



# Persistent panmixia despite extreme habitat loss and population decline in the threatened tricolored blackbird (*Agelaius tricolor*)

Kelly Barr<sup>1,2</sup>  | Annabel C. Beichman<sup>2</sup> | Pooneh Kalhori<sup>2</sup> | Jasmine Rajbhandary<sup>2</sup> | Rachael A. Bay<sup>3</sup> | Kristen Ruegg<sup>4</sup> | Thomas B. Smith<sup>1,2</sup> 

<sup>1</sup>Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA

<sup>2</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, USA

<sup>3</sup>Department of Evolution and Ecology, University of California, Davis, Davis, CA, USA

<sup>4</sup>Department of Biology, Colorado State University, Fort Collins, CO, USA

## Correspondence

Kelly Barr, Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA.  
Email: kellybarr@gmail.com

## Funding information

Pasadena Audubon Society; La Kretz Center of UCLA; California Conservation Genomics Consortium

## Abstract

Habitat loss and alteration has driven many species into decline, often to the point of requiring protection and intervention to avert extinction. Genomic data provide the opportunity to inform conservation and recovery efforts with details about vital evolutionary processes with a resolution far beyond that of traditional genetic approaches. The tricolored blackbird (*Agelaius tricolor*) has suffered severe losses during the previous century largely due to anthropogenic impacts on their habitat. Using a dataset composed of a whole genome paired with reduced representation libraries (RAD-Seq) from samples collected across the species' range, we find evidence for panmixia using multiple methods, including PCA (no geographic clustering), admixture analyses (ADMIXTURE and TESS conclude  $K = 1$ ), and comparisons of genetic differentiation (average  $F_{ST} = 0.029$ ). Demographic modeling approaches recovered an ancient decline that had a strong impact on genetic diversity but did not detect any effect from the known recent decline. We also did not detect any evidence for selection, and hence adaptive variation, at any site, either geographic or genomic. These results indicate that species continues to have high vagility across its range despite population decline and habitat loss and should be managed as a single unit.

## KEYWORDS

conservation genetics, demographic modeling, effective population size, gene flow, genomics, habitat loss, threatened species

## 1 | INTRODUCTION

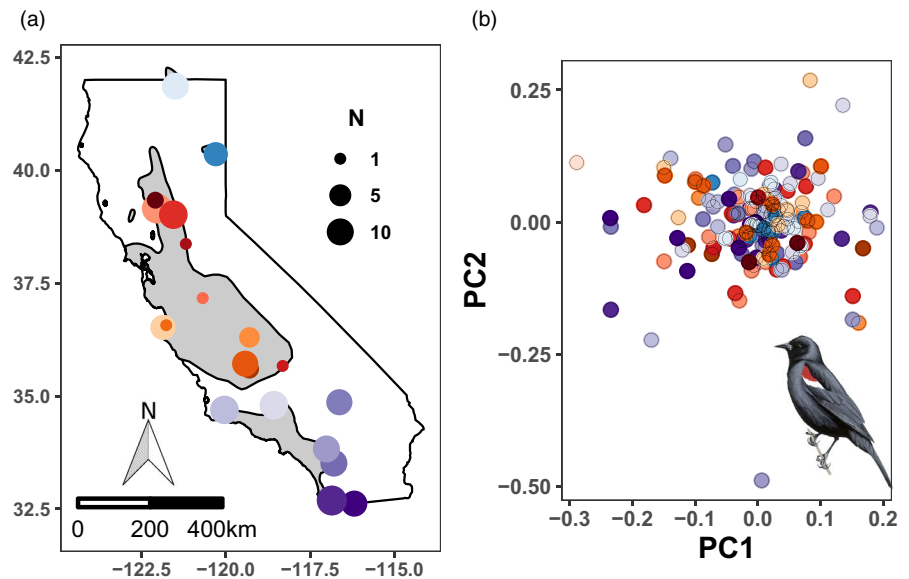
Rising anthropogenic pressures over the past century have created a global biodiversity crisis (Ceballos et al., 2015). Countless species have experienced significant population decline due to habitat reduction and alteration in the course of human activities, and many

are now threatened with extinction (Zalasiewicz et al., 2011). Efforts to slow and reverse these trends are often limited by a deficiency of information regarding the evolutionary processes that dictate long-term species survival (Smith & Bernatchez, 2008). Historically, attempts to fill this information gap employed genetic markers with the capacity for evaluating only coarse genetic patterns (e.g.,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

**FIGURE 1** (a) Tricolored blackbird range (gray) and locations of breeding colonies where samples were collected ( $N = 153$ ). Additional details about samples and collection are available in Table S1. (b) Plot of principal components analysis (PCA) of 153 tricoloreds genotyped at 70,933 single nucleotide polymorphisms (SNPs). Each point is an individual colored by general sample location indicated in (a). Clear mixing of genotypes and very low loadings (both  $<1\%$ ) are consistent with panmixia across the breeding range. Warm colors represent colonies sampled in the range core and cool colors are peripheral or southern California



microsatellites or mitochondrial DNA sequences). A primary objective of these approaches, for instance, was the identification of evolutionarily significant units (ESUs) that may encompass unique and possibly adaptive variation and hence warrant targeted protection (Moritz, 1994; Ryder, 1986). Given the increasing accessibility of genome-wide data, we can now move toward more precise evaluations of evolutionary processes by directly assessing adaptive variation (Bay et al., 2018; Funk et al., 2012; Ruegg et al., 2018), analyzing fine-scale gene flow patterns and hierarchical genetic structure (Hendricks et al., 2017; Ruegg et al., 2014; Younger et al., 2017), and estimating recent and historical demographic trends (Beichman et al., 2018, 2019; Oh et al., 2019).

One such species experiencing severe impacts in the course of anthropogenic activities is the tricolored blackbird (tricoloreds; *Agelaius tricolor*), a colonial songbird that is near endemic to California (Beedy et al., 2018). Tricoloreds are now listed as threatened at the state level after declining by an estimated 63% from 1935 to 1975 (Graves et al., 2013) and another 34% from 2007 to 2016 (Robinson et al., 2018). These losses are primarily due to the destruction of the species' historically preferred habitats for nesting, wetlands, and foraging, grasslands, by extensive agricultural and urban development (Beedy et al., 2018). As a consequence, whereas 93% of surveyed colonies nested in wetlands in the 1930s (Neff, 1937), tricolored colonies today use a broad range of nesting substrates, often including croplands and invasive species (Meese, 2017). Led by a multiagency collaboration of public and private interests (The Tricolored Blackbird Working Group; Kester, 2007), substantial time and financial resources have been committed toward their conservation and recovery over the past two decades. The lone genetic study guiding these efforts, Berg et al. (2010), reported no differentiation and varying levels of genetic diversity among colonies using a small suite of microsatellites and mitochondrial sequences. The limitations of these data leave many questions about range-wide genetic connectivity and the impacts of population decline on the overall genetic diversity in the species.

Here, we offer a comprehensive examination of current levels of gene flow and genetic diversity in the tricolored blackbird using genome-wide data. We sample numerous colonies breeding at the range periphery that were not covered by Berg et al. (2010) and where the earliest impacts of declining population sizes and restricted gene flow is expected. Using multiple demographic modeling approaches to distinguish between recent and historical events, we assess genetic diversity at multiple temporal scales. Finally, we investigate evidence for local adaptation using outlier and genotype–environment association (GEA) analyses. Our primary objectives are to (a) assess gene flow and genetic diversity, both neutral and adaptive (the latter being the product of local environmental selection), (a) estimate current and long-term effective population sizes ( $N_e$ ), and (c) provide management recommendations based upon our results.

## 2 | MATERIALS AND METHODS

### 2.1 | Genetic sampling

We obtained tissue samples from breeding tricolored colonies throughout their range (Figure 1a). Detailed information about sample sites, numbers of individuals, tissue types, and sources are provided in Table S1. From tissue samples, we purified DNA using DNeasy Blood and Tissue Kits (Qiagen) and assessed extract quantity using a Qubit (ThermoFisher) and quality with an agarose gel. We collected genetic data through two means: (a) whole genome sequencing with deep coverage ( $n = 1$ ) and, (b) restriction-site associated sequencing (RAD-Seq;  $n = 329$ ).

### 2.2 | Genome sequencing

We prepared DNA for whole genome sequencing using the Illumina TruSeq DNA PCR-Free LT kit (Illumina). After fragmenting  $1 \mu\text{g}$  of

DNA to 400 bp using a Diagenode sonicator and cleaning with magnetic beads at a ratio of 105  $\mu$ l of beads/79  $\mu$ l of water to select for >400 bp fragments, bioanalyzer traces were collected by the University of California, Los Angeles GenoSeq Core to verify library quality. We sequenced a final library with fragments averaging ~500 bp using a 250 bp paired-end run on an Illumina HiSeq2500 at the University of California QB3 Vincent J. Coates Genomics Sequencing Laboratory. Scaffolds were assembled from resulting sequence data via the Discovar DeNovo assembler (Broad Institute), and those <5,000 bp were removed. We used BUSCO (Simão et al., 2015) to estimate genome completeness by searching for single copy orthologs common to all species in the class Aves.

## 2.3 | Variant discovery

We developed genomic libraries using bestRAD (Ali et al., 2016). For these, DNA was digested using the SbfI restriction enzyme (New England Biolabs, NEB), cleaned using 1X Agencourt AMPure XP beads (Beckman Coulter), ligated with biotinylated adaptors, and sheared to 400bp fragments with a Bioruptor NGS sonicator (Diagenode). We filtered out nonligated fragments using magnetic beads (Dynabeads M-280; Life Technologies). Blunt ends were repaired and ligated with adaptors via the Illumina NEBNext Ultra DNA Library Prep Kit (NEB), and 500 bp fragments were selected with AMPure beads. PCR-enrichment was tested using 5  $\mu$ l of library with a maximum of 15 cycles. Based upon product brightness on an agarose gel, 15  $\mu$ l of library was then amplified for an appropriate number of cycles, cleaned with AMPure beads, and verified via bioanalyzer traces at the UCLA Technology Center for Genomics and Bioinformatics.

We sequenced RAD-Seq libraries over four lanes of 100bp paired-end reads on an Illumina HiSeq2500 at the UC-Davis DNA Technologies Core, and used the “process\_radtags” function in STACKS (Catchen et al., 2013) to demultiplex, filter, trim adaptors, and remove low-quality reads. PCR duplicates were removed using the “clone\_filter” function. We mapped reads to the genome assembly with bowtie2 (Langmead & Salzberg, 2012) and identified single nucleotide polymorphisms (SNPs) using the Haplotype Caller module in the Genome Analysis Toolkit (McKenna et al., 2010). We removed low-quality variants (genotype quality <30, depth <8, minor allele frequency <0.01), indels, and nonbiallelic SNPs with vcfTools (Danecek et al., 2011). To determine filtering levels for missing data, we visualized and assessed missingness using the R package “genoscapeRtools” (Anderson, 2019).

## 2.4 | Population structure

Population structure is in part a consequence of recent gene flow, and hence is indicative both of a species' natural and recently developed changes in movement and dispersal patterns. Since closely related individuals can bias signatures of population structure and genetic

diversity, we used KING (Manichaikul et al., 2010) to estimate kinship and removed individuals from pairs detected to have first-order relationships (kinship > 0.177). We conducted principal components analyses (PCA) with the R package SNPRelate (Zheng et al., 2012) and sequentially removed visual outliers. Heterozygosity, both observed ( $H_O$ ) and expected ( $H_E$ ), of detected variants and the inbreeding coefficient ( $F_{IS}$ ) were estimated with the STACKS POPULATIONS module. We calculated global Tajima's D that is bias-corrected for missing data and tested for significance with 1,000 simulations in the R package “r2vcftools” (Pope, 2019). Using ADMIXTURE (Alexander et al., 2009) and the spatially explicit Bayesian clustering algorithm TESS (Caye et al., 2016), we estimated the number of genetic clusters in the dataset and assessed individual-level admixture. We calculated pairwise  $F_{ST}$  among sample groups (with  $N \geq 3$ ) using the POPULATIONS module and tested for isolation by distance (IBD) with a Mantel test.

## 2.5 | Historical demography and effective population size

While tricoloreds experienced a sharp decline through the 20th century, it is possible that older events also impacted genome-wide diversity patterns. We examined the species' demographic history using multiple approaches to understand the impacts of population declines on genetic diversity. These include inferences from the folded site frequency spectrum (SFS), via  $\partial a \partial i$  (Gutenkunst et al., 2009) and fastsimcoal2 (Excoffier et al., 2013), and scenario tests using approximate Bayesian computation (ABC) as implemented in DIY-ABC v2.0 (Cornuet et al., 2014). We also estimated  $N_e$  for the current generation based upon linkage disequilibrium (Waples & Do, 2010) using the program NEESTIMATOR v2 (Do et al., 2014).

For both  $\partial a \partial i$  and fastsimcoal2, we used an SFS generated via a modification of easySFS (<https://github.com/isaacovercast/easySFS>) to smooth over missing data and maximize the total number of SNPs using a hypergeometric projection of a dataset filtered to remove loci with >75% heterozygosity. Focusing on historical population size changes for a single population (see results), we compared multiple demographic models: a nested “one epoch” model with no size changes, a “two epoch” model with a single size change, and a “three epoch” model with two size changes (Figure 3). We assumed a mutation rate ( $\mu$ ) of  $4.6 \times 10^{-9}$  (Smeds et al., 2016), generation time ( $g$ ) of 2 years for both  $\partial a \partial i$  and fastsimcoal2, and a sequence length ( $L$ ) of 60,429,389 bps. This  $L$  is based on sites that had at least 190X coverage ( $10X \times$  the number of individuals used in the SFS projection) across a merged BAM file composed of the individuals that passed quality filters with and with no close relatives or PCA outliers (see results). Meanwhile,  $g$  is based upon a robust estimation from another Passerine (Brommer et al., 2004), and the known ages of first breeding of one year for female and two years for male tricoloreds (Beedy et al., 2018).

In  $\partial a \partial i$ , demographic parameters for each model are determined by solving an approximation to the diffusion equation (Gutenkunst

MODEL		$\delta a\delta i$	fastsimcoal2
1 EPOCH	LL	-7,948	-2,506,605
	$Ne_a$	$Ne_a$	151,228
2 EPOCHS	LL	-121	-2,503,286
	$Ne_a$	$Ne_a$	174,455
	$Ne_{it}$ ↑ $T_{it}$	$T_{it}$	21,317
	$Ne_{it}$	$Ne_{it}$	91,315
3 EPOCHS	LL	-118	-2,503,287
	$Ne_{a2}$	$Ne_{a2}$	174,932
	$Ne_{a1}$ ↑ $T_{a1}$	$T_{a1}$	24,101
	$Ne_{a1}$	$Ne_{a1}$	101,739
	$Ne_{it}$ ↑ $T_{it}$	$T_{it}$	2,501
	$Ne_{it}$	$Ne_{it}$	78,023

**FIGURE 3** Log-likelihoods (LL) and parameter estimations (including Time,  $T$ , in generations and effective population size,  $Ne$ ) for the three demographic models examined using  $\delta a\delta i$  and fastsimcoal2. The “2 epoch” model was resolved to be best fitting as adding an additional change in  $Ne$  had a marginal impact on the LL. NB: The LLs are calculated differently between these two analytical frameworks and hence are not comparable

et al., 2009). We set extrapolation grid points for simulations as the haploid sample size of the projected SFS and that plus 5, 15, and 25. Using 50 independent replicates, we carried out inferences with permuted starting parameter values and assessed the fit of the expected SFS under the inferred model parameters to the empirical SFS using a multinomial log-likelihood. For each model, the maximum likelihood estimate (MLE) for the set of 50 runs was selected. The best-fit  $\theta$  (population scaled mutation rate) for each replicate's parameters was inferred using  $\delta a\delta i$ , and scaled by  $L$  and  $\mu$  to calculate the ancestral size in diploids ( $Ne_{anc}$ ):

$$Ne_{anc} = \frac{\theta}{4\mu L}$$

Population sizes inferred in  $\delta a\delta i$  were then scaled by  $Ne_{anc}$ , and times were scaled by  $2 * Ne_{anc} * g$ . After determining the best-fit model, we used a grid-search approach to refine the bounds of our inferred parameters. We examined a 100x100 grid of values of  $nu$  (contraction size relative to  $Ne_{anc}$ ) and  $T$  (contraction duration in terms of  $2 * Ne_{anc} * g$ ) spaced evenly along a  $\log_{10}$  scale. We obtained the expected SFS for each of the 10,000 parameter pairs using  $\delta a\delta i$  and calculated the multinomial log-likelihood. We then plotted the delta log-likelihood between each parameter pair and the MLE as a heatmap.

Fastsimcoal2 offers an alternative analytical framework for using the SFS for demographic inferences using coalescent simulations. For each of the previously described demographic models, we ran 100,000 coalescent simulations and used 50 Expected/Conditional Maximization (ECM) parameter-optimization cycles to estimate the expected SFS for each set of parameters.

We conducted ABC analyses in a sequential, hypothesis-testing manner with at least 100,000 simulations of each scenario (Figure S6). First, we compared contraction and expansion scenarios, then single versus multiple contractions, and finally, we compared scenarios with a single contraction at four-time frames: recent ( $T_a = 0-99$  g ago), older ( $T_a = 100-999$  g ago), historical ( $T_a = 1,000-9,999$  g ago), and deeply historical ( $T_a = 10,000-99,999$  g ago). For computational efficiency, we used 1,000 randomly selected loci from the real dataset of 153 individuals and simulated the same number through the scenarios. Using proportion of monomorphic loci, Nei's (1987) mean gene diversity, variance of gene diversity across polymorphic loci, and mean gene diversity across all loci as summary statistics, we assessed scenario accuracy with a PCA and estimated posterior probabilities of scenarios using both direct and logistic regression approaches. Sampling priors used for simulations are provided in Figure S6.

Estimations of contemporary, short-term  $Ne$  using the LD method are downwardly biased by the presence of overlapping generations (Waples et al., 2014). To limit this effect, we focused estimations on adults sampled in 2017 and 2018, removing nestlings and samples collected in 2002 and 2008. We considered the 132 remaining samples to be a single population based upon the lack of genetic structure (see results). For these samples, we used PLINK2.0 (Chang et al., 2015) to prune varying numbers of SNPs to assess the effect of using reduced representation of the genome on calculations of  $Ne$  and report results with a minimum allele frequency (MAF) of 0.01.

## 2.6 | Testing for selection

The presence of adaptive variants would be critical information for conservation planning. We used both outlier analyses and genotype-environment association (GEA) methods to detect loci potentially under selection. For the former, we used PCAdapt (Luu et al., 2017) to jointly estimate genetic structure and assess significantly differentiated loci. We used the R package “qvalue” (Storey et al., 2019) to adjust p-values using the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995). GEAs were assessed using both redundancy analyses (RDA; Forester et al., 2018) and a machine learning approach (gradient forests; Breiman, 2001). For RDA, we used the R package “vegan” (Oksanen et al., 2019) to conduct a permutation test for constrained correspondence for significance testing. Our gradient forest approach detects areas of genotypic transitions associated with environmental conditions. For this, we used the R package “gradientForest” (Ellis et al., 2012) using the following parameters:  $n_{tree} = 100$ ,  $n_{bin} = 101$ ,  $corr.threshold = 0.5$ . We ran 10 additional gradient forests with randomized environmental variables for confidence testing. Both of these GEAs were based upon 19 climate variables downloaded from WorldClim (Hijmans et al., 2005), the vegetation indices NDVI and NDVIstd for May of 2018 (Carroll et al., 2004), tree cover (Sexton et al., 2013), elevations from the Global Land Cover Facility (www.landcover.org), and surface water measurements (QuickScat; from scp.byu.edu).

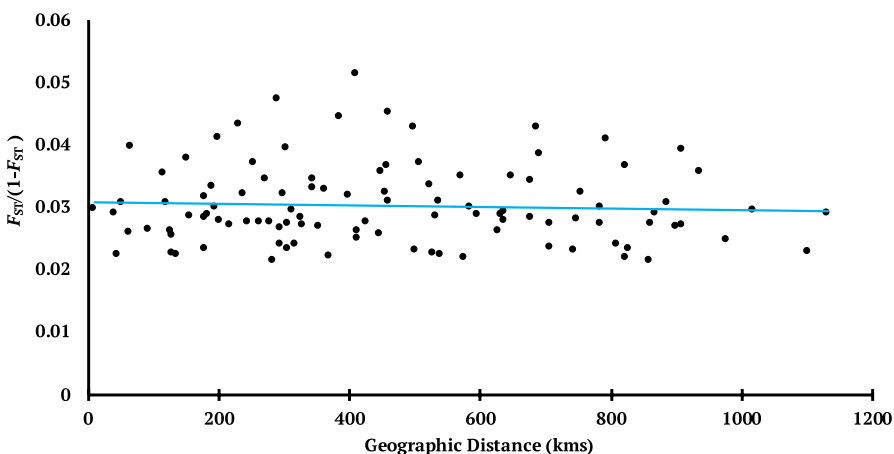
### 3 | RESULTS

#### 3.1 | Data quality

We obtained 389 samples from throughout the tricolored's breeding range, including many colonies sampled along range periphery (Figure 1a; Table S1). The genome we assembled for the species is 1.08 Gb in total length across 70,524 scaffolds with an N50 of 103,912 and >100× coverage. Of a total 4,915 known single copy orthologs in Aves, the de novo genome assembly includes 87.2% of these represented completely, 8.3% are fragmented, and 4.5% missing. RAD-Seq libraries were created for 329 individuals that passed DNA quality standards. After filtering individuals and loci with missing data >10% ( $N = 219$ ) to maximize quantity and quality (Figure S1) and removing close relatives ( $N = 6$ ) and PCA outliers ( $N = 11$ ), a dataset with 153 tricoloreds genotyped at an average of 68,366 SNPs was used for analyses except where indicated. Additional data assessments are provided in Figures S14–S16.

#### 3.2 | Gene flow and genetic diversity

We found no evidence for population structure in either PCA (Figure 1b) or clustering analyses—both ADMIXTURE and TESS indicated  $K = 1$ —suggesting high gene flow across the species' range. This pattern is supported by similar levels of heterozygosity ( $H_O$ : 0.23–0.24;  $H_E$ : 0.19–0.22) and inbreeding ( $F_{IS}$ : -0.012 to -0.002) across sample sites (Table S1), which also suggests no individual breeding colonies are in genetic isolation. An excess of rare alleles is suggested in a bias-corrected Tajima's  $D$  was significant ( $p < .001$ ) and positive (1.95; CI: 1.929–1.972), which is a pattern usually attributed to a population expansion after a decline. Average pairwise  $F_{ST}$  between sampled colonies is quite low at 0.029 (0.02–0.049) and there was no correlation with geographic distance (Figure 2, Mantel's  $r = -0.11$ ,  $p = .84$ ). This along with a lack of significant differentiation anywhere in the examined genome (Figure S2) further indicates that gene flow is ongoing with no restrictions by either habitat fragmentation or distance.



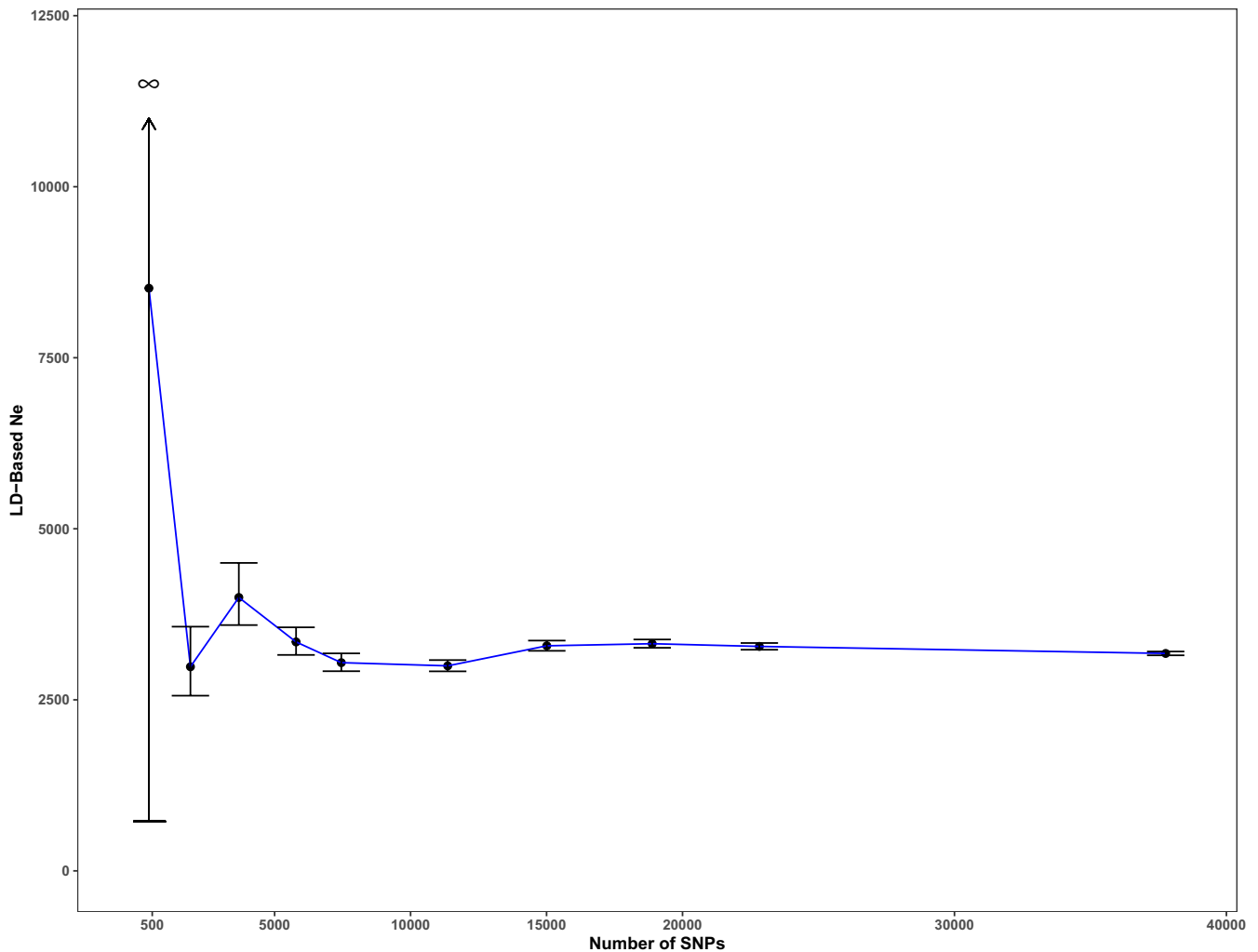
**FIGURE 2** Plot of pairwise genetic distances (normalized  $F_{ST}$ ) versus geographic distance (kms) between tricolored blackbird colonies with more than three samples. No significant relationship between differentiation and distances between colonies (Mantel's  $r = -0.11$ ;  $p = .84$ ) indicates high gene flow across the range

#### 3.3 | Historical demography and effective population size

All three demographic modeling approaches indicated a strikingly similar pattern, with the strongest event shaping the species' genetic diversity being a single contraction that is much deeper in the tricolored's evolutionary history than the known 20th-century decline. Using an SFS composed of 704,884 SNPs (Figures S3 and S4) from a hypergeometric projection of 153 diploid individuals down to 19, both  $\partial a\partial i$ , and fastsimcoal2 rejected the single-epoch (no size change) model in favor of a two-epoch contraction (Figure 3;  $p$ -value based on likelihood ratio test with two degrees of freedom  $< .00001$ ). Additional size changes did not significantly improve the fit to the data beyond the two-epoch model ( $p$ -value  $> .05$ ). ABC results were similar. The expansion and two-contraction scenarios were sequentially rejected with high confidence (Figures S7 and S8). Meanwhile, the deeply historical contraction scenario was the strongest scenario, suggesting a population decline >10,000 generations ago (Figure S9).

Each of the SFS-based approaches also arrived at similar parameter estimates of the time since the population contraction and  $N_e$ , both  $N_{e_{anc}}$  and long-term ( $N_{e_{lt}}$ ). Specific parameter estimates for the two-epoch model from the grid search in  $\partial a\partial i$  suggest ~50% population size decline occurred 21,317 (19,541 – 23,018) g/ago from an  $N_{e_{anc}}$  of 174,455 (173,734 – 175,249) to a  $N_{e_{lt}}$  of 91,315 (89,385 – 92,500; parameter ranges are within 5 log-likelihood units of the MLE; Figure 3 and S5). The parameters inferred using fastsimcoal2 were highly concordant, exhibiting a decline at 20,144 g/ago from an  $N_{e_{anc}}$  of 174,617 to a  $N_{e_{lt}}$  of 90,941.

While these methods infer a long-term  $N_{e_{lt}}$  of ~91,000, the LD method suggests current  $N_e$  is much lower at ~3,100 (Figure 4). These methods are not directly comparable as SFS-based approaches are more influenced by ancient events and the LD estimation is the product of recent genetic drift. Further, all methods are influenced by model violations in different ways. SFS-based inferences, for instance, maybe impacted by incorrect mutation rates, and the LD method may be sensitive to cryptic linkage. For instance, if we examine the impacts on  $\partial a\partial i$  results of alternative  $\mu$  used by previous authors for other birds (Nadachowska-Brzyska et al., 2015), it is



**FIGURE 4** Plot of current effective population size ( $N_e$ ) estimated from linkage disequilibrium among varying numbers of single nucleotide polymorphisms. The current  $N_e$  is  $\sim 3,100$

apparent that a lower rate assumed for a domestic pigeon (*Columba livia*;  $\mu = 4.598e-10$ ) results in  $N_{e,t}$  and  $g$  since the contraction are 10X as large; meanwhile, a higher rate such as that used for the rhinoceros hornbill (*Buceros rhinoceros*;  $\mu = 6.999e-10$ ) results in estimations that are 50% smaller (Table S3). The  $\mu$  we choose to use here from Smeds et al. (2016) seems to be the most robustly estimated one available for a fellow songbird.

### 3.4 | Testing for selection

There was no evidence for strong selection across the species' genome. The scree plot of eigenvalues calculated in PCAdapt exhibited no sharp changes in proportion of explained variation, which is the pattern expected when there is no population structure or significant outlier loci (Figure S10). Arbitrarily selecting the first four eigenvalues, no loci were detected as significant outliers after accounting for multiple tests. The GEA correlation in the RDA also found to be insignificant ( $p = .44$ ). Meanwhile, gradient forest analyses revealed both average  $r^2$  and numbers of correlated SNPs from randomized

environmental datasets were higher than those detected from the empirical data (Figure S11), suggesting any correlations detected are spurious and likely variable between runs. Despite a sampling effort aimed to capture a broad range of habitat and environmental conditions experienced by tricolors, there were no adaptive variants apparent in our dataset.

## 4 | DISCUSSION

### 4.1 | Gene flow and genetic diversity

Our data provide multiple lines of evidence indicating that tricolors persist as a single, panmictic population. The low genetic differentiation we observed both among breeding colonies and across the genome is the product of high gene flow that encompasses their complete range. Breeding colonies separated by the greatest distances, often with little suitable breeding or foraging habitat in between, are no more differentiated than the closest ones. This is particularly striking given the focus on peripheral, disparate colonies

in the sample design. Most of the birds breed in the core of their contemporary range in the Central Valley of California, and colonies at the periphery tend to be comparatively small. If isolated, genetic drift would be expected to rapidly cause differentiation and loss of genetic diversity in these smaller colonies. Here, however, the high genetic connectivity we observed even at the extremes of the breeding range further supports the conclusion that the species is in panmixia.

Our genome-wide analyses also illustrate higher and more far-reaching vagility by tricoloreds than previously known, though multiple studies employing bands on thousands of individuals have revealed large-scale movements over major portions of the breeding range. These include, for instance, across the Central Valley (Beedy et al., 2018; DeHaven et al., 1975; Neff, 1942), along the coast (Wilson et al., 2016), and throughout southern California (Neff, 1942). However, because movements were not observed between the Central Valley and southern California, in particular by Neff (1942), the prevailing notion was that tricoloreds should be considered two separate demes split between these areas. This is illustrated in an equivocal presentation of genetic diversity calculations for these two areas by Berg et al. (2010) even though they observed no significant genetic differentiation. High vagility and panmixia are important conclusions for conservation efforts, as the species as a whole can be considered a single deme.

It is notable that high genetic connectivity persists across the range despite the severe population decline and habitat loss experienced by the species over the past century. This may not be surprising for a vagile, volant species; however, genetic isolation associated with habitat fragmentation over similarly narrow extents was previously reported in other songbirds (Barr et al., 2008, 2015; Lindsay et al., 2008) and species with high dispersal distances would be expected to exhibit the earliest impacts of a barrier to gene flow (Landguth et al., 2010). It is possible that the species' historical preference for nesting in seasonal wetland habitat, which is inherently ephemeral in western North America, contributes to their capacity for maintaining genetic connectivity despite severe habitat loss. Since wetlands may not develop in the same geographic locations on an annual basis, the species is likely adapted for searching over greater areas for suitable habitat (Cerame et al., 2014). Indeed, the existence of the many colonies in extreme geographic isolation we sampled for this study further suggests a broadly ranging habitat searching behavior by tricoloreds.

#### 4.2 | Historical demography and effective population size

Considering the documented extreme decline tricoloreds experienced during the past century, it is surprising that the multiple demographic modeling approaches we employed uniformly conclude an ancient contraction ~ 20,000 generations ago has more significant impact on shaping genetic diversity in the species. Similar patterns are observed in other species, such as yellow-bellied toads (*Bombina*

*variegata*; Cornetti et al., 2016) and orcas (*Orcinus orca*; Moura et al., 2014), both of which experienced known recent declines but demographic modeling illustrates events deeper in evolutionary time are more impactful for shaping long-term genetic diversity. Though ABC methods are frequently employed for examining recent bottlenecks (e.g., Cammen et al., 2018; Richmond et al., 2013; Xenikoudakis et al., 2015), our dataset may be too coarse for detecting the impacts of the known 20th-century decline in tricoloreds given its recency and the relatively high remaining population size (2017  $N_c = 177,656$ ; Meese, 2017). It is clear from our analyses, though, that an ancient contraction occurred was highly consequential in shaping overall species genetic diversity.

While there are no clear causes of the inferred ancient decline, plausible explanations include climate change or species divergence. During high glacial periods, precipitation was high in western North America (Allen & Anderson, 1993; Oster et al., 2015) resulting in more abundant wetland habitat and likely higher tricolored population sizes. Assuming a generation time of 2 years (Brommer et al., 2004), the decline was older than the Last Glacial Maximum (20,000 y/ago); however there were numerous climatic oscillations between 20,000 and 60,000 year/ago (Petit et al., 1999) that could have resulted in significant increases or decreases in breeding habitat. As for the contraction signal being a recent species divergence, it is notable that the node between tricolored and red-winged blackbirds (*Agelaius phoeniceus*) is less well resolved than most others in an Icteridae phylogeny (Powell et al., 2014). This apparent incomplete lineage sorting is suggestive of a recent divergence between the species and possibly postdivergence hybridization. We also collected RAD-Seq libraries for ten red-winged (sampled next to tricolored colonies) to test for the possibility of hybridization but found no evidence for admixture between these closely related species that often share breeding habitat (Figures S12 and S13). TimeTree (Kumar et al., 2017) indicates these two blackbird species diverged long before (>3 Mya; based upon (Barker et al., 2015; Powell et al., 2014)) the estimated time of population contraction detected here (~40,000 years ago). This suggests the ancient cause of population decline is more likely associated with Pleistocene climate change rather than speciation.

Both the long-term  $N_e$  (~90,000) and recent  $N_e$  (~3,100) estimated here are surprisingly low given the 400,000 birds reported in a 2008 census (Kelsey, 2008) and early 1900s estimates numbering in the millions.  $N_e$  is generally smaller than  $N_c$ , and the ratio between these varies between species based upon life history characteristics (Frankham, 1995). Long-term  $N_e$ , for instance, is influenced by population fluctuations over time, with small sizes having the strongest effect (Vucetich et al., 1997). This is relevant to tricoloreds as their populations likely fluctuated throughout its evolutionary history due to interannual variation in both in prey abundance (Meese, 2013) and habitat availability. Another factor that influences long-term  $N_e$  is variance in reproductive success, with high variance reducing the ratio of  $N_e$  to  $N_c$  (Sugg & Chessler, 1994). Thus, polygyny, which is thought to be relatively high in tricoloreds (Liu, 2014), may also impact  $N_e$  in an unpredictable direction that would require additional information about the mating system to ascertain (Liu, 2015). Finally,

beyond the aforementioned biological causes, there is also an analytical component that should be considered. Our current  $N_e$  calculation may be reduced by the presence of overlapping generations (Waples et al., 2014), which is likely in our dataset because adults are long-lived (~12 years) and a lack of variation in molt and plumage beyond the second year limits age assessments (Beedy et al., 2018).

### 4.3 | Evidence for selection

While full whole genome sequences are indisputably better for assessing subtle genetic variation patterns, our results illustrate no major selective sweeps affecting large regions of the tricolored genome. It is notable that such evidence has been reported with less sequencing effort in other species with much larger genomes (e.g., Hohenlohe et al., 2010; White et al., 2013). This lack of evidence for selection may not be surprising given both the low standing genetic variation apparent in our  $N_e$  estimates and the recent, widescale shifts in habitat uses from primarily wetlands to highly variable alternative nesting substrates (Beedy et al., 2018; Meese, 2017). Another possible limitation to the development of adaptive diversity is gene swamping (Lenormand, 2002) across the range by dispersers from the Central Valley, where most of the species breeds. The high gene flow we detected would likely preclude the rise of large-effect alleles around the range periphery, where variance in environmental conditions is highest (Kirkpatrick & Barton, 1997). Additional sequencing effort focused on full genomes would be helpful for analyzing adaptive variation associated with alleles of weaker effect that may develop despite gene flow (Tigano & Friesen, 2016) and examining targeted regions directly relevant to a species' long-term viability, such as MHC loci (Agudo et al., 2012).

## 5 | CONCLUSIONS AND RELEVANCE TO MANAGEMENT

Our results illustrate the analytical power and additional information gained from reexamining a system only informed by classical genetic markers with a modern genomic approach. Berg et al. (2010) were generally inconclusive about gene flow in the species, reporting at the same time a lack of genetic differentiation but also differences in genetic diversity among sample sites. Our data allow us to conclude that tricoloreds may be managed without concern for gene flow, directed preservation of unique genetic variation, or a focus on recovery of any particular local aggregation anywhere in their range. Genetic diversity, while seemingly low overall, is homogenous across breeding colonies. These results indicate the species as a whole may justifiably be considered a single management unit.

It seems that the tricolored's natural history modulates species-wide genetic diversity and that the current level is quite a bit lower than one might predict given their recent census population sizes. While we detected no evidence for inbreeding, whether through population-level estimations of  $F_{IS}$  (Table S1) or analyses

of runs of homozygosity within individuals (data not shown), the current  $N_e$  suggests that ongoing genetic monitoring should occur to supplement censuses. Moreover, given the heterozygote excess we observed at all sample sites (Table S1), tricoloreds are likely in "drift debt" (Gilroy et al., 2017) and will experience further erosion of genetic diversity as they settle into mutation-drift equilibrium.

Our dataset and sample design should be quite powerful for the analyses we report here; however, we cannot entirely discount the possibility that our reduced representation dataset may miss weak or burgeoning genetic differentiation. Future additional sequencing effort aimed at whole genomes would significantly increase our power for detecting weak genetic differentiation or selection, and allow for a finer-scaled assessment of genetic diversity by estimating genome-wide heterozygosity. The relative impacts of the ancient and recent contraction events may be further examined through alternative analytical techniques that are less sensitive to departures from model assumptions, such as using identity by descent segments (Browning & Browning, 2015). Finally, museum samples may be used to better understand the impacts of recent population decline on the genetic diversity of the species.

## ANIMAL WELFARE AND PERMIT STATEMENTS

Samples were collected under Tom Smith's Federal Bird Banding Permit, #21901, and Kelly Barr's Scientific Collecting Permit, #SC-11568, and an MOU with California Fish and Wildlife. Animal handling and sampling protocols were conducted with the approval of UCLA's Animal Research Committee (ARC), agreement #2017-073-03.

## ACKNOWLEDGEMENTS

We thank collaborators for providing samples, including Christy Wyckoff of the Santa Lucia Conservancy, Emilie Graves and Robert Meese of UC-Davis, and Jesse Bahm of the US Dept. of Agriculture. Our extensive sampling would have been impossible without property access approval by numerous owners and managers. We thank Holt Ag Solutions of Merrill, OR, for driving 30 min out onto the loneliest road in California to jump-start our 30-year-old field vehicle. We thank Vinh Le for performing laboratory work and joining us for a day of sampling. We thank Christen Bossu for advice and bioinformatic assistance, and members of the Smith lab for comments on several manuscript draft versions. We thank Kirk E. Lohmuller for advice on demographic inferences and comments on the manuscript. Finally, we thank the California Conservation Genomics Consortium, the Audubon Society of Pasadena, and the La Kretz Center of UCLA for providing funds for this project. [Correction added on 08 November 2020, after first online publication: acknowledgement section has been updated.]

## CONFLICT OF INTEREST

None declared.



## AUTHOR CONTRIBUTIONS

KRB, ACB, PK, RAB, KR, and TBS helped design the research and write the paper. RAB designed the bioinformatic pipeline for processing raw data, and both assembled and assessed the quality of the de novo genome. KRB and JR conducted fieldwork and laboratory work. ACB and PK ran SFS inferences. KRB completed all remaining analyses.

## DATA AVAILABILITY STATEMENT

All of the genetic data collected for this study are available in public databases. These include the genome assembly in the NCBI Sequence Read Archive, accession SAMN16392922, and vcf files, both filtered and unfiltered, on Dryad, <https://doi.org/10.5068/D1DM4H>.

## ORCID

Kelly Barr  <https://orcid.org/0000-0001-9193-1094>

Thomas B. Smith  <https://orcid.org/0000-0002-5978-6912>

## REFERENCES

- Agudo, R., Carrete, M., Alcaide, M., Rico, C., Hiraldo, F., & Donazar, J. A. (2012). Genetic diversity at neutral and adaptive loci determines individual fitness in a long-lived territorial bird. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741), 3241–3249. <https://doi.org/10.1098/rspb.2011.2606>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (rapture): Flexible and efficient sequence-based genotyping. *Genetics*, 202(2), 389–400. <https://doi.org/10.1534/genetics.115.183665>
- Allen, B. D., & Anderson, R. Y. (1993). Evidence from Western North America for rapid shifts in climate during the last glacial maximum. *Science*, 260(5116), 1920–1923. <https://doi.org/10.1126/science.260.5116.1920>
- Anderson, E. C. (2019). *GenoscapeRtools: Tools for building migratory bird genoscapes*. R package version 0.1.0. <https://doi.org/10.5281/zenodo.848279>
- Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2015). New insights into New World biogeography: An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *Nuevas Revelaciones sobre la Biogeografía del Nuevo Mundo: Una Visión Integrada desde la Filogenia de los Tordos, Cardenales, Gorriones, Tangaras, Reinitas y AliadosEmberizoid phylogeny and biogeography*. *The Auk*, 132(2), 333–348. <https://doi.org/10.1642/AUK-14-110.1>
- Barr, K. R., Kus, B. E., Preston, K. L., Howell, S., Perkins, E., & Vandergast, A. G. (2015). Habitat fragmentation in coastal southern California disrupts genetic connectivity in the cactus wren (*Campylorhynchus brunneicapillus*). *Molecular Ecology*, 24(10), 2349–2363. <https://doi.org/10.1111/mec.13176>
- Barr, K. R., Lindsay, D. L., Athrey, G., Lance, R. F., Hayden, T. J., Tweddale, S. A., & Leberg, P. L. (2008). Population structure in an endangered songbird: Maintenance of genetic differentiation despite high vagility and significant population recovery. *Molecular Ecology*, 17(16), 3628–3639. <https://doi.org/10.1111/j.1365-294X.2008.03868.x>
- Bay, R. A., Harrigan, R. J., Underwood, V. L., Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, 359(6371), 83–86. <https://doi.org/10.1126/science.aan4380>
- Beedy, E. C., Hamilton, W. J. III, Meese, R. J., Airola, D. A., & Pyle, P. (2018). Tricolored Blackbird (*Agelaius tricolor*), version 1.0. In P. G. Rodewald (Ed.), *Birds of the World*. Ithaca, NY: Cornell Lab of Ornithology. <https://doi.org/10.2173/bow.tribla.01>
- Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, 49(1), 433–456. <https://doi.org/10.1146/annurev-ecolsys-110617-062431>
- Beichman, A. C., Koepfli, K.-P., Li, G., Murphy, W., Dobrynin, P., Kliver, S., ... Wayne, R. K. (2019). Aquatic adaptation and depleted diversity: A deep dive into the genomes of the sea otter and giant otter. *Molecular Biology and Evolution*, 36(12), 2631–2655. <https://doi.org/10.1093/molbev/msz101>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berg, E. C., Pollinger, J., & Smith, T. B. (2010). Population structure of the Tricolored Blackbird (*Agelaius tricolor*) in California: are northern and southern populations genetically distinct? Nongame Wildlife Program Report 2010–05. California Department of Fish and Wildlife.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brommer, J. E., Gustafsson, L., Pietiäinen, H., & Merilä, J. (2004). Single-generation estimates of individual fitness as proxies for long-term genetic contribution. *The American Naturalist*, 163(4), 505–517. <https://doi.org/10.1086/382547>
- Browning, S. R., & Browning, B. L. (2015). Accurate non-parametric estimation of recent effective population size from segments of identity by descent. *The American Journal of Human Genetics*, 97(3), 404–418. <https://doi.org/10.1016/j.ajhg.2015.07.012>
- Cammen, K. M., Schultz, T. F., Bowen, W. D., Hammill, M. O., Puryear, W. B., Runstadler, J., ... Kinnison, M. (2018). Genomic signatures of population bottleneck and recovery in Northwest Atlantic pinnipeds. *Ecology and Evolution*, 8(13), 6599–6614. <https://doi.org/10.1002/ece3.4143>
- Carroll, M. L., DiMiceli, C. M., Sohlberg, R. A., & Townshend, J. R. G. (2004). 250m MODIS normalized difference vegetation index. University of Maryland, College Park, Maryland.
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, 16(2), 540–548. <https://doi.org/10.1111/1755-0998.12471>
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), e1400253. <https://doi.org/10.1126/sciadv.1400253>
- Cerame, B., Cox, J. A., Brumfield, R. T., Tucker, J. W., & Taylor, S. S. (2014). Adaptation to ephemeral habitat may overcome natural barriers and severe habitat fragmentation in a fire-dependent species, the Bachman's sparrow (*Peucaea aestivalis*). *PLoS One*, 9(9), e105782. <https://doi.org/10.1371/journal.pone.0105782>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Cornetti, L., Benazzo, A., Hoban, S., Vernesi, C., & Bertorelle, G. (2016). Ancient, but not recent, population declines have had a genetic

- impact on alpine yellow-bellied toad populations, suggesting potential for complete recovery. *Conservation Genetics*, 17(3), 727–743. <https://doi.org/10.1007/s10592-016-0818-5>
- Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., ... Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30(8), 1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- DeHaven, R. W., Crase, F. T., & Woronecki, P. P. (1975). Movements of tricolored blackbirds banded in the central valley of California, 1965–1972. *Bird-Banding*, 46(3), 220–229. <https://doi.org/10.2307/4512139>
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Molecular Ecology Resources*, 14(1), 209–214. <https://doi.org/10.1111/1755-0998.12157>
- Ellis, N., Smith, S. J., & Pitcher, C. R. (2012). Gradient forests: Calculating importance gradients on physical predictors. *Ecology*, 93(1), 156–168. <https://doi.org/10.1890/11-0252.1>
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP Data. *PLOS Genetics*, 9(10), e1003905. <https://doi.org/10.1371/journal.pgen.1003905>
- Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multi-variant genotype–environment associations. *Molecular Ecology*, 27(9), 2215–2233. <https://doi.org/10.1111/mec.14584>
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genetics Research*, 66(2), 95–107.
- Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, 27(9), 489–496. <https://doi.org/10.1016/j.tree.2012.05.012>
- Gilroy, D. L., Phillips, K. P., Richardson, D. S., & van Oosterhout, C. (2017). Toll-like receptor variation in the bottlenecked population of the Seychelles warbler: Computer simulations see the 'ghost of selection past' and quantify the 'drift debt'. *Journal of Evolutionary Biology*, 30(7), 1276–1287. <https://doi.org/10.1111/jeb.13077>
- Graves, E. E., Holyoak, M., Rodd Kelsey, T., & Meese, R. J. (2013). Understanding the contribution of habitats and regional variation to long-term population trends in tricolored blackbirds. *Ecology and evolution*, 3(9), 2845–2858.
- Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, 5(10), e1000695. <https://doi.org/10.1371/journal.pgen.1000695>
- Hendricks, S., Epstein, B., Schönfeld, B., Wiench, C., Hamede, R., Jones, M., ... Hohenlohe, P. (2017). Conservation implications of limited genetic diversity and population structure in Tasmanian devils (*Sarcophilus harrisii*). *Conservation Genetics*, 18(4), 977–982. <https://doi.org/10.1007/s10592-017-0939-5>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, 6(2), e1000862. <https://doi.org/10.1371/journal.pgen.1000862>
- Kelsey, R. (2008). Results of the tricolored blackbird 2008 census. Report submitted to US Fish and Wildlife Service, Portland, OR, USA.
- Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *The American Naturalist*, 150(1), 1–23. <https://doi.org/10.1086/286054>
- Kumar, S., Stecher, G., Suleski, M., & Heddes, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, 34(7), 1812–1819. <https://doi.org/10.1093/molbev/msx116>
- Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M., & Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, 19(19), 4179–4191. <https://doi.org/10.1111/j.1365-294X.2010.04808.x>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, 17(4), 183–189. [https://doi.org/10.1016/S0169-5347\(02\)02497-7](https://doi.org/10.1016/S0169-5347(02)02497-7)
- Lindsay, D. L., Barr, K. R., Lance, R. F., Tweddle, S. A., Hayden, T. J., & Leberg, P. L. (2008). Habitat fragmentation and genetic diversity of an endangered, migratory songbird, the golden-cheeked warbler (*Dendroica chrysoparia*). *Molecular Ecology*, 17(9), 2122–2133. <https://doi.org/10.1111/j.1365-294X.2008.03673.x>
- Liu, I. A. (2014). Molecular Causes and Consequences of Sperm Competition in *Agelaius* Blackbirds (Doctoral dissertation, Duke University). Retrieved from <https://dukespace.lib.duke.edu/dspace/handle/10161/8795>
- Liu, I. A. (2015). Conservation genetics and genetic mating system of the yellow-shouldered blackbird (*Agelaius xanthomus*), an endangered island endemic. *Conservation Genetics*, 16(5), 1041–1053. <https://doi.org/10.1007/s10592-015-0721-5>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust relationship inference in genomewide association studies. *Bioinformatics*, 26(22), 2867–2873.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Meese, R. J. (2013). Chronic low breeding success in the tricolored blackbird from 2006–2011. *Western Birds*, 44, 98–113.
- Meese, R. J. (2017). Results of the 2017 Tricolored Blackbird Statewide Survey. Calif. Dept. of Fish and Wildlife, Wildlife Branch, Nongame Wildlife Program Report 2017-04, Sacramento, CA. [California Department of Fish and Wildlife, Nongame Wildlife Program Report 2017-04]. 1–27.
- Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution*, 9(10), 373–375. [https://doi.org/10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4)
- Moura, A. E., Janse van Rensburg, C., Pilot, M., Tehrani, A., Best, P. B., Thornton, M., ... Hoelzel, A. R. (2014). Killer whale nuclear genome and mtDNA reveal widespread population bottleneck during the last glacial maximum. *Molecular Biology and Evolution*, 31(5), 1121–1131. <https://doi.org/10.1093/molbev/msu058>
- Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during pleistocene revealed by whole-genome sequences. *Current Biology*, 25(10), 1375–1380. <https://doi.org/10.1016/j.cub.2015.03.047>
- Neff, J. A. (1937). Nesting distribution of the tri-colored red-wing. *The Condor*, 39(2), 61–81. <https://doi.org/10.2307/1363776>

- Neff, J. A. (1942). Migration of the tricolored red-wing in Central California. *The Condor*, 44(2), 45–53. <https://doi.org/10.2307/1364223>
- Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.
- Oh, K. P., Aldridge, C. L., Forbey, J. S., Dadabay, C. Y., & Oyler-McCance, S. J. (2019). Conservation genomics in the sagebrush sea: Population divergence, demographic history, and local adaptation in sage-grouse (*Centrocercus* spp.). *Genome Biology and Evolution*, 11(7), 2023–2034. <https://doi.org/10.1093/gbe/evz112>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ... Wagner, H. H. (2019). *vegan: Community Ecology Package*. R package version 2.5-6. R package version 1.17-2. R Package Version 2.5-4, 2. <https://CRAN.R-project.org/package=vegan>
- Oster, J. L., Ibarra, D. E., Winnick, M. J., & Maher, K. (2015). Steering of westerly storms over western North America at the Last Glacial Maximum. *Nature Geoscience*, 8(3), 201–205. <https://doi.org/10.1038/ngeo2365>
- Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J.-M., Basile, I., ... Stievenard, M. (1999). Climate and atmospheric history of the past 420,000 years from the Vostok ice core. *Antarctica. Nature*, 399(6735), 429–436. <https://doi.org/10.1038/20859>
- Pope, N. (2019). *r2vcftools: An R interface for vcftools (0.0.0.9000)* [R]. <https://github.com/npope/r2vcftools>
- Powell, A. F. L. A., Barker, F. K., Lanyon, S. M., Burns, K. J., Klicka, J., & Lovette, I. J. (2014). A comprehensive species-level molecular phylogeny of the New World blackbirds (Icteridae). *Molecular Phylogenetics and Evolution*, 71, 94–112. <https://doi.org/10.1016/j.ympev.2013.11.009>
- Richmond, J. Q., Barr, K. R., Backlin, A. R., Vandergast, A. G., & Fisher, R. N. (2013). Evolutionary dynamics of a rapidly receding southern range boundary in the threatened California Red-Legged Frog (*Rana draytonii*). *Evolutionary Applications*, 6(5), 808–822. <https://doi.org/10.1111/eva.12067>
- Robinson, O. J., Ruiz-Gutierrez, V., Fink, D., Meese, R. J., Holyoak, M., & Cooch, E. G. (2018). Using citizen science data in integrated population models to inform conservation. *Biological Conservation*, 227, 361–368. <https://doi.org/10.1016/j.biocon.2018.10.002>
- Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., ... Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23(23), 5726–5739. <https://doi.org/10.1111/mec.12977>
- Ruegg, K., Bay, R. A., Anderson, E. C., Saracco, J. F., Harrigan, R. J., Whitfield, M., ... Smith, T. B. (2018). Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters*, 21(7), 1085–1096. <https://doi.org/10.1111/ele.12977>
- Ryder, O. A. (1986). Species conservation and systematics: The dilemma of subspecies. *Trends in Ecology & Evolution*, 1, 9–10. [https://doi.org/10.1016/0169-5347\(86\)90059-5](https://doi.org/10.1016/0169-5347(86)90059-5)
- Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., ... Townshend, J. R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, 6(5), 427–448. <https://doi.org/10.1080/17538947.2013.786146>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Smeds, L., Qvarnström, A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome Research*, 26(9), 1211–1218. <https://doi.org/10.1101/gr.204669.116>
- Smith, T. B., & Bernatchez, L. (2008). Evolutionary change in human-altered environments. *Molecular Ecology*, 17(1), 1–8. <https://doi.org/10.1111/j.1365-294X.2007.03607.x>
- Storey, J. D., Bass, A. J., Dabney, A., Robinson, D., & Warnes, G. (2019). *qvalue: Q-value estimation for false discovery rate control* [R package version 2.16.0]. Retrieved from <http://github.com/jdstorey/qvalue>
- Sugg, D. W., & Chesser, R. K. (1994). Effective population sizes with multiple paternity. *Genetics*, 137(4), 1147–1155.
- Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, 25(10), 2144–2164. <https://doi.org/10.1111/mec.13606>
- Tricolored Blackbird Working Group. (2007). Conservation Plan for the Tricolored Blackbird (*Agelaius tricolor*). S. Kester (ed.). San Francisco, CA: Sustainable Conservation.
- Vucetich, J. A., Waite, T. A., & Nunney, L. (1997). Fluctuating population size and the ratio of effective to census population size. *Evolution*, 51(6), 2017–2021.
- Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, 197(2), 769–780. <https://doi.org/10.1534/genetics.114.164822>
- Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3), 244–262. <https://doi.org/10.1111/j.1752-4571.2009.00104.x>
- White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, 22(11), 2971–2985. <https://doi.org/10.1111/mec.12343>
- Wilson, C. R., Meese, R. J., & Wyckoff, A. C. (2016). Breeding chronology, movements, and life history observations of Tricolored Blackbirds in the California Central Coast. *California Fish and Game*, 102(4), 162–174.
- Xenikoudakis, G., Ersmark, E., Tison, J.-L., Waits, L., Kindberg, J., Swenson, J. E., & Dalén, L. (2015). Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear. *Molecular Ecology*, 24(13), 3441–3454. <https://doi.org/10.1111/mec.13239>
- Younger, J. L., Clucas, G. V., Kao, D., Rogers, A. D., Gharbi, K., Hart, T., & Miller, K. J. (2017). The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Molecular Ecology*, 26(15), 3883–3897. <https://doi.org/10.1111/mec.14172>
- Zalasiewicz, J., Williams, M., Haywood, A., & Ellis, M. (2011). The Anthropocene: A new epoch of geological time? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1938), 835–841. <https://doi.org/10.1098/rsta.2010.0339>
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Barr K, Beichman AC, Kalhori P, et al. Persistent panmixia despite extreme habitat loss and population decline in the threatened tricolored blackbird (*Agelaius tricolor*). *Evol Appl*. 2021;14:674–684. <https://doi.org/10.1111/eva.13147>