697	15969	2645	0.375	27.7	-
		3763	YOSE	7	6.888	0.957	0.065	0.058	0.063	-0.003	689848	15555	2658	0.385		11.7
		4025	YOSE	30	9.747	0.950	0.070	0.067	0.071	0.004	706993	15972	3357	0.475		22.1
	Rodgers	3384	YOSE	1	1.995	0.963	0.063	0.045	0.061	-0.003	683673	15432	1661	0.243		-
		3273	YOSE	3	9.406	0.963	0.058	0.050	0.053	-0.011	661464	14958	2153	0.326	19	18.2
		3452	YOSE	21	9.183	0.958	0.062	0.057	0.060	-0.002	632392	14271	2526	0.399		40.4
	Rancheria	3537	YOSE	10	9.697	0.965	0.051	0.047	0.049	-0.003	626263	14208	1990	0.318	31.7	31.7
	Thompson	3613	YOSE	13	9.588	0.966	0.053	0.045	0.048	-0.011	677517	15461	1972	0.291	9.3	9.3
	Kerrick	4136	YOSE	50	9.412	0.954	0.066	0.063	0.067	0.005	629398	14188	2842	0.452		22.4
		4146	YOSE	14	4.655	0.958	0.065	0.055	0.062	-0.004	615031	13838	2093	0.340		-
		4164	YOSE	2	4.860	0.964	0.065	0.046	0.052	-0.026	686336	15450	1829	0.267	22	-
		4324	YOSE	45	9.608	0.954	0.066	0.063	0.067	0.005	673526	15185	3059	0.454		24.4
		4391	YOSE	10	4.769	0.961	0.067	0.050	0.057	-0.019	706978	16006	2137	0.302		-

Wells	4365	YOSE	23	9.647	0.958	0.061	0.056	0.059	0.001	637383	14369	2414	0.379		17.3	17.3
Tilden	4370	YOSE	26	9.728	0.960	0.058	0.053	0.055	-0.005	680026	15422	2360	0.347		8.2	8.2
Twin Lakes	4317	YOSE	36	9.765	0.967	0.056	0.044	0.047	-0.018	636930	14392	1832	0.288		11	11
		All	**10.471**	**6.084**	**0.963**	**0.059**	**0.049**	**0.054**	**-0.007**	**669088.755**	**15099.725**	**2057.500**	**0.309**	**25.767**	**23.332**	**29.125**
		KICA	**24.417**	**8.699**	**0.955**	**0.066**	**0.060**	**0.064**	**-0.002**	**611002.500**	**14019.667**	**2532.583**	**0.414**	**13.800**	**25.217**	**53.667**
		YOSE	**8.611**	**5.735**	**0.964**	**0.058**	**0.047**	**0.053**	**-0.008**	**676833.589**	**15243.733**	**1994.156**	**0.295**	**31.750**	**22.880**	**19.522**

Supplementary Table 3. Summary of population genetics by ADMIXTURE cluster.

Clade	Cluster	Park	PA	N	P	Ho	He	Pi	Fis	Sites	Var Sites	Pol Sites	% Pol Loci	Clade Ne	Cluster Ne
Evolution	Sapphire	KICA	28	6.631	0.966	0.055	0.045	0.050	-0.011	605443	13951	1758	0.291	3.3	3.3
Goddard	Emerald	KICA	47	21.270	0.957	0.062	0.059	0.060	-0.002	640827	14686	2766	0.432	24.3	4.2
	Finger	KICA	34	5.838	0.962	0.061	0.049	0.054	-0.015	622232	14348	1770	0.285		60.2
	Martha	KICA	22	8.116	0.951	0.071	0.067	0.071	0.002	610523	13966	2900	0.474		53.7
	Peak 12,434	KICA	9	9.474	0.950	0.072	0.068	0.072	0.003	598925	13766	2902	0.485		9.3
	Tehipite	KICA	16	5.847	0.961	0.061	0.052	0.057	-0.006	596020	13740	2038	0.342		20.6
South	Bridalveil	YOSE	9	5.235	0.968	0.051	0.041	0.047	-0.007	688178	15534	1723	0.250	23.6	13.2
	Chilnualna	YOSE	3	4.890	0.969	0.052	0.039	0.044	-0.015	686272	15480	1622	0.236		15.2
	Starr King	YOSE	6	5.277	0.968	0.052	0.042	0.047	-0.009	664892	15008	1745	0.262		29.7
	Summit	YOSE	5	9.739	0.971	0.053	0.037	0.039	-0.028	715850	16155	1551	0.217		-
	Wawona	YOSE	5	9.726	0.973	0.045	0.036	0.038	-0.014	697445	15691	1621	0.232		15.7
East	Cathedral	YOSE	2	4.353	0.971	0.052	0.037	0.042	-0.019	700741	15793	1422	0.203	51.7	3
	Cockscomb	YOSE	6	9.685	0.963	0.055	0.049	0.051	-0.007	669010	14904	2121	0.317		12
	Conness	YOSE	8	4.623	0.961	0.061	0.051	0.058	-0.006	666520	14995	2120	0.318		29.8
	Ireland	YOSE	5	4.646	0.963	0.062	0.048	0.054	-0.014	699901	15740	2048	0.293		19.3
	Isberg	YOSE	7	10.136	0.957	0.063	0.058	0.061	-0.001	671512	15079	2669	0.398		17.3
	Lyell	YOSE	6	3.924	0.961	0.064	0.050	0.060	-0.009	592648	13342	1752	0.294		11.9
	Miller	YOSE	7	4.485	0.961	0.061	0.052	0.058	-0.005	679164	15273	2150	0.316		9.2
	Polly	YOSE	3	9.244	0.970	0.053	0.039	0.042	-0.024	637011	14315	1426	0.224		-
	Tioga	YOSE	9	5.042	0.961	0.059	0.051	0.058	0.000	688806	15507	2203	0.321		181.2
	Tresidder	YOSE	2	9.669	0.962	0.061	0.051	0.054	-0.015	664369	14919	2244	0.338		11.4
West	Bald	YOSE	4	9.725	0.968	0.048	0.043	0.045	-0.006	694699	15542	1888	0.272	33.1	12
	Porcupine	YOSE	3	4.890	0.971	0.051	0.037	0.041	-0.019	695976	15639	1451	0.208		2.8
	Ribbon 1	YOSE	8	4.829	0.965	0.052	0.046	0.051	-0.002	675097	15154	1893	0.281		19.3
	Ribbon 2	YOSE	1	4.907	0.972	0.049	0.035	0.040	-0.019	707100	15924	1368	0.194		-
	White Wolf	YOSE	7	4.835	0.966	0.054	0.044	0.049	-0.008	691704	15564	1859	0.269		22.8
North	Kerrick	YOSE	24	6.661	0.958	0.066	0.056	0.061	-0.008	662254	14933	2392	0.363	18.6	22
	Rancheria	YOSE	10	9.697	0.965	0.051	0.047	0.049	-0.003	626263	14208	1990	0.318		31.7
	Rodgers	YOSE	8	6.861	0.961	0.061	0.051	0.058	-0.005	659176	14887	2113	0.323		19
	Slide	YOSE	15	6.296	0.957	0.064	0.057	0.063	-0.001	697078	15754	2639	0.378		27.7
	Thompson	YOSE	13	9.588	0.966	0.053	0.045	0.048	-0.011	677517	15461	1972	0.291		9.3
	Tilden	YOSE	26	9.728	0.960	0.058	0.053	0.055	-0.005	680026	15422	2360	0.347		8.2
	Twin Lakes	YOSE	36	9.765	0.967	0.056	0.044	0.047	-0.018	636930	14392	1832	0.288		11
	Wells	YOSE	23	9.647	0.958	0.061	0.056	0.059	0.001	637383	14369	2414	0.379		17.3
		All	**12.224**	**7.508**	**0.963**	**0.057**	**0.048**	**0.052**	**-0.009**	**662867.432**	**14983.513**	**2021.230**	**0.307**		**23.724**
		KICA	**25.972**	**9.529**	**0.958**	**0.064**	**0.057**	**0.061**	**-0.005**	**612328.389**	**14076.056**	**2355.583**	**0.384**		**25.217**
		YOSE	**9.278**	**7.075**	**0.965**	**0.056**	**0.046**	**0.051**	**-0.010**	**673697.227**	**15177.968**	**1949.582**	**0.290**		**22.880**