Received: 2 August 2023 | Revised: 18 October 2023 | Accepted: 27 October 2023

DOI: 10.1111/mec.17199

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genomics-informed conservation units reveal spatial variation in climate vulnerability in a migratory bird

Caitlin V. Miller[1] | Christen M. Bossu[1] | James F. Sarraco[2] | David P. L. Toews[3] | Clark S. Rushing[4] | Amélie Roberto-Charron[5] | Junior A. Tremblay[6] | Richard B. Chandler[4] | Matthew G. DeSaix[1] | Cameron J. Fiss[7] | Jeff L. Larkin[7] | Samuel Haché[8] | Silke Nebel[9] | Kristen C. Ruegg[1]

[1]Department of Biology, Colorado State University, Fort Collins, Colorado, USA

[2]The Institute for Bird Populations, Petaluma, California, USA

[3]Department of Biology, Pennsylvania State University, State College, Pennsylvania, USA

[4]Warnell School of Forestry and Natural Resources, University of Georgia, Athens, Georgia, USA

[5]Department of Environment, Government of Nunavut, Iqaluit, Canada

[6]Wildlife Research Division, Environment and Climate Change Canada, Québec, Quebec, Canada

[7]Department of Biology, Indiana University of Pennsylvania, Indiana, Pennsylvania, USA

[8]Canadian Wildlife Service, Environment Climate Change Canada, Yellowknife, Northwest Territories, Canada

[9]Birds Canada, Port Rowan, Ontario, Canada

Correspondence
Caitlin V. Miller, Department of Biology, Colorado State University, Fort Collins, Colorado, USA.
Email: mcaitlinv@gmail.com

Funding information
CanSeq150; Environment Climate Change Canada; National Science Foundation, Grant/Award Number: 1942313

Handling Editor: Jason Bragg

Abstract

Identifying genetic conservation units (CUs) in threatened species is critical for the preservation of adaptive capacity and evolutionary potential in the face of climate change. However, delineating CUs in highly mobile species remains a challenge due to high rates of gene flow and genetic signatures of isolation by distance. Even when CUs are delineated in highly mobile species, the CUs often lack key biological information about what populations have the most conservation need to guide management decisions. Here we implement a framework for CU identification in the Canada Warbler (*Cardellina canadensis*), a migratory bird species of conservation concern, and then integrate demographic modelling and genomic offset to guide conservation decisions. We find that patterns of whole genome genetic variation in this highly mobile species are primarily driven by putative adaptive variation. Identification of CUs across the breeding range revealed that Canada Warblers fall into two evolutionarily significant units (ESU), and three putative adaptive units (AUs) in the South, East, and Northwest. Quantification of genomic offset, a metric of genetic changes necessary to maintain current gene–environment relationships, revealed significant spatial variation in climate vulnerability, with the Northwestern AU being identified as the most vulnerable to future climate change. Alternatively, quantification of past population trends within each AU revealed the steepest population declines have occurred within the Eastern AU. Overall, we illustrate that genomics-informed CUs provide a strong foundation for identifying current and future regional threats that can be used to inform management strategies for a highly mobile species in a rapidly changing world.

KEYWORDS
climate change, conservation units, genomic offset, genomics, isolation by distance, population decline

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Recent estimates of biodiversity loss suggest that up to 60% of animal species are at risk of decline (Grooten & Almond, 2018), leading to an urgent need to identify and conserve threatened species. Conservation efforts to stop biodiversity loss focus on preserving biodiversity at the ecosystem, species, and genetic levels (Coates et al., 2018). While ecosystem and species level protections have historically been easier to quantify and administer, maintenance of genetic diversity is equally important for long-term ecosystem viability (Exposito-Alonso et al., 2022; Ralls et al., 2018; Ruegg & Turbek, 2022). Species with low or declining genetic diversity are threatened by inbreeding depression (Frankham, 2003) and the loss of adaptive capacity (Thurman et al., 2020), which may lead to high extinction risk (Forester et al., 2022). With climate change further accelerating biodiversity loss across biological scales (Dale et al., 2001), it is increasingly important to maintain genetic diversity within vulnerable populations to allow them to adapt. However, identifying populations most vulnerable to climate change and developing strategies for protecting them is not always straightforward.

Current strategies to protect populations or species based on genetic diversity often rely on the designation of intraspecific conservation units (CUs) to guide conservation and management decisions (Paetkau, 1999). There are many approaches to designating CUs, depending on the conservation priorities for the species. One of the most recognizable intraspecific CUs is the evolutionarily significant unit (ESU). ESUs are generally designated as evolutionarily or ecologically distinct populations within a broader species, though exact definitions may vary (de Guia & Saitoh, 2007; Moritz, 1994; Ryder, 1986; Waples, 1991). Adaptive units (AUs) are intraspecific groups that share similar adaptive traits and represent groups both within and across ESUs that are adapted to similar environments (Barbosa et al., 2018; Funk et al., 2012). Historically, CU designation has focused on ESUs and management units (MUs; smaller demographically independent populations inside ESUs), as AUs were difficult to define due to limited genomic tools needed to identify adaptive genetic markers (Luikart et al., 2003). However, with the advent of high-throughput sequencing and landscape genomic methods, it is now possible to identify putatively adaptive loci, how they are linked with environmental variation, and how changing climate conditions may affect these gene–environment relationships. While each type of CU is important for preserving different aspects of genetic diversity, here we focus on ESUs and AUs to investigate regional variation in climate change responses due to evolutionary isolation and adaptive differentiation.

Identifying CU boundaries that rigorously integrate key biological information critical to conservation in the face of climate change is not always straightforward. One issue that often arises but has not always been adequately addressed is the need to identify CUs in highly mobile species. Establishing clear CU boundaries in organisms with high capacity for dispersal (e.g. many migratory birds, bats and marine organisms) using genomics alone can be difficult because high levels of dispersal can lead to high gene flow between nearby populations. High gene flow can result in a signature of isolation by distance- where increasing distance correlates with decreased genetic similarity. Isolation by distance can make it difficult to differentiate CUs despite clear genetic variation throughout a species' range (Kekkonen et al., 2011; Palumbi, 1994; Veith et al., 2004). However, recent work in Turbek et al. (2023) has conceptualized a framework to 'draw the lines' to define CUs in highly mobile species. In brief, the framework uses contemporary genomic methods to define ESUs, such as principal component analysis and hierarchical models such as ADMIXTURE, but when isolation by distance is present defines ESUs using breakpoints in isolation by distance. Then AUs are defined using genomic population structure of only the adaptive loci. This framework also explicitly includes options for using alternate data sources, such as demographic or ecological data, when genomic methods alone do not provide CU resolution. Here we implement this framework for rigorous CU identification in the Canada Warbler (*Cardellina canadensis*), a migratory bird species of conservation concern, and demonstrate how key biological information (e.g. putative adaptive variation and population demography) can be integrated within the CU framework to guide conservation decisions at a regional level.

The Canada Warbler is a migratory songbird whose breeding range extends from Northwestern Canada to the Southeastern United States. Populations across the breeding range have declined 1.9% per year on average from 1966 to 2019 (Sauer et al., 2020). Currently, Canada Warblers have federal protection in Canada under the Species at Risk Act (COSEWIC, 2008) but are considered Least Concern under IUCN red-list designation (BirdLife International, 2021) partially due to their large range and heterogeneous declines. Canada Warblers, despite their potentially high dispersal capacity, exhibit high breeding site fidelity which could lead to genomic population structure across the breeding range (Hallworth et al., 2008). Previous genetic research, using eight microsatellite markers from three breeding sites, found that birds in the Southern portion of the range were genetically distinct from birds in the Eastern and Northwestern portions of the range, but the Eastern and Northwestern birds were not distinct from each other (Ferrari et al., 2018). There is a clear need for genetic information about population structure within the species, whether declines have been focused in areas that contain unique genetic diversity, and which populations are likely to be most vulnerable to changing climate conditions.

We used whole-genome resequencing to examine population structure across the Canada Warbler breeding range and identified putatively adaptive loci and neutral loci. We used the framework proposed by Turbek et al. (2023) to guide CU designation in highly mobile organisms. In addition, we assessed where management interventions would be most important by quantifying abundance and trend with demographic data, and estimating which populations may face the most climate-related vulnerability due to gene–environment mismatch (i.e. genomic offset; Capblancq et al., 2020; Fitzpatrick & Keller, 2015; Rellstab, 2021; Ruegg et al., 2018). The resulting data provide a framework for integrating CU designations with estimates

of climate vulnerability to improve our ability to identify and manage at-risk populations in a changing world.

2 | METHODS

2.1 | Reference genome

To create a reference genome, we captured a male Canada Warbler (record #SF12T03) to obtain blood for a high molecular weight (HMW) DNA sample. We affixed an aluminium band (#284029445) and took standard measurements, then drew ~10 μL of blood with a capillary tube from a brachial vein puncture. Using the blood, we extracted DNA using the Qiagen MagAttract HMW DNA Mini Kit (cat. no. 67563) with minor modifications to the standard elution protocol. We found that, likely because avian blood is nucleated, the DNA became tightly bound to the beads and the standard elution protocol would not yield sufficient HMW DNA. Instead, we eluted the DNA in 200 μL of water and left it on the mixer (at low speed) for approximately 1 h.

Using the HMW DNA obtained for the reference genome, we used 10X linked read sequencing to generate a whole-genome reference sequence of a Canada Warbler. Sequencing was part of the 'CanSeq150' project (https://www.cgen.ca/canseq150-project-list) in partnership with Birds Canada/Oiseaux Canada. 10× Genomics libraries were prepared at The Centre for Applied Genomics at The Hospital for Sick Children (Toronto, Canada) and libraries were sequenced on a HiSeq X machine (Illumina, San Diego) lane, with 150-bp paired-end reads. We assembled the reference genome using Supernova v2.1.1 (Weisenfeld et al., 2017) with default settings, except setting maximum reads to use all reads (485 M), on the Pennsylvania State University's Institute for Computational Data Sciences' Roar supercomputer. We used Benchmarking Universal Single-Copy Orthologs (BUSCO) to assess genome functional completeness (Simão et al., 2015) against the passeriformes_odb10 database.

2.2 | Resequencing sample collection and DNA extraction

We collected samples from an additional 181 breeding adult Canada Warblers from across the breeding range in North America in collaboration with multiple university researchers, private environmental companies and state and federal agencies (Figure S1). For DNA extraction, we collected blood from 134 Canada Warblers (~80 μL), via brachial venipuncture and preserved it in Queen's lysis buffer and stored at room temperature. Blood (50–80 μL) was extracted using Qiagen DNeasy Blood and Tissue Kits (QIAGEN) and eluted into 100 μL of provided AE buffer. For the remaining 47 Canada Warblers, we collected tail feathers by pulling two tail feathers from each bird and storing feathers at −20°C. We cut the calamus of one feather from the shaft and extracted the calamus using the modified Qiagen DNeasy Blood and Tissue protocol (Schweizer & DeSaix, 2023). After DNA extraction, we quantified samples using Qubit dsDNA assay.

2.2.1 | DNA resequencing

We prepared the samples for low coverage whole genome sequencing using a modified Nextera prep (Schweizer & DeSaix, 2023) with normalized DNA input. We sequenced samples in two libraries, 110 samples on an Illumina HiSeq 4000 using paired end 150 bp reads and 71 samples on an Illumina NovaSeq 6000 using paired end 150 bp reads. The 71 samples on the NovaSeq were sequenced across multiple lanes to get to the targeted sequencing depth of 2–3X coverage per sample (for sequencing scheme, see Table S1) and included replicates of 32 samples with lower than 1.5X coverage from the HiSeq 4000 run.

2.2.2 | Bioinformatic processing

We used Conda v4.13.0 (Anaconda Documentation, 2020) environments to manage bioinformatic packages on the RMACC Summit supercomputer managed jointly by Colorado State University and University of Colorado, Boulder. To process raw fastqs from the 181 samples that underwent low coverage whole genome sequencing, we used Trim Galore v0.6.7 (Krueger, 2012), a wrapper for cutadapt v1.18 (Martin, 2011) and FastQC v0.11.9 (Andrews, 2010) to trim any remaining Illumina adaptors in the fastqs. Next, based on recommendations for low coverage data generated with NovaSeq platforms (Lou & Therkildsen, 2022), we performed a sliding window cut of the 3-prime end of the reads to remove low-quality tails, defined as four bases in a row with mean QUAL scores of <20, using fastp v0.22.0 (Chen et al., 2018). We checked fastqs for quality using FastQC and MultiQC v1.0.dev0 (Ewels et al., 2016) before and after trimming reads.

After processing raw fastqs, we aligned samples to the Canada Warbler reference genome using Burrows-Wheeler Alignment mem v0.7.17 (Li & Durbin, 2009). Then we added read group information using Picard v2.26.11 AddorReplaceReadGroups (*Picard Toolkit*, 2014/2019) and marked duplicate reads using samtools v1.11 markdup (Danecek et al., 2011) before merging samples with multiple bams. After merging bams, we checked sample coverage using bedtools v2.30.0 genomecov (Quinlan & Hall, 2010) and samples with <1X coverage were removed, leaving 169 Canada Warblers in the analysis.

We used the processed bams to call variants in two separate pipelines using GATK v4.2.5.0 HaplotypeCaller (McKenna et al., 2010) and BCFtools v1.15.1 mpileup (Danecek et al., 2021). Then, we stringently filtered the two variant sets using BCFtools, allowing only biallelic sites, a minor allele frequency of >5%, QUAL score of >30, and <10% missing across the 169 individuals. We intersected the filtered variant sets from HaplotypeCaller and mpileup to create a

bootstrapped high-quality variant set to use for base quality score recalibration. Using the intersected variants, we recalibrated the sample bams using GATK BaseRecalibrator and ApplyBQSR. With the recalibrated bams, we used HaplotypeCaller to call a recalibrated set of variants. Then we filtered the recalibrated variant set allowing only biallelic sites, a minor allele frequency of >5%, QUAL score of >30, and <20% missing data across the 169 samples.

Using the recalibrated, filtered variant set, we performed an exploratory analysis using R (R Core Team, 2022) and the package srsStuff (Anderson, 2020) to produce single-read sampling principal components analysis (PCA) of whole genome structure. We used single-read sampling because differences in the average coverage across samples can be mistaken for population structure on PCA in low coverage data (Lou & Therkildsen, 2022). Single read sampling equalizes coverage for all samples. Despite equalizing coverage, we found significant platform effects (Figures S2 and S3), where samples sequenced on different platforms have inherent bias that can be mistaken for population structure (for example of platform effects on low coverage data, see Lou & Therkildsen, 2022). We removed platform-associated variants from the dataset and proceeded with the analysis once samples no longer clustered in platform groups by PCA and RDA (for full methods to remove platform effects, see Supplemental Methods).

2.3 | ESU identification

To identify ESUs, we used the criteria set out in Turbek et al. (2023) as a guide to delineate where genetic discontinuities exist in a species with high gene flow. We decided to delineate ESUs based on population structure that was supported with two out of three complementary, but different, approaches to finding breaks in genetic variation across the breeding range: PCA for a model-free approach, ADMIXTURE for a hierarchical model and estimated effective migration surfaces (EEMS) to model potential barriers to gene flow (Figure 1). We investigated population structure using the filtered variant set after removing platform effects. We first used the package srsStuff (Anderson, 2020) to produce single-read sampling principal components analysis (PCA) of whole genome structure. We evaluated the first six PCs and chose to represent the data with PC1 and PC2 due to the low overall additional variation explained after PC2.

As called genotypes on low coverage data are low confidence and often result in missing data for any given SNP, we imputed missing genotypes using Beagle v4.1 (Browning et al., 2018) using the genotype probabilities from GATK. Using the imputed data, we then removed linked SNPs using linkage disequilibrium ($r > .5$) in PLINK v2.0 (Purcell et al., 2007) and further investigated the potential for population structure using the program ADMIXTURE (Alexander et al., 2009). We used 5 runs of ADMIXTURE with K values 1–6 with the full set of variants but different random seeds. In order to visualize the different values of K and identify the most supported value of K based on cross-validation, we used the R package pophelper (Francis, 2017).

To further investigate if structure within the PCA or ADMIXTURE was due to a subtle barrier to gene flow, we used EEMS to check for potential barriers to gene flow (Petkova et al., 2016). Using the imputed dataset and 200 demes to test for potential barriers to gene flow, we then created a raster of estimated effective migration rates, which was mapped across the breeding grounds to visualize potential barriers to gene flow on the breeding range.

2.4 | AU identification

Similarly to ESUs, we used the criteria set out in Turbek et al. (2023) for guidance in the delineation of AUs from a landscape genetics approach. We defined AUs as breaks in putative adaptive genetic variation across the breeding range using only the adaptive loci set (see below) with two complementary methods: PCA for a model-free approach and ADMIXTURE for a hierarchical model (Figure 1). To select environmental variables, we used gradient forest (Ellis et al., 2012), an extension of random forest (Liaw & Wiener, 2002) and 23 environmental variables potentially important to Canada Warbler breeding ecology based on previous research (Table S3, Ferrari et al., 2018; Reitsma et al., 2020). While gradient forest has

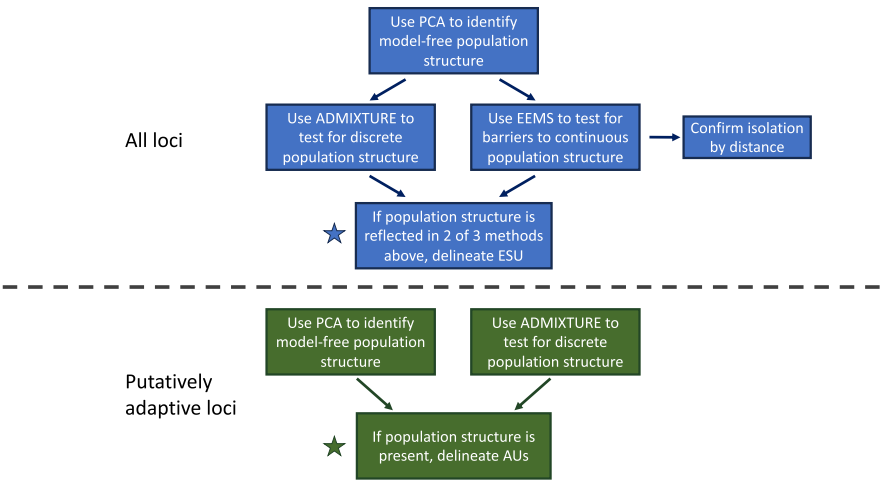

FIGURE 1 Overview of the methods and criteria used to delineate evolutionarily significant units (ESUs) and adaptive units (AUs). Boxes represent separate steps used to define each unit adapted from Turbek et al. (2023). Starred boxes indicating criteria specific to our delineation of ESUs and AUs.

been previously used to identify adaptive loci (Bay et al., 2018), we decided to use it to inform our environmental variable choice by using the highest-ranked environmental variables. We did not use gradient forest to identify adaptive loci as neutral population structure can confound gradient forest methods which was likely within a species with genetic patterns consistent with isolation by distance (Láruson et al., 2022). In this case, we chose not to use proxy variables for population structure, as we were not identifying loci themselves but only environmental variables. Environmental data were extracted from each of 16 sampling locations, excluding two sampling sites with fewer than four individuals. We used ANGSD v0.935 (Korneliussen et al., 2014) to calculate the allele frequency for the included sampled sites from genotype likelihood estimation using only SNPs with a minor allele frequency of >5% and removed SNPs that had missing data for any sampled site. Gradient forest was run with the R package gradientForest (Ellis et al., 2012) on five different subsets of 50,000 random SNPs using the environmental variables as predictors for the genomic data (ntree = 500, nbin = 101, corr. threshold = 0.5). To ensure that the models inferred from the data explained more than could be expected by random chance, we used 100 different randomizations of the data to create random models. Using these random models, we compared the distribution of randomized r-squared values of the SNPs to the r-squared values for the five models inferred from the data and ensured the models inferred from the data were above the 95th percentile of the r-squared values taken from the random models (Figure S5). We then chose the top four uncorrelated ($|r| < 0.75$) environmental variables ranked as most important to explaining genetic variation shared across the five models inferred from the data. These environmental variables were used as a reduced variable set for the rest of the analyses: mean temperature of the warmest quarter (BIO10), precipitation of the wettest month (BIO13), precipitation seasonality (BIO15) and tree cover.

To identify putatively adaptive loci, we used two approaches, redundancy analysis (RDA) and latent factor mixed models (LFMM). LFMM is a univariate approach that controls for population structure with latent factors, while RDA is a multivariate constrained ordination approach that performs better at finding many loci of small effect (Forester et al., 2018). Both approaches find SNPs associated with differences in environment at each individual's given latitude and longitude. To account for population structure in our RDA, we generated spatial variables using Moran's eigenvector maps (MEMs) (Dray et al., 2006) using the R package adespatial v0.3-16 (Guénard & Legendre, 2022). Then we ran the RDA using the R package vegan with individual genotypes as the response and the reduced environmental variable set as the predictors, conditioned on the MEMs to account for underlying population structure and geographic distance. We used the axes RDA1 and RDA2 with variations of 41.5% and 20.38% respectively. We selected loci that were above three standard deviations away from the mean.

We then used LFMM to find putatively adaptive loci by an alternate method. To account for population structure in our LFMM, we used $K = 3$. We chose $K = 3$ from the screeplot of the environmental PCs suggesting $K = 3$ was appropriate, as well as the initial PCA of whole genome population structure that showed subtle structure within the dataset with sites in the Northwest, East, and South appearing to cluster weakly together (see Figure 1). We ran LFMM using the R package lfmm (Jumentier, 2021) with individual genotypes as the response and, as LFMM is a univariate test, used the first principal component of the reduced environmental predictors (56.26%) to reduce the need for multiple corrections due to multiple tests. Using LFMM best practices (Forester et al., 2018), we adjusted an initial genomic inflation factor of >2.5 to 1.0 and identified loci using a false discovery rate of 1%.

Once loci were identified using both RDA and LFMM, the union of loci discovered by both methods was used as our set of candidate adaptive loci. To identify population genetic structure among the putatively adaptive loci we used PCA and ADMIXTURE. The resulting posterior probabilities of genetic group membership estimated from ADMIXTURE were visualized as transparency levels of different colours overlaid and clipped to a map of the Canada Warbler breeding range using the R packages sp, RGDAL and RASTER (Fick & Hijmans, 2017; Pebesma & Bivand, 2005) creating a spatially explicit map of putatively adaptive groups.

2.5 | Testing for isolation by distance versus isolation by environment

To determine if the geographically relevant clustering in the PCA and ADMIXTURE plot was a result of geography or environment, we used a combination of mantel tests and partial mantel tests. We generated pairwise F_{ST} comparisons using all loci between sites with at least four individuals, excluding two sampling sites with fewer than four individuals, using ANGSD v0.935. We calculated the site allele frequency (SAF) likelihoods for each sampled site from genotype likelihood estimation using only SNPs with a minor allele frequency of >5% and <30% missingness within the sampling site. We then calculated the 2D site frequency spectrum (SFS) for each pair of sites and, with the per-site SAF files as priors, we estimated pairwise F_{ST} between each sampled site. We then linearized $F_{ST}\left(\frac{F_{ST}}{1-F_{ST}}\right)$values for each pairwise comparison. We calculated pairwise Euclidean distance between each site's latitude and longitude using the R package sp (Pebesma & Bivand, 2005). To determine if genetic variation was more closely linked to environment or geography, we extracted environmental values for the reduced environmental variable set for each site's latitude and longitude. Then we centred each environmental variable to control for differences in absolute values of each variable, and calculated a pairwise environmental distance using the R package stats (R Core Team, 2022). We tested for isolation by distance and isolation by environment with Mantel tests in the R package vegan v2.6-2 (Oksanen et al., 2022) and partial Mantel tests conditioned on environmental distance and geographic distance respectively.

2.6 | Genomic offset analysis

Genomic offset is the magnitude of change necessary to keep gene–environment associations the same given changing climate conditions (Capblancq et al., 2020; Fitzpatrick & Keller, 2015). Higher genomic offset indicates that populations will need to change allele frequencies more to maintain current gene–environment relationships in future conditions. Though some caution should be used when evaluating genomic offset metrics due to the assumptions of the model (for a thorough discussion of assumptions and limitations, see Ahrens et al., 2023; Capblancq et al., 2020; DeSaix et al., 2022; Rellstab, 2021), here we use genomic offset to identify which AUs may be the most at risk of climate-related vulnerability. Using the adaptive loci found with LFMM and RDA as our response, we ran gradient forest (Ellis et al., 2012) and used the reduced environmental variable set as predictors to generate a model of allele frequency turnover across the breeding range. We used this model as a baseline to predict expected allele frequencies in 2061–2080 using the predicted environmental raster values for the years 2061–2080 under Shared-Socioeconomic Pathways 126 and 585 (SSP126 and SSP585) at 100,000 random points throughout the breeding range. We chose SSP126 and SSP585 as representative of the lowest and highest warming scenarios (Hausfather, 2019) as predicted from the Coupled Model Intercomparison Projects. We calculated the genomic offset between current allele frequencies and predicted future allele frequencies using Euclidean distance (Bay et al., 2018). Given the inherent uncertainty in predicting if or where range shifts will occur (Sofaer et al., 2018), we did not predict potential gene–environment associations or genomic offset outside of the current breeding range. Using the spatially explicit map of adaptive groups, we created shapefiles of each of the putative adaptive groups identified across the breeding range. We then extracted genomic offset values for SSP585 inside the boundaries of each adaptive group shapefile and calculated the median genomic offset within each adaptive group. We also extracted the genomic offset values across the entire breeding range and calculated the median genomic offset.

2.7 | Demographic analysis

We estimated relative population size indices and population trends from 1968 to 2019 for each of the three AUs and for all AUs combined based on the hierarchical over-dispersed Poisson model of Sauer and Link (2011) applied to Breeding Bird Survey data (Pardieck et al., 2020). While there are alternative data sources for estimating trends in migratory birds (e.g. eBird), the BBS provides the longest-running avian monitoring dataset for North American breeding birds, and shorter-term trend analyses produce similar results for broadly distributed species using either BBS or unstructured community-science data (Barker et al., 2015; Horns et al., 2018). The fixed strata effects in the model were defined based on the AUs, with BBS routes assigned to AUs if they ever had a Canada Warbler detection on the route and if the coordinates of the route starting point were contained within the AU polygon boundary. In addition, routes with Canada Warbler detections that were outside of AU polygons but within a 50-km buffer of an AU boundary were assigned to the nearest AU. Population size indices were derived by summarizing posterior distributions of mean route-level counts weighted by AU area and proportions of routes with Canada Warbler detections (Sauer & Link, 2011). We estimated population size indices for each AU by summarizing posterior distributions over the most recent 5 years (2015–2019). Long-term trends for each AU and for the overall population (based on the summed population indices across AUs) were estimated as the geometric mean of yearly changes in population size from 1968 to 2019 (Sauer & Link, 2011). We implemented the BBS model with JAGS 4.3.1 (Plummer, 2003) via the jagsUI (Kellner & Meredith, 2021) package in R (R Core Team, 2022). We assigned vague prior distributions for all model parameters and hyperparameters. Posterior distributions were derived from 40,000 simulated values of four chains from the posterior distribution after an adaptive phase of 20,000 iterations and burn-in of 10,000 samples of the Gibbs sampler and thinning by 3. Markov chains were determined to have successfully converged based on $\hat{R} < 1.1$ for posterior estimates of all parameters (Gelman & Hill, 2007). We also derived estimates of population size indices in the year 2069, assuming the current estimated trend and annual variance remain constant, based on posterior samples from the model at a time point 50 years in the future.

3 | RESULTS

3.1 | Reference genome

From the 10X sequencing libraries for genome assembly, we obtained 485.36 million reads. We produced an assembly with a 'raw' coverage of 48.6X, a scaffold N50 size of 7.51 Mb and genome size of 1.03 Gbp as estimated by *Supernova* for 3.04 $\times 10^3$ scaffolds >10 kb. Complete BUSCOS found in the assembly totalled 93%, with 91% complete and single copy, indicating high completeness. The genome assembly is deposited at NCBI with accession number PRJNA689308.

3.2 | Bioinformatics processing for DNA resequencing

Of the 181 samples from 18 different sites across the breeding range in our initial dataset (Figure S1), 12 samples with <1X average coverage were removed as part of quality filtering, leaving 169 total samples (Table S2). The median number of samples per site was 7.5 (range 3–22), with an average depth of coverage of 2.6X (range 1–22X). After filtering out low-quality SNPs and indels, we found 672,053 variants. After filtering for platform effects (Figure S2), we retained 654,226 SNPs.

3.3 | ESU identification

Using PCA, we found there was a subtle population structure throughout the breeding range (Figure 2a). Samples were grouped into three regional clusters: the Northwest with sites from Alberta, the East with sites ranging from Manitoba to Pennsylvania, and the South with sites from West Virginia to North Carolina. Samples from the Northwest and South separated along PC2, with samples from the Eastern portion of the range falling between these clusters. Overall variation explained by the first two PC axes was low- 0.94% and 0.90% on PCs 1 and 2 respectively. We filtered out variants in linkage disequilibrium and retained 451,571 SNPs, then assessed population structure for values of *K* 1–6 using ADMIXTURE (Figure S4). Results from ADMIXTURE suggest that the most supported *K* was 1.

Using EEMS with the filtered variants, we found that there is a strong barrier to gene flow between the Southern population and the Eastern populations in the Pennsylvania/New York region (Figure 2b).

3.4 | AU identification

We found 4832 SNPs and 9212 SNPs associated with the environmental variables using RDA (Figure 3) and LFMM respectively. To create the set of putatively adaptive SNPs, we combined SNPs found by both methods for a dataset of 11,441 unique SNPs.

We used PCA with the putatively adaptive SNPs and found three potential clusters (Figure S6). We used ADMIXTURE to assess population structure for values of *K* 1–6 (Figure S7) and found the best supported *K* was 3 (CV = 0.33822) for putatively adaptive loci (Figure 4a). We used the best supported *K* to assign individuals to putative AUs and produced a spatially explicit map of putative AUs (Figure 4b). The AUs are grouped into three regional clusters: the Northwest extending from the western border of Manitoba to the Northwestern Territories, the East extending from the eastern edge of Manitoba to eastern seaboard and into Pennsylvania, and the South representing the Appalachian Mountains through North Carolina.

3.5 | Testing for isolation by distance or isolation by environment

Pairwise FST across all quality-filtered SNPs ranged from 0 to 0.02767 (Table S4). Mantel tests revealed a strong correlation between environment and genetics ($r = 0.5984$, $p = .001$), as well as geography and genetics ($r = .6699$, $p = .001$). When we used a partial mantel test, the correlation between environment and genetics did not remain significant when accounting for geography ($r = .2157$, $p = .1$), but the correlation between geography and genetics remained significant when accounting for environment ($r = .4256$, $p = .001$).

3.6 | Genomic offset analysis

Using the model of future climate under the emissions pathway in SSP585 that assumes the highest level of emissions pathways,

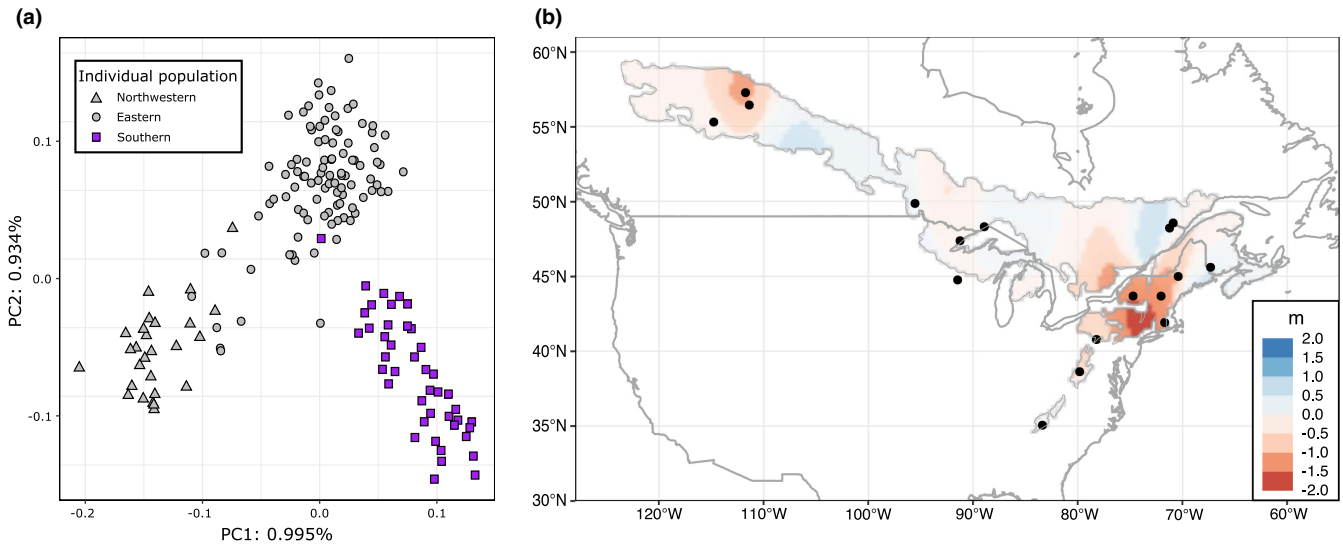

FIGURE 2 Population structure of Canada Warbler using whole genome loci of 654,226 single nucleotide polymorphisms (SNPs). (a) Principal components analysis representing whole genome structure. Points represent individual birds assigned to populations from latitude and longitude coordinates of capture site. Northwestern populations are grey triangles, Eastern populations are grey circles, while Southern populations in purple squares. (b) Estimated posterior mean migration rates on a log10 scale from estimated effective migration surfaces (EEMS). Areas with positive migration in blue are estimated to have greater gene flow than expected, while areas with negative migration in red are estimated to have less gene flow than expected. Transparency is scaled to reflect magnitude of estimated migration. Grey outline reflects the breeding range.

genomic offset was predicted to be highest in the Northern-most sections of the breeding range (Figure 5a). When genomic offset was assessed across putative adaptive groups, the Northwestern group had the highest predicted genomic offset, followed by the Eastern and Southern groups respectively (Figure 5b).

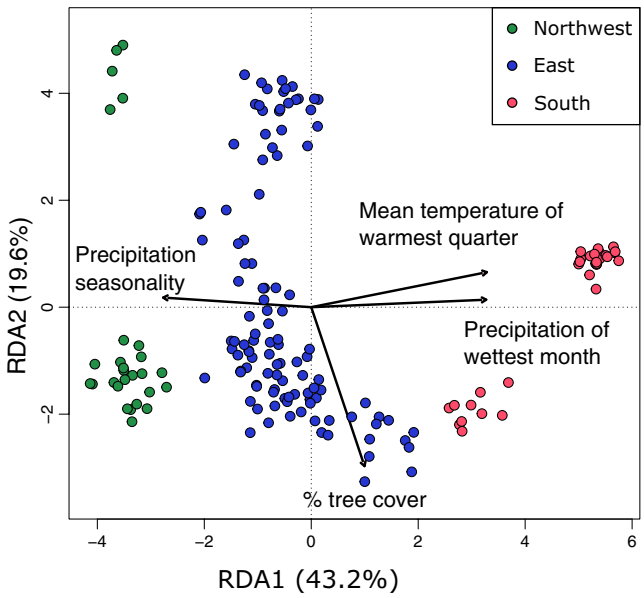

FIGURE 3 Principal component analysis of redundancy analysis axes 1 and 2 to delineate adaptive units (AUs). Coloured points are individuals sampled with the Northwest AU in green, the Eastern AU in blue, and the Southern AU in pink. Arrows represent the magnitude and direction of environmental variables.

3.7 | Demographic analysis

We used 819 BBS routes to estimate that the breeding range had a range-wide declining trend of −2.05% per year (CI −2.49% to −1.58%) between 1968 and 2019. We split the breeding range into the three putative adaptive units, with 28 routes in the Northwestern AU, 748 routes in the Eastern AU and 33 routes in the Southern AU. We found that the Eastern AU had the highest area-weighted abundance of 0.3763 (CI 0.3020–0.4598), followed by the Northwest AU with 0.0637 (CI 0.0363–0.1141), and then the Southern AU with 0.0037 (CI 0.0022–0.0059) (Figure 4c) based on recent population indices (2015–2019). Trends in abundance from 1968 to 2019 in the Southern AU (0.4416%; CI −0.9532% to 1.7961%) and Northwestern AU (−1.0682%; CI −2.8545% to 0.7831%) were not clearly positive or negative, but the Eastern AU (−2.1803%; CI −2.6276% to −1.7166%) had a strongly negative trend (Figure 4d). Predicted area-weighted abundance in 2069 was highly variable for both the Southern AU (0.0048; CI 0.0017–0.0130) and Northwestern AU (0.0363; CI 0.0098–0.1332), making it unclear if there will be declines or increases, but the Eastern AU (0.0866; CI 0.0609–0.1222) was predicted to decline steeply (Figure 4c).

4 | DISCUSSION

While preserving genetic diversity is important for the maintenance of current and future adaptive potential, defining CUs in a highly mobile species remains a challenge. Here, we demonstrate the value of a genomics-informed approach by identifying CUs in

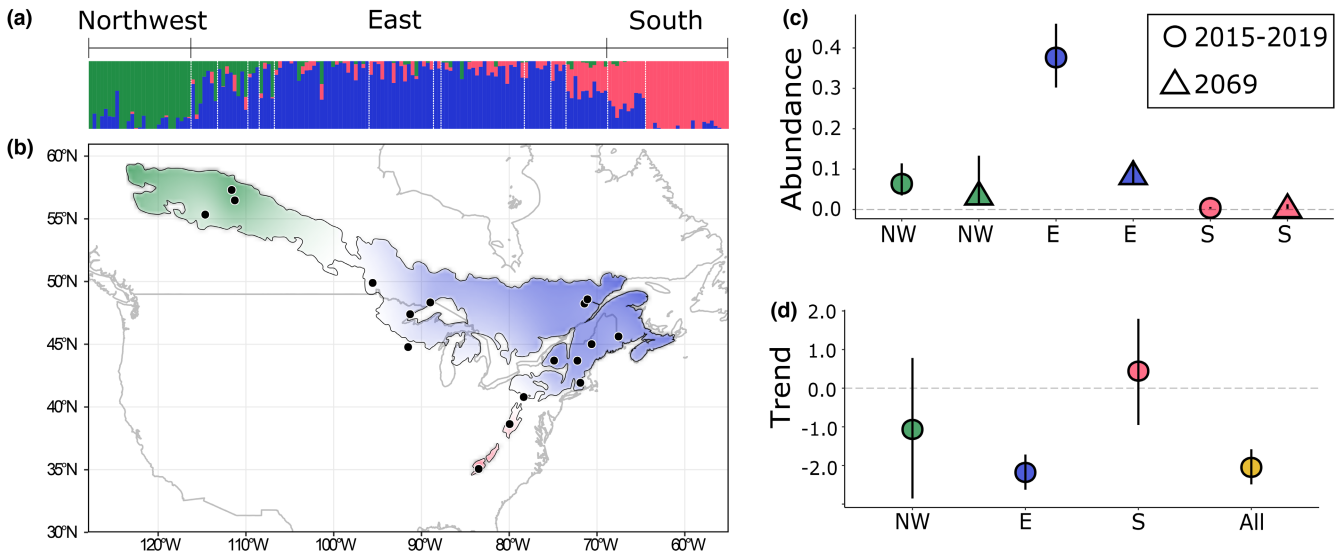

FIGURE 4 Putative adaptive units (AUs) of Canada Warbler using adaptive loci. Colours are used to represent AUs, green for the Northwestern AU, blue for the Eastern AU and pink for the Southern AU. (a) Best supported ADMIXTURE plot of *K* of 3 using only putatively adaptive loci. (b) Map of AU designations. Colours represent the AUs determined by ADMIXTURE groups, while points are sampled sites. Transparency is scaled to the predicted accuracy of assignment. (c) Area-weighted median abundance estimates in each AU. Points represent the median estimates for 1968–2019 and triangles represent predicted median estimates for 2069. (d) Estimated trend in area-weighted percent per year for each AU and range-wide calculated for 1968–2019.

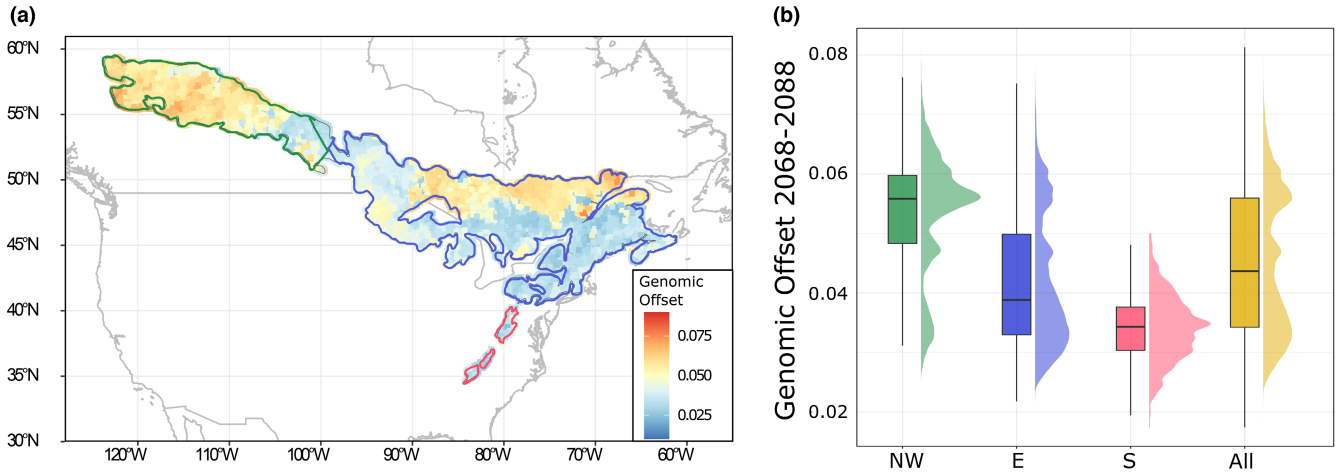

FIGURE 5 Predicted genomic offset across the Canada Warbler breeding range for 2061–2080 using shared socioeconomic pathway 585 (SSP585). Colours are used to represent AUs, green for the Northwestern AU, blue for the Eastern AU and pink for the Southern AU. (a) Map of predicted genomic offset at 100,000 random points across the breeding range. Coloured outlines represent the predicted AUs. (b) Box plots and density curves of genomic offset values for each AU and the entire range.

the Canada Warbler, a migratory songbird with a large range and heterogeneous declines. We found that overall genomic differentiation in Canada Warblers was low, with little population structure across the breeding range. In contrast, population structure at putatively adaptive loci was associated with significant differences in abundance, trend, and potential vulnerability to a changing climate as determined by estimates of genomic offset in the future. Overall, our results point more generally to the conclusion that genomics-informed CUs provide insight into region-specific trends across a broad species range and can help guide management in a changing world.

4.1 | ESU identification

Designating ESUs based on genomics can be challenging in species that remain highly connected across their range. Here, we use a recently described method for defining CUs in highly mobile species outlined by Turbek et al. (2023) and find that Canada Warblers fall into two ESUs. Specifically, analysis of all loci revealed that the South was distinct from the rest of the range, confirming results from a previous microsatellite-bases analysis of Canada Warblers (Ferrari et al., 2018). Although our ADMIXTURE analysis indicated a single ESU ($K=1$) with a very strong pattern of isolation by distance, further analysis with EEMS supported the existence of two ESUs. In particular, EEMs identified a potential barrier to gene flow in the Pennsylvania/New York region, which aligned with a previously documented ecological transition from high elevation breeding sites in the Southern Appalachian Mountains to lower elevation breeding sites in the Northeast (Howell, 1910). Overall, our results support the idea that a multifaceted approach for ESU identification can be instrumental in identifying barriers to gene flow in species characterized by limited population structure and strong isolation by distance patterns.

4.2 | AU identification

AUs are a relatively recent addition to CU delineation, as it has only recently been possible to sequence the large amount of genetic data necessary to identify putatively adaptive loci (Funk et al., 2012). For conservation and management of species with high motility, the utility of AUs lies in finding groups that share adaptive differences that may not have strong genetic structure otherwise (de Guia & Saitoh, 2007; Whitlock, 2014). Here, we analysed population genetic structure at putatively adaptive loci and found support for three distinct AUs within the Canada Warbler: a Northwestern, an Eastern, and a Southern AU, with evidence of admixture between the three AUs at areas transitioning from one AU to the next. Further, genetic variation in the Southern AU was associated with warmer mean temperatures, the Eastern AU was associated with higher amounts of precipitation during the wettest month, and the Northwestern AU was associated with high seasonality of precipitation. Because each AU is associated with distinct environmental parameters, identifying AUs on the breeding range provides a strong foundation for analysing how past and future environmental change may influence population trends within and between ecologically distinct regions.

While it has been recently suggested that whole genome structure could be used as proxy for adaptive variation without the need to identify putatively adaptive loci (Fernandez-Fournier et al., 2021), we found that analysing population structure at putatively adaptive loci separately allowed us to identify AUs that may otherwise have been overlooked. Specifically, when all loci were analysed together, the best supported K-value was one, but when putatively adaptive loci were analysed separately, the best supported K-value was three. These results were robust to randomizing training and test sets, suggesting that they are not a result of ascertainment bias (Anderson, 2010). The difference in the population structure results between all loci and adaptive loci is

likely because strong isolation by distance at neutral loci swamps signatures of population structure at adaptive loci when all loci are analysed together. Overall, our results support the idea that in highly mobile species with high gene flow, putatively adaptive variation may be the strongest signal of genetic differentiation (Yeaman & Whitlock, 2011).

4.3 | Identifying threats with genomics-informed CUs

Highly mobile species like birds, bats and fish often exhibit continuous genetic variation across space which has historically posed a challenge to identifying breakpoints for intraspecific CUs (Kekkonen et al., 2011; Palumbi, 1994; Veith et al., 2004). Here, we used genomics to identify CUs in the Canada Warbler and provide insight into understanding population declines and assessing vulnerability to future environmental threats. While across the range the species has been declining by ~2% per year since 1966 (Sauer et al., 2020), our study shows that segmenting the species' range into CUs reveals important spatial variation in population declines and abundance that may otherwise be overlooked. Among the three AUs, the Eastern AU was found to be the most abundant, but also predicted to decline by 77% by 2069. In contrast, the Southern and Northwestern AUs, which currently have smaller populations, are not predicted to decline as steeply. One possible explanation for this difference could be the lower number of routes associated with the South and Northwest, as fewer overall routes may lead to higher year to year variability adding uncertainty in our predictions. While the Eastern AU declines are the steepest, additional research into the vulnerability of the Southern AU may be important as it is located at the southernmost edge of the species distribution, where the effects of climate change are anticipated to be most severe (Lewis et al., 2023).

Our genomic offset analysis suggests that different CUs may be at risk from predicted climate changes than our demographic analyses. Previous work has found that genomic offset is highest in regions where current population declines are highest, and that this may be due to climate change already impacting populations (Bay et al., 2018; Ruegg et al., 2018). Interestingly, here we found a different pattern. Specifically, the AU showing the steepest population declines, the Eastern AU, had a moderate level of genomic offset when compared to the other AUs. A possible explanation for this result is that factors outside of the breeding range play a bigger role in explaining current population declines in this species. For example, previous research found that low recruitment and survival rates in Canada Warblers were linked to the nonbreeding period (Wilson et al., 2018), suggesting that changes to the wintering grounds may be driving breeding ground declines. While there are some studies connecting breeding ground populations to wintering ground populations using stable isotope data (González-Prieto et al., 2017; Wilson et al., 2018), future research linking breeding and wintering populations at finer spatial scales using genomic data could disentangle factors driving current population trends within each AU.

5 | CONCLUSION

In this study, we used a robust framework to delineate genomics-informed CUs for the Canada Warbler. Our analysis indicates that the population structure in this highly mobile species is low and is driven primarily by adaptive variation. Our results suggest that the Canada Warbler can be divided into two ESUs and three AUs that have spatial variation in both current declines and future climate vulnerability. Our findings suggest that one conservation management strategy could be to focus on addressing the current declines in the Eastern AU by increasing overwinter survival of juvenile birds on the wintering grounds. Alternately, maintaining or increasing habitat where populations have yet to decline steeply, such as the Northwestern and Southern AUs, could prevent declines in regions that contain unique putative adaptation. Furthermore, our research indicates that the effects of future climate change on the species range is expected to vary across regions, which may alter the current trends in decline. Overall, this work helps illustrate how combining the identification of CUs with future climate modelling can help identify populations in need of further protection to preserve future genetic diversity.

AUTHOR CONTRIBUTIONS

CVM and KCR conceptualized the study. KCR and SH acquired funding. ARC, JAT, RBC, KRF, MGD, CJF, JLL and TBS facilitated sample collection. SN facilitated the genome sequencing by the 'Canada 150 Sequencing Initiative'. CVM, DPLT, and CMB performed all genomic analyses. CVM, JFS and CSR performed the demographic analyses. CVM wrote the original manuscript draft and all authors participated in the review process.

ACKNOWLEDGEMENTS

This project was funded by Environment Climate Change Canada and an National Science Foundation CAREER grant (#1942313) to K. Ruegg, as well as sequencing as part of the 'CanSeq150' project. We thank Kevin Fraser, Len Reitsma, Nick Russo, Lori Walenski, Wolf Ridge Environmental Center, John Woodcock, the Institute for Bird Populations, and Owl Moon Environmental Inc. who contributed feather samples to the project. We also thank Eric Anderson for bioinformatics and coding support, Teia Schweizer and Christine Rayne for library preparation, and the Ruegg and Funk labs at Colorado State University for detailed feedback on early drafts of the manuscript. Computations for this research were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer. This content is solely the responsibility of the authors and does not necessarily represent the views of the Institute for Computational and Data Sciences. This work utilized the RMACC Summit supercomputer, which is supported by the National Science Foundation

(awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder and Colorado State University. The RMACC Summit supercomputer is a joint effort of the University of Colorado Boulder and Colorado State University.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Genomic data with associated individual metadata have been deposited in a Dryad repository https://doi.org/10.5061/dryad.wh70rxwt9 (Miller et al., 2023). All analysis scripts have been made available at https://github.com/mcaitlinv/cawa-breeding.

BENEFIT-SHARING STATEMENT

An international research collaboration was developed with scientists facilitating genetic sample collection and all collaborators are included as co-authors. The contributions of all individuals to the research are described in the acknowledgements. The results of the research have been shared with the broader scientific community and address a priority concern of the conservation of highly mobile species, in particular migratory birds.

ORCID

Caitlin V. Miller https://orcid.org/0000-0003-0850-1486
Christen M. Bossu https://orcid.org/0000-0002-0458-9305
David P. L. Toews https://orcid.org/0000-0002-9763-0476
Clark S. Rushing https://orcid.org/0000-0002-9283-6563
Amélie Roberto-Charron https://orcid.org/0000-0001-5917-4919
Junior A. Tremblay https://orcid.org/0000-0003-4930-0939
Richard B. Chandler https://orcid.org/0000-0003-4930-2790
Matthew G. DeSaix https://orcid.org/0000-0002-5721-0311
Cameron J. Fiss https://orcid.org/0000-0002-5649-6880
Samuel Haché https://orcid.org/0000-0003-3952-009X
Kristen C. Ruegg https://orcid.org/0000-0001-5579-941X

REFERENCES

Ahrens, C., Rymer, P., & Miller, A. (2023). *Genetic offset and vulnerability modelling: Misinterpretations of results and violations of evolutionary principles*. https://doi.org/10.22541/au.168727971.18670759/v1

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

Anaconda documentation. (2020). *Anaconda software distribution*. https://docs.anaconda.com/

Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly biased. *Molecular Ecology Resources*, *10*(4), 701–710. https://doi.org/10.1111/j.1755-0998.2010.02846.x

Anderson, E. C. (2020). *srsStuff: R package for inference by sampling single reads from sequencing data*. https://github.com/eriqande/srsStuff

Andrews, S. (2010). *FASTQC: A quality control tool for high throughput sequence data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barbosa, S., Mestre, F., White, T. A., Paupério, J., Alves, P. C., & Searle, J. B. (2018). Integrative approaches to guide conservation decisions: Using genomics to define conservation units and functional corridors. *Molecular Ecology*, *27*(17), 3452–3465. https://doi.org/10.1111/mec.14806

Barker, N. K. S., Fontaine, P. C., Cumming, S. G., Stralberg, D., Westwood, A., Bayne, E. M., Sólymos, P., Schmiegelow, F. K. A., Song, S. J., & Rugg, D. J. (2015). Ecological monitoring through harmonizing existing data: Lessons from the boreal avian modelling project. *Wildlife Society Bulletin*, *39*(3), 480–487. https://doi.org/10.1002/wsb.567

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

BirdLife International. (2021). *Cardellina canadensis. The IUCN Red List of Threatened Species*. https://doi.org/10.2305/IUCN.UK.2021-3.RLTS.T22721882A137213211.en

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, *103*(3), 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Coates, D. J., Byrne, M., & Moritz, C. (2018). Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*, *6*. https://doi.org/10.3389/fevo.2018.00165

COSEWIC. (2008). *COSEWIC assessment and status report on the Canada Warbler Wilsonia Canadensis in Canada. Committee on the Status of Endangered Wildlife in Canada* [Assessments;research]. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/canada-warbler.html

Dale, V. H., Joyce, L. A., McNulty, S., Neilson, R. P., Ayres, M. P., Flannigan, M. D., Hanson, P. J., Irland, L. C., Lugo, A. E., Peterson, C. J., Simberloff, D., Swanson, F. J., Stocks, B. J., & Wotton, B. M. (2001). Climate change and Forest disturbances: Climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. *Bioscience*, *51*(9), 723–734. https://doi.org/10.1641/0006-3568(2001)051[0723:CCAFD]2.0.CO;2

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

de Guia, A. P. O., & Saitoh, T. (2007). The gap between the concept and definitions in the evolutionarily significant unit: The need to integrate neutral genetic variation and adaptive variation. *Ecological Research*, *22*(4), 604–612. https://doi.org/10.1007/s11284-006-0059-z

DeSaix, M. G., George, T. L., Seglund, A. E., Spellman, G. M., Zavaleta, E. S., & Ruegg, K. C. (2022). Forecasting climate change response

in an alpine specialist songbird reveals the importance of considering novel climate. *Diversity and Distributions*, *28*(10), 2239–2254. https://doi.org/10.1111/ddi.13628

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*(3), 483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015

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

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Exposito-Alonso, M., Booker, T. R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C. C., Lang, P. L. M., Leventhal, L., Nogues-Bravo, D., Pagowski, V., Ruffley, M., Spence, J. P., Toro Arana, S. E., Weiß, C. L., & Zess, E. (2022). Genetic diversity loss in the Anthropocene. *Science*, *377*(6613), 1431–1435. https://doi.org/10.1126/science.abn5642

Fernandez-Fournier, P., Lewthwaite, J. M. M., & Mooers, A. Ø. (2021). Do we need to identify adaptive genetic variation when prioritizing populations for conservation? *Conservation Genetics*, *22*(2), 205–216. https://doi.org/10.1007/s10592-020-01327-w

Ferrari, B., Shamblin, B., Chandler, R., Tumas, H., Hache, S., Reitsma, L., & Nairn, C. (2018). Canada Warbler (*Cardellina canadensis*): Novel molecular markers and a preliminary analysis of genetic diversity and structure. *Avian Conservation and Ecology*, *13*(1), 8. https://doi.org/10.5751/ACE-01176-130108

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, *18*(1), 1–16. https://doi.org/10.1111/ele.12376

Forester, B. R., Beever, E. A., Darst, C., Szymanski, J., & Funk, W. C. (2022). Linking evolutionary potential to extinction risk: Applications and future directions. *Frontiers in Ecology and the Environment*, *20*, 507–515. https://doi.org/10.1002/fee.2552

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, *27*(9), 2215–2233. https://doi.org/10.1111/mec.14584

Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32. https://doi.org/10.1111/1755-0998.12509

Frankham, R. (2003). Genetics and conservation biology. *Comptes Rendus Biologies*, *326*, 22–29. https://doi.org/10.1016/S1631-0691(03)00023-4

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

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. In *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

González-Prieto, A. M., Bayly, N. J., Colorado, G. J., & Hobson, K. A. (2017). Topography of the Andes Mountains shapes the wintering distribution of a migratory bird. *Diversity and Distributions*, *23*(1/2), 118–129.

Grooten, M., & Almond, R. E. A. (2018). *Living planet report—2018: Aiming higher*. World Wildlife Fund. https://www.worldwildlife.org/pages/living-planet-report-2018?_hsenc=p2ANqtz-9DJ-2Fh_gaupLvz8xfsFchP9DBn_QUFDUj0kJd-3saW9lSZzAIgyEZQXoQrOrOtXTDwB2IQtHoWcp_XPv8qmftM4T5FA

Guénard, G., & Legendre, P. (2022). Hierarchical clustering with contiguity constraint in R. *Journal of Statistical Software*, *103*(7), 1–26. https://doi.org/10.18637/jss.v103.i07

Hallworth, M., Ueland, A., Anderson, E., Lambert, J. D., & Reitsma, L. (2008). Habitat selection and site Fidelity of Canada warblers (Wilsonia Canadensis) in Central New Hampshire. *The Auk*, *125*(4), 880–888. https://doi.org/10.1525/auk.2008.07115

Hausfather, Z. (2019). *CMIP6: The next generation of climate models explained*. Carbon Brief. https://www.carbonbrief.org/cmip6-the-next-generation-of-climate-models-explained/

Horns, J. J., Adler, F. R., & Şekercioğlu, Ç. H. (2018). Using opportunistic citizen science data to estimate avian population trends. *Biological Conservation*, *221*, 151–159. https://doi.org/10.1016/j.biocon.2018.02.027

Howell, A. H. (1910). Notes on the summer birds of Kentucky and Tennessee. *The Auk*, *27*(3), 295–304. https://doi.org/10.2307/4071313

Jumentier, B. (2021). *lfmm: Latent factor mixed models*. https://CRAN.R-project.org/package=lfmm

Kekkonen, J., Seppä, P., Hanski, I. K., Jensen, H., Väisänen, R. A., & Brommer, J. E. (2011). Low genetic differentiation in a sedentary bird: House sparrow population genetics in a contiguous landscape. *Heredity*, *106*(1), Article 1. https://doi.org/10.1038/hdy.2010.32

Kellner, K., & Meredith, M. (2021). *jagsUI: A Wrapper Around "rjags" to Streamline "JAGS" Analyses* (1.5.2) [Computer software]. https://CRAN.R-project.org/package=jagsUI

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Krueger, F. (2012). *TrimGalore: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data*. https://github.com/FelixKrueger/TrimGalore

Láruson, Á. J., Fitzpatrick, M. C., Keller, S. R., Haller, B. C., & Lotterhos, K. E. (2022). Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. *Evolutionary Applications*, *15*(3), 403–416. https://doi.org/10.1111/eva.13354

Lewis, W. B., Cooper, R. J., Chandler, R. B., Chitwood, R. W., Cline, M. H., Hallworth, M. T., Hatt, J. L., Hepinstall-Cymerman, J., Kaiser, S. A., Rodenhouse, N. L., Sillett, T. S., Stodola, K. W., Webster, M. S., & Holmes, R. T. (2023). Climate-mediated population dynamics of a migratory songbird differ between the trailing edge and range core. *Ecological Monographs*, *93*(1), e1559. https://doi.org/10.1002/ecm.1559

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Lou, R. N., & Therkildsen, N. O. (2022). Batch effects in population genomic studies with low-coverage whole genome sequencing data: Causes, detection and mitigation. *Molecular Ecology Resources*, *22*(5), 1678–1692. https://doi.org/10.1111/1755-0998.13559

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, *4*(12), Article 12. https://doi.org/10.1038/nrg1226

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), Article 1. https://doi.org/10.14806/ej.17.1.200

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & 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

Miller, C. V., Bossu, C. M., & Ruegg, K. C. (2023). *Genomics-informed conservation units reveal spatial variation in climate vulnerability in a migratory bird* [dataset]. Dryad. https://doi.org/10.5061/dryad.wh70rxwt9

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

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M. D., Durand, S., … Weedon, J. (2022). *vegan: Community ecology package*. https://CRAN.R-project.org/package=vegan

Paetkau, D. (1999). Using genetics to identify intraspecific conservation units: A critique of current methods. *Conservation Biology*, *13*(6), 1507–1509. https://doi.org/10.1046/j.1523-1739.1999.98507.x

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, *25*(1), 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555

Pardieck, K. L., David, Z. J., Lutmerding, M., Aponte, V., & Hudson, M.-A. R. (2020). *North American breeding bird survey dataset 1966–2019, version 2019.0*. [dataset]. U.S. Geological Survey. https://doi.org/10.5066/P9J6QUF6

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, *5*(2), 9–13.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. https://doi.org/10.1038/ng.3464

Plummer, M. (2003). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*. Working Papers.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Ralls, K., Ballou, J. D., Dudash, M. R., Eldridge, M. D. B., Fenster, C. B., Lacy, R. C., Sunnucks, P., & Frankham, R. (2018). Call for a paradigm shift in the genetic Management of Fragmented Populations. *Conservation Letters*, *11*(2), e12412. https://doi.org/10.1111/conl.12412

Reitsma, L. R., Hallworth, M. T., McMahon, M., & Conway, C. J. (2020). Canada Warbler (*Cardellina canadensis*), version 2.0. In P. G. Rodewald, & B. K. Keeney (Eds.), *Birds of the World*. Cornell Lab of Ornithology. https://doi.org.ezproxy2.library.colostate.edu/10.2173/bow.canwar.02

Rellstab, C. (2021). Genomics helps to predict maladaptation to climate change. *Nature Climate Change*, *11*(2), Article 2. https://doi.org/10.1038/s41558-020-00964-w

Ruegg, K., Bay, R. A., Anderson, E. C., Saracco, J. F., Harrigan, R. J., Whitfield, M., Paxton, E. H., & 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

Ruegg, K., & Turbek, S. (2022). Estimating global genetic diversity loss. *Science*, *377*(6613), 1384–1385. https://doi.org/10.1126/science.add0007

Ryder, O. A. (1986). Species conservation and systematics: The dilemma of subspecies. *Trends in Ecology & Evolution*, *1*(1), 9–10. https://doi.org/10.1016/0169-5347(86)90059-5

Sauer, J. R., & Link, W. A. (2011). Analysis of the north American breeding bird survey using hierarchical models. *The Auk*, *128*(1), 87–98. https://doi.org/10.1525/auk.2010.09220

Sauer, J. R., Link, W. A., & Hines, J. E. (2020). *The north American breeding bird survey, analysis results 1966–2019* [dataset]. U.S. Geological Survey. https://doi.org/10.5066/P96A7675

Schweizer, T. M., & DeSaix, M. G. (2023). Cost-effective library preparation for whole genome sequencing with feather DNA. *Conservation Genetics Resources*, *15*(1), 21–28. https://doi.org/10.1007/s12686-023-01299-2

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 (Oxford, England)*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sofaer, H. R., Jarnevich, C. S., & Flather, C. H. (2018). Misleading prioritizations from modelling range shifts under climate change. *Global Ecology and Biogeography*, *27*(6), 658–666. https://doi.org/10.1111/geb.12726

Thurman, L. L., Stein, B. A., Beever, E. A., Foden, W., Geange, S. R., Green, N., Gross, J. E., Lawrence, D. J., LeDee, O., Olden, J. D., Thompson, L. M., & Young, B. E. (2020). Persist in place or shift in space? Evaluating the adaptive capacity of species to climate change. *Frontiers in Ecology and the Environment*, *18*(9), 520–528. https://doi.org/10.1002/fee.2253

Picard Toolkit. (2019). [Java]. *Broad Institute*. (Original work published 2014) https://github.com/broadinstitute/picard

Turbek, S. P., Funk, W. C., & Ruegg, K. C. (2023). Where to draw the line? Expanding the delineation of conservation units to highly mobile taxa. *Journal of Heredity*, *114*(4), 300–311. https://doi.org/10.1093/jhered/esad011

Veith, M., Beer, N., Kiefer, A., Johannesen, J., & Seitz, A. (2004). The role of swarming sites for maintaining gene flow in the brown long-eared bat (*Plecotus auritus*). *Heredity*, *93*(4), Article 4. https://doi.org/10.1038/sj.hdy.6800509

Waples, R. S. (1991). Pacific Salmon, Oncorhynchus spp., and the definition of "species" under the endangered species act. *Marine Fisheries Review*, *53*(3), Article 3.

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, *27*(5), 757–767. https://doi.org/10.1101/gr.214874.116

Whitlock, R. (2014). Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: A meta-analysis. *The Journal of Ecology*, *102*(4), 857–872. https://doi.org/10.1111/1365-2745.12240

Wilson, S., Saracco, J. F., Krikun, R., Flockhart, D. T. T., Godwin, C. M., & Foster, K. R. (2018). Drivers of demographic decline across the annual cycle of a threatened migratory bird. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-25633-z

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration–selection balance. *Evolution*, *65*(7), 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

SUPPORTING INFORMATION

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

How to cite this article: Miller, C. V., Bossu, C. M., Sarraco, J. F., Toews, D. P. L., Rushing, C. S., Roberto-Charron, A., Tremblay, J. A., Chandler, R. B., DeSaix, M. G., Fiss, C. J., Larkin, J. L., Haché, S., Nebel, S., & Ruegg, K. C. (2023). Genomics-informed conservation units reveal spatial variation in climate vulnerability in a migratory bird. *Molecular Ecology*, *00*, 1–13. https://doi.org/10.1111/mec.17199