*SPECIAL FEATURE: Species Limits and Taxonomy in Birds*

# The American Kestrel (*Falco sparverius*) genoscape: Implications for monitoring, management, and subspecies boundaries

**Kristen C. Ruegg,**[1,2,#,*] **Michaela Brinkmeyer,**[3,#] **Christen M. Bossu,**[1,2] **Rachael A. Bay,**[2,4]
**Eric C. Anderson,**[5] **Clint W. Boal,**[6] **Russell D. Dawson,**[7] **Amber Eschenbauch,**[8] **Christopher J. W. McClure,**[9]
**Karl E. Miller,**[10] **Lance Morrow,**[11] **Jill Morrow,**[11] **M. David Oleyar,**[12] **Bill Ralph,**[13] **Sarah Schulwitz,**[9]
**Ted Swem,**[14] **Jean-Francois Therrien,**[15] **Rich Van Buskirk,**[16] **Thomas B. Smith,**[2,17] and **Julie A. Heath**[3]

[1]Biology Department, Colorado State University, Fort Collins, Colorado, USA
[2]Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, California, USA
[3]Department of Biological Sciences and Raptor Research Center, Boise State University, Boise, Idaho, USA
[4]Department of Evolution and Ecology, University of California Davis, Davis, California, USA
[5]Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, Santa Cruz, California, USA
[6]U.S. Geological Survey, Texas Cooperative Fish and Wildlife Research Unit, Texas Tech University, Lubbock, Texas, USA
[7]Ecosystem Science and Management Program, University of Northern British Columbia, Prince George, British Columbia, Canada
[8]Central Wisconsin Kestrel Research, Amherst, Wisconsin, USA
[9]The Peregrine Fund, Boise, Idaho, USA
[10]Fish and Wildlife Research Institute, Florida Fish and Wildlife Conservation Commission, Gainesville, Florida, USA
[11]Shenandoah Valley Raptor Study Area, Timberville, Virginia, USA
[12]HawkWatch International, Salt Lake City, Utah, USA
[13]Yosemite Area Audubon Society, Raymond, California, USA
[14]U.S. Fish and Wildlife Service, Fairbanks, Alaska, USA
[15]Hawk Mountain Sanctuary, Orwigsburg, Pennsylvania, USA
[16]Department of Environmental Studies, Pacific University, Forest Grove, Oregon, USA
[17]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, USA
[#]These authors contributed equally to this work.
[*]Corresponding author: kristen.ruegg@colostate.edu

## ABSTRACT

Identifying population genetic structure is useful for inferring evolutionary process and comparing the resulting structure with subspecies boundaries can aid in species management. The American Kestrel (*Falco sparverius*) is a widespread and highly diverse species with 17 total subspecies, only 2 of which are found north of U.S./Mexico border (*F. s. paulus* is restricted to southeastern United States, while *F. s. sparverius* breeds across the remainder of the U.S. and Canadian distribution). In many parts of their U.S. and Canadian range, American Kestrels have been declining, but it has been difficult to interpret demographic trends without a clearer understanding of gene flow among populations. Here we sequence the first American Kestrel genome and scan the genome of 197 individuals from 12 sampling locations across the United States and Canada in order to identify population structure. To validate signatures of population structure and fill in sampling gaps across the U.S. and Canadian range, we screened 192 outlier loci in an additional 376 samples from 34 sampling locations. Overall, our analyses support the existence of 5 genetically distinct populations of American Kestrels—eastern, western, Texas, Florida, and Alaska. Interestingly, we found that while our genome-wide genetic data support the existence of previously described subspecies boundaries in the United States and Canada, genetic differences across the sampled range correlate more with putative migratory phenotypes (resident, long-distance, and short-distance migrants) rather than a priori described subspecies boundaries per se. Based on our results, we suggest the resulting 5 genetically distinct populations serve as the foundation for American Kestrel conservation and management in the face of future threats.

*Keywords:* conservation unit, genoscape, migratory bird

**LAY SUMMARY**

- The American Kestrel (*Falco sparverius*) is a widespread iconic raptor species that has shown highly variable trends in abundance over the last several decades.
- Here we sequence the first American Kestrel genome and scan the genome for genetic variation in order to identify five genetically distinct populations across the U.S. and Canadian breeding range.
- The resulting map of genetic variation (the genoscape) can serve as a foundation for testing hypotheses to explain observed population-specific responses to climate change and other stressors.

**Paisaje genómico de *Falco sparverius*: Implicancias para monitoreo, manejo y límites entre subespecies**

**RESUMEN**
Identificar la estructura genética molecular es útil para inferir procesos evolutivos, y comparar la estructura resultante con los límites entre subespecies puede ayudar en el manejo de especies. *Falco sparverius* es una especie ampliamente distribuida y altamente diversa con 17 subespecies en total, solo dos de las cuales se encuentran al norte de la frontera de Estados Unidos y México (*F. s. paulus* está restringida al sudeste de Estados Unidos, mientras que *F. s. sparverius* cría a través del resto de la distribución en Estados Unidos y Canadá). En muchas partes de su rango de EEUU y Canadá, *F. sparverius* ha estado disminuyendo, pero ha sido difícil interpretar las tendencias demográficas sin un entendimiento más claro del flujo génico entre las poblaciones. Aquí secuenciamos el primer genoma de *F. sparverius* y escaneamos el genoma de 197 individuos a partir de 12 localidades de muestreo a través de EEUU y Canadá con el propósito de identificar la estructura poblacional. Para validar firmas de estructura poblacional y llenar agujeros de muestreo a través del rango de EEUU y Canadá, seleccionamos 192 loci atípicos usando 376 muestras adicionales provenientes de 34 localidades de muestreo. De forma global, nuestros análisis apoyan la existencia de 5 poblaciones genéticamente distintas de *F. sparverius*—Este, Oeste, Texas, Florida y Alaska. Curiosamente, encontramos que mientras nuestros datos genéticos para todo el genoma apoyan la existencia de los límites entre subespecies previamente descriptos en EEUU y Canadá, las diferencias genéticas a través del rango muestreado se correlacionan más con fenotipos migratorios putativos (residentes y migrantes de larga y corta distancia), que con los límites entre subespecies descritos a priori per se. Basados en nuestros resultados, sugerimos que las cinco poblaciones genéticamente distintas resultantes sirven como base para la conservación y el manejo de *F. sparverius* de cara a futuras amenazas.

*Palabras clave:* ave migratoria, paisaje genómico, unidad de conservación

## INTRODUCTION

An important application of population genetics is the identification of genetically distinct populations within species that can be used to guide conservation and management efforts. Depending on the context, such groups are often referred to as subspecies, management units (MUs), evolutionary significant units (ESUs), conservation units, or genetically distinct populations (Moritz 1994, Allendorf and Luikart 2007, Funk et al. 2007). Population genetic structure below the species level has frequently been used to delineate units for conservation and management (Moritz 1994, Allendorf and Luikart 2007, Funk et al. 2007), but other factors including behavior and morphological variations are also important, particularly in species for which genetic data are absent or lacking in resolution (Mayr 1982, Waples et al. 2007). In highly mobile species, it has historically been difficult to identify subspecies that correlate with genetically distinct populations because gene flow often homogenizes the diversifying effects of local adaptation and drift (Waples 1998, Willoughby et al. 2017, Doyle et al. 2018, Medina et al. 2018). As a result, identifying genetically distinct populations in migratory animals, such as migratory birds, remains a challenge (Larson et al. 2014, Zink 2014, Freer et al. 2015, Mura-Jornet et al. 2018).

Traditionally, genetic studies focused on identifying genetically distinct populations relied on a limited number of molecular markers (e.g., microsatellites, mitochondrial (mt) DNA sequences, and allozymes) to make inferences about population genetic structure (Ryman et al. 2006, Morin et al. 2009, Rowe et al. 2011, Mura-Jornet et al. 2018). However, recent advances in sequencing technology have made it possible to screen tens of thousands to millions of genetic markers and reveal patterns of population structure that may have previously gone undetected (Rowe et al. 2011). For many species, greatly increasing the number of loci included in population genetic analyses has improved the precision of population genetic parameters (Egger et al. 2017), increased the resolution of detectable population genetic structure (Ruegg et al. 2014, Benestan et al. 2015, Jahner et al. 2016), and provided opportunities for fine-scale investigations of genetically distinct populations and their relationship to subspecies boundaries (Larson et al. 2014, Fredrickson et al. 2015, Bussche et al. 2017, Mura-Jornet et al. 2018). In migratory birds, Ruegg et al. (2014) coined the term "genoscapes" to refer to maps of genetic variation across geographic space, but the relationship among genetically distinct populations within a

genoscape and previously defined subspecies boundaries has yet to be explored. Here we use next-generation sequencing (NGS) technology to create a genoscape for the American Kestrel, assess its relationship to current subspecies boundaries, and provide a framework for conservation and management of this and other highly mobile species with a high capacity for dispersal.

The American Kestrel (*Falco sparverius*) is a widely distributed species that breeds throughout North and South America (Smallwood and Bird 2020) and has upwards of 17 recognized subspecies (Ferguson-Lees and Christie 2006). American Kestrels show highly variable migration strategies across their range, including individuals that migrate long distances, short distances, or do not migrate, and populations that are completely migratory, partially migratory, or non-migratory (Layne 1982, Bird and Palmer 1988, Henny and Brady 1994, Smallwood and Bird 2020). Here we focus on 2 American Kestrel subspecies found north of the Mexico and U.S. border, the non-migratory subspecies (*F. s. paulus*) found breeding in the southeastern United States and the widespread subspecies (*F. s. sparverius*) found throughout the remainder of the United States and Canada (Hoffman and Collopy 1988, Smallwood 1990). In general, it is believed that populations of *F. s. sparverius* follow a pattern of leap-frog migration, where migratory distance decreases on a latitudinal gradient, with birds in the northernmost part of the breeding range migrating the farthest and birds in the southernmost part of the range remaining year-round residents (Heath et al. 2012). In addition, there is growing evidence that American Kestrel populations are declining (Bird 2009, Farmer and Smith 2009, Smallwood et al. 2009, Hinnebusch et al. 2010), but estimates of demographic trends differ regionally (McClure et al. 2017). One hypothesis to explain regional variation in demographic trends is that genetically distinct populations with different migratory strategies are exposed to different stressors across the annual cycle. As a result, identifying genetically distinct populations in American Kestrels and how they correspond with previously defined subspecies will improve our ability to interpret recent demographic trends and appropriately focus conservation actions.

Previous American Kestrel genetic research using 5 microsatellite loci and 1 mitochondrial DNA marker identified no strong signal of population genetic structure across the U.S. range, with only subtle differences in allele frequencies between the 2 recognized subspecies (Miller et al. 2012). Here we employ NGS-sequencing technology to screen 3 times the number of samples and 10,000 times the number of genetic loci relative to previous work, and use the resulting data to re-evaluate patterns of population structure in American Kestrels across their U.S. and Canadian breeding distribution. Specifically, we ask the following research questions: (1) Does genome-wide single nucleotide polymorphism (SNP) data provide higher resolution of population structure than previous work based on fewer markers? (2) How does the resulting population genetic structure relate to previously defined subspecies boundaries and variation in migratory behavior across the range? and (3) How can the resulting genoscape be used to help identify genetically distinct populations and develop hypotheses to explain regional variation in demographic trends?

## METHODS

### Sample Collection and DNA Extraction
Genetic samples were collected from 683 breeding adult or nestling American Kestrels from across the breeding range in North America in collaboration with several non-profit organizations, state agencies, university researchers, and citizen scientists (Table 1, Supplementary Material Appendix 1). Blood (~30 μL) or pin feather samples were collected from 287 individuals for the construction of restriction site-associated DNA sequencing (RADseq) libraries. Blood was collected via brachial or jugular venipuncture, preserved in Queen's lysis buffer (Seutin et al. 1991), and stored at –80°C or 3–4 pin feathers containing a small amount of blood in the base of the feather were collected from the breast of one nestling per brood and stored in envelopes at room temperature. The remainder of the feather samples used for high-throughput SNP analyses (*n* = 396) were collected from the breast of adult birds and stored in envelopes at room temperature. All samples were extracted using a QIAGEN DNeasy Blood and Tissue Kits (San Francisco, CA), and blood and pin feather extractions were further quantified using Qubit dsDNA HS157 Assay kits (Thermo Fisher Scientific) and visually inspected via gel electrophoresis to ensure selection of high-quality, intact DNA for construction of RADseq libraries. Remaining tissue and blood samples, as well as remaining extractions, were curated and made available for future use in –20°C and –80°C freezers, respectively, in the Conservation Genomics Laboratory at Colorado State University.

### Genome Sequencing, Assembly, and Annotation
To create a reference genome, the Illumina TruSeq DNA PCR-Free LT kit (Illumina) was used to prepare a genomic DNA library from a single individual from Boise, Idaho, following the adjustments made by Ruegg et al. (2018). The resulting library was sequenced on 2 lanes of an Illumina HiSeq 2500 using 250 base-pair (bp) paired-end sequencing at the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center (Davis, CA). Initial contigs were assembled with the Discovar DeNovo assembler from the Broad Institute (http://www.broadinstitute.org), discarding contigs <1,000 bp in length. We also sequenced mate-pair libraries with 2 insert sizes (4 kb and 8 kb) on

one-third of an Illumina HiSeq 2500 2 × 100 bp lane at the University of Utah Huntsman Cancer Center. Mate-pair reads were trimmed with NxTrim (O'Connell et al. 2015) and scaffolds were generated with both paired-end and mate-pair libraries with SSPACE (overlap requirement $k = 3$; Boetzer et al. 2011). The assembly was then broken at likely error regions using REAPR (Hunt et al. 2013) and SSPACE scaffolding was repeated with $k = 5$ and scaffolds <1,000 bp were discarded for the final assembly.

For annotation purposes, *RepeatMasker* (-species birds) (Tarailo-Graovac and Chen 2009) was used to replace repetitive regions of the final genome assembly with *N*'s. Two different ab initio gene predictions were used within the *Maker* pipeline (Cantarel et al. 2008): *Snap* and *Augustus.* The Zebra Finch (*Taeniopygia guttata*) cDNA and protein sequences were downloaded from Ensembl and used to train *Snap*, and the available chicken (*Gallus gallus domesticus*) training dataset was used to train *Augustus.* *Iterproscan* (Zdobnov and Apweiler 2001) was used to add Pfam protein annotation and gene ontology (GO) terms and identified 13,342 genes.

### SNP Discovery and SNP Filtering

High-density RADseq was carried out on 287 individuals from 12 sampling locations following a modified version of the BestRAD library preparation protocol (Table 1; Ali et al. 2016). In short, DNA was normalized to a final concentration of 100 ng in a 10 μL volume, digested with restriction enzyme SBfl (New England Biolabs [NEB]). The fragmented DNA was then ligated with SBfI specific adapters prepared with biotinylated ends and samples were pooled and cleaned using 1X Agencourt AMPure XP beads (Beckman Coulter). Pooled and clean libraries were sheared to an average length of 400 bp with 10 cycles on the Bioruptor NGS sonicator (Diagenode) to ensure appropriate length for sequencing and an Illumina NEBNext Ultra DNA Library Prep Kit (NEB) was used to repair blunt ends and ligate on NEBNext Adaptors to the resulting DNA fragments. Agencourt AMPure XP beads (Beckman Coulter) were then used to select DNA fragments with an average length of 500 bp, libraries were enriched with PCR, and cleaned again with Agencourt AMPure XP beads. The resulting libraries were sequenced on 3 lanes of an Illumina HiSeq 2500 at the UC Davis Genome Center using 250 bp, paired-end sequencing, and 66 individuals with low coverage were re-sequenced on a fourth lane.

The program *Stacks* (Catchen et al. 2013) was used to demultiplex, filter, and trim adapters from the data with the *process_radtags* function and remove duplicate read pairs using the *clone_filter* function. *Bowtie2* was used to map reads to the genome (Langmead and Salzberg 2012), and the *HaplotypeCaller* in the Genome Analysis Toolkit was used to identify SNPs (McKenna et al. 2010, Van der Auwera et al. 2013). *VCFtools* (Danecek et al. 2011) was used to remove indels, non-biallelic SNPs, and low-quality and rare variants (genotype quality: 20; coverage depth: 10; minor allele frequency: 0.05). The final number of SNPs and individuals to be retained for further analyses was assessed by visualizing the tradeoff between discarding low-coverage SNPs and discarding individuals with missing genotypes using custom scripts within the R package *GenoscapeRtools* (Anderson 2019). Because preliminary analyses revealed outliers in the principal component analysis (PCA) and we were concerned about sample contamination among individuals during the library preparation stage, we filtered out individuals with >40% heterozygosity as heterozygosity is expected to be higher than expected in cases where multiple individuals are combined into a single well (Supplementary Material Figure 2).

### Identification of Outlier SNPs for Population Assignment

Population genomic analyses were conducted on all SNPs that passed our filters to assess genome-wide patterns of genetic divergence and identify SNPs for population assignment and assay design. Population genetic structure was assessed by calculating pairwise population-level $F_{ST}$ (with different sampling sites representing different populations) with bootstrapped confidence intervals using the R package *Assigner* (Gosselin et al. 2019), and PCA using *SNPRelate* (Zheng et al. 2012). To test for isolation by distance, we compared linearized $F_{ST}$ with pairwise geographic distance calculated from the central longitudinal and latitudinal coordinates of each location using the Vincenty ellipsoid method in the R package *geosphere* (Hijmans 2019). Because the PCA of all SNPs from the genome-wide analysis revealed 5 major groups, including Alaska, Texas, the western, the eastern, and Florida (see Results, Figure 1), subsequent analyses focused on developing SNPs for population assignment within and among these groups.

To identify SNPs useful for population assignment between the 5 genetically distinct populations, we used *VCFtools* (Danecek et al. 2011) to calculate site-wise $F_{ST}$ between populations and identify individual SNPs with the most power for discriminating between populations (SNPs with the biggest allele frequency differences). It is important to note that population genetic summary statistics were based on the full RADseq dataset (see above) rather than downstream SNP dataset in order to avoid potential biases associated with selecting SNPs with the highest discriminatory power for population assignment. Custom R-scripts were used to evaluate which of our top-ranking SNPs would generate designable assays based on the following parameters: (1) Guanine–Cytocene content was <0.65; (2) there were no insertions or deletions (indels) within 30 bp of the variable site; or (3) there were no ambiguous codes within 20 bp of the variable site. Additionally,

**TABLE 1.** Sampling locations and sample types used in the construction of the genoscape map for American Kestrels. The samples are broken down into those used for the RAD-sequencing analysis (pre- and post-filters) and the SNP genotyping (pre- and post-filtering). State and province location codes correspond to Figure 1, while letter location codes correspond to the locations in Figure 2.

| Study site | Latitude | Longitude | Sample type | N_RAD_NoFilter | N_RAD_Filter | N_SNP_NoFilter | N_SNP_NoFilter | Location code (Figure 1/Figure 2) |
|---|---|---|---|---|---|---|---|---|
| Alberta, Canada | 56.16262 | −117.34 | Feather | | | 12 | 12 | E |
| Alaska, USA | 63.96041 | −145.136 | Blood, feather | 25 | 18 | 7 | 7 | AK/A |
| Arizona, USA | 33.45007 | −111.943 | Feather | | | 20 | 15 | G |
| California (1), USA | 37.23551 | −120.244 | Blood, feather | 26 | 10 | 22 | 22 | CA/D1 |
| California (2), USA | 33.7315 | −117.878 | Blood, feather | 4 | 1 | 3 | 3 | CA/D2 |
| Connecticut (1), USA | 41.95001 | −73.364 | Feather | | | 10 | 10 | Y1 |
| Connecticut (2), USA | 41.94374 | −72.5072 | Feather | | | 10 | 7 | Y2 |
| Delaware, USA | 39.75047 | −75.6969 | Feather | | | 4 | 4 | W |
| Florida, USA | 29.54151 | −82.2334 | Blood, feather | 18 | 16 | 20 | 20 | FL/AA |
| Idaho (1), USA | 43.51209 | −116.377 | Blood, feather | 33 | 29 | 18 | 11 | ID/F1 |
| Idaho (2), USA | 43.27597 | −115.003 | Blood, feather | 9 | 2 | 6 | 5 | ID/F2 |
| Illinois, USA | 38.30563 | −89.7969 | Feather | | | 5 | 3 | P |
| Kansas, USA | 38.77533 | −95.2486 | Feather | | | 4 | 4 | M |
| Massachusetts, USA | 42.15064 | −71.8496 | Feather | | | 20 | 18 | Z |
| Michigan (1), USA | 42.3216062 | −86.182447 | Feather | | | 6 | 6 | R1 |
| Michigan (2), USA | 44.9774725 | −85.712026 | Feather | | | 6 | 6 | R2 |
| Michigan (3), USA | 44.9034875 | −83.4873 | Feather | | | 8 | 8 | R3 |
| Missouri (1), USA | 39.04824 | −94.3663 | Feather | | | 8 | 5 | N1 |
| Missouri (2), USA | 38.3804 | −90.925 | Feather | | | 1 | 1 | N2 |
| Montana, USA | 45.76764 | −111.154 | Feather | | | 19 | 15 | I |
| North Carolina, USA | 35.73285 | −82.6776 | Feather | | | 17 | 16 | S |
| Nebraska, USA | 40.99115 | −96.8957 | Feather | | | 6 | 6 | L |
| New Jersey, USA | 40.4297 | −74.9271 | Feather | | | 30 | 28 | X |
| New York, USA | 41.8773 | −74.7309 | Feather | | | 4 | 3 | U |
| Oregon (1), USA | 45.57362 | −123.048 | Blood, feather | 14 | 11 | 3 | 3 | OR/C1 |
| Oregon (2), USA | 44.36466 | −121.412 | Feather | | | 25 | 24 | C2 |
| Pennsylvania, USA | 40.60183 | −75.5961 | Blood, feather | 37 | 23 | 13 | 13 | PA/V |
| Saskatchewan, Canada | 55.01968 | −105.893 | Blood, feather | 40 | 28 | 12 | 11 | SK/K |
| Tennessee, USA | 35.14944 | −90.0489 | Feather | | | 5 | 5 | O |
| Texas, USA | 33.63184 | −101.889 | Blood, feather | 13 | 12 | 1 | 1 | TX/B |
| Utah, USA | 40.62255 | −111.987 | Feather | | | 30 | 28 | H |
| Virginia, USA | 38.6457 | −78.7522 | Blood, feather | 38 | 29 | 17 | 17 | VA/T |
| Wisconsin, USA | 43.77114 | −89.4059 | Blood, feather | 30 | 18 | 24 | 24 | WI/Q |
| Wyoming, USA | 43.60488 | −110.658 | Feather | | | 17 | 15 | J |
| TOTAL | | | | 287 | 197 | 396 | 376 | |
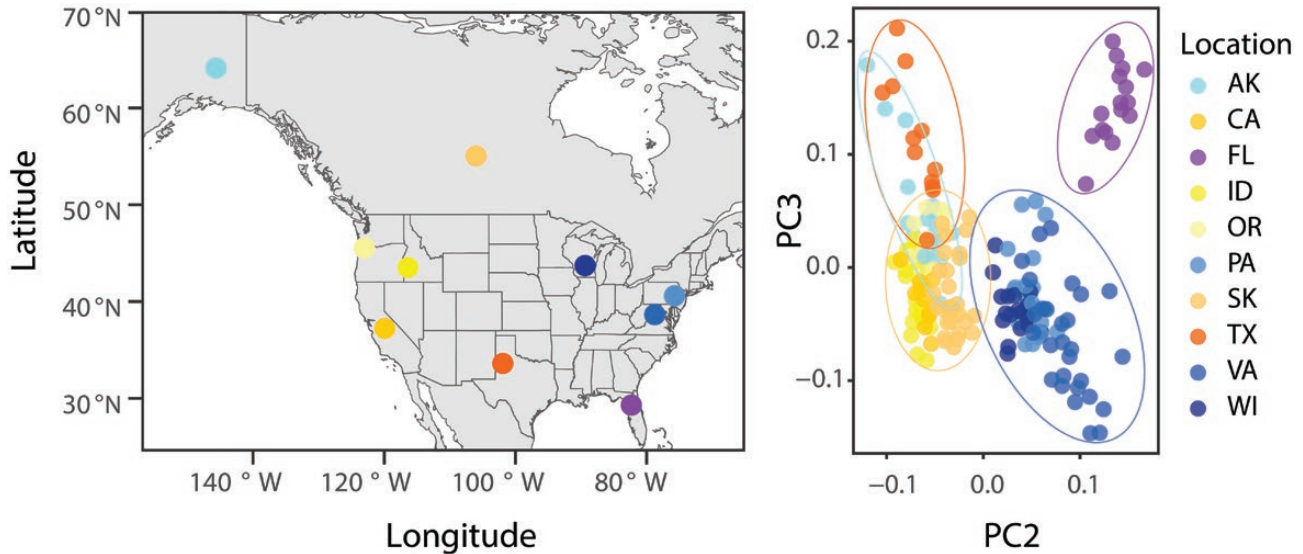
**FIGURE 1.** Principal component analysis of 72,263 SNP markers from across the breeding range of American Kestrels showing separation between eastern, western, Alaska, Texas, and the Florida subspecies. Each putative conservation unit is encircled by an ellipse. Separation between Alaska and Texas conservation units occurred along PC axes not shown and was also evident in pairwise $F_{ST}$ calculations.

we used BWA-MEM (Burrows-Wheeler Aligner; Li and Durbin 2009) to determine which of our designable SNPs mapped uniquely to the reference genome. Fluidigm SNP-type assays (Fluidigm Inc) were then developed in the 216 top-ranking SNPs that passed our filters.

**Genetic Screening and Building the Genoscape**
Ninety-three samples and 3 non-template controls were screened on the Fluidigm Corporation EP1 Genotyping System (Fluidigm Inc) and assays were ranked by variability and call rate to identify the most reliable 192 SNP assays of 216 that were designed. The 192 variable SNP assays with the highest call rate were used to screen an additional 396 American Kestrel feather samples from 34 breeding locations in the United States and Canada in order to fill in sampling gaps and refine the resulting map of population genetic structure (Table 1). Following the methods described in Ruegg et al. (2014), we amplified PCR products using fluorescently labeled allele-specific primers and then used the EP1 Array Reader and Fluidigm's automated Genotyping Analysis Software (Fluidigm Inc) to call alleles with a confidence threshold of 90%. Each genotype was also visually inspected for potential irregularities and uncertain genotype calls were removed from the analysis. Samples with missing genotypes at >25% of SNP assays were removed from the analyses and SNP loci with >25% missing genotypes were removed, resulting in a total of 376 additional individuals at 186 SNP loci that could be used to identify genetic structure across the range (Table 1).

The final analysis of population genetic structure at 186 loci was conducted on a subset of the loci from the RADseq dataset combined with the SNP genotype-only dataset for a total of 683 individuals. The program STRUCTURE (version 2.3.4; Pritchard et al. 2000) was used to assess how genetic variation is distributed across geographic space. The admixture model with the locprior option was run with uncorrelated allele frequencies, a burn-in period of 50,000, a total run length of 150,000, and assuming the number of genetic clusters (K) ranged from 1 to 10 (with 5 iterations run at each assumed value of K). We used the Evanno method to determine the number of genetic clusters. The Evanno method (Evanno et al. 2005), implemented in *pophelper* in R (Francis 2017), is an ad hoc method to determine the most probable number of population genetic clusters based on the rate of change in the log probability of data between successive K values. We used this algorithm to detect the uppermost hierarchical level of structure across the Kestrel breeding range and visually inspected subsequent structure plots to identify regions where geographic barriers to gene flow exist and/or where admixture homogenizes population structure. The resulting posterior probabilities of genetic group membership estimated from structure were visualized as transparency levels of different colors overlaid upon a base map from Natural Earth (naturalearthdata.com) and clipped to a map of the American Kestrel breeding range (NatureServe 2012), making use of the R packages *sp*, *rgdal*, and *raster* (Bivand et al. 2013, 2017, Hijmans 2017). We scaled the transparency of colors within each distinguishable group, so that the highest posterior probability of membership in the group according to the structure is opaque and the smallest is transparent. This creates a
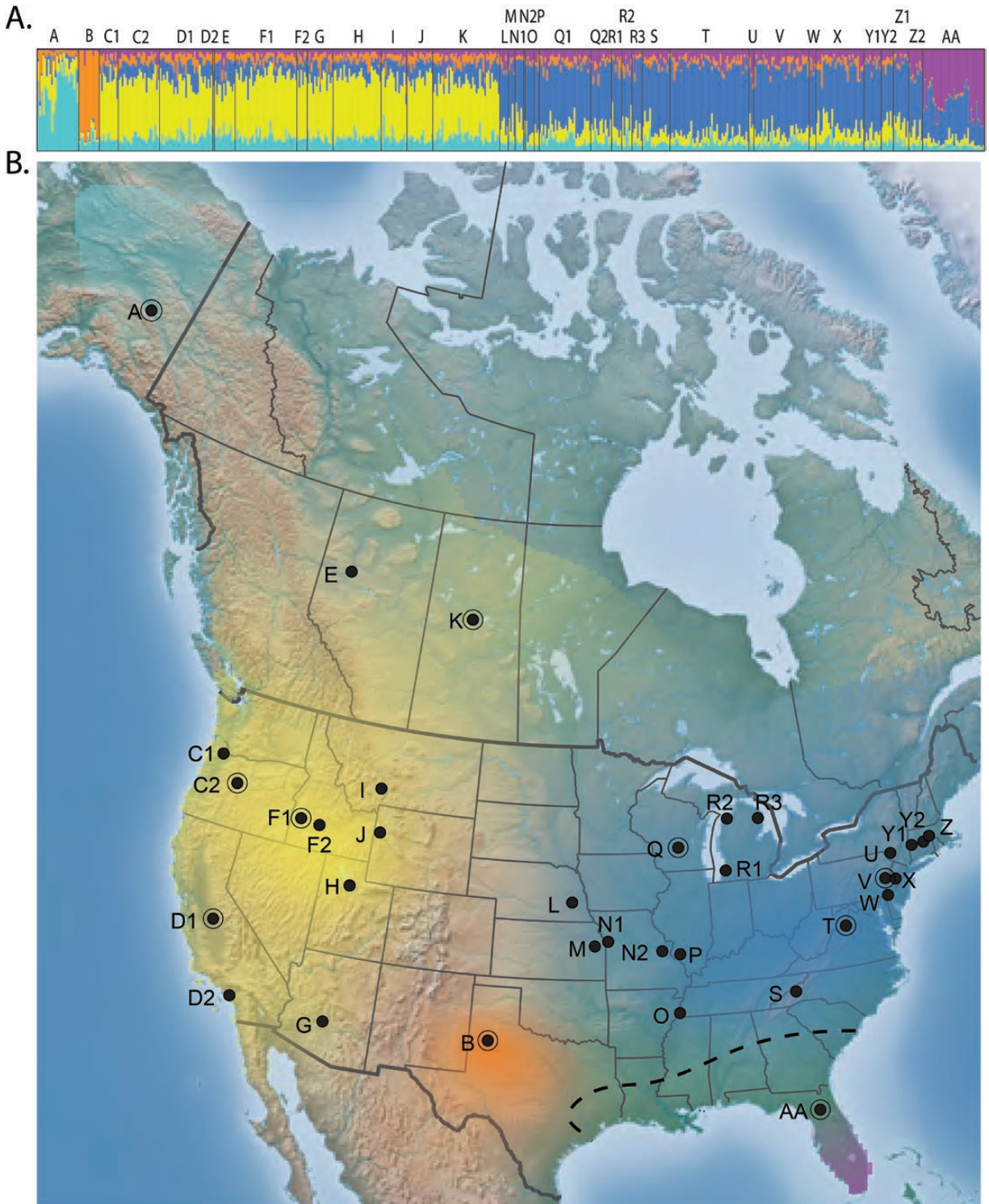
**FIGURE 2.** The American Kestrel genoscape. (A) Structure plot showing support for $K = 5$ genetic groups. Letters correspond to population locations on the map as well as sample numbers listed in Table 1. (B) A spatially explicit representation of the population structure results showing the biggest genetic differences between eastern, western, Texas, Florida, and Alaska genetic groups. Dots with circles around them indicate sampling locations were both RADseq and SNP genotyping was conducted. The dashed line indicates the hypothesized northern boundary of *F. s. sparvarius* (Lane and Fischer 1997).

spatially explicit map of the population structure analysis that we call the genoscape of the American Kestrel (Figure 2).

## RESULTS

### Genome Sequencing, Assembly, and Annotation
The final American Kestrel genome assembly is 1.23 Gb in length and consists of 5,096 scaffolds with an N50 of 941 kb.

### SNP Discovery and SNP Filtering
RAD-sequencing data from 287 individuals resulted in the identification of 199,705 biallelic loci with a minor allele frequency greater than 5%, minimum quality score greater than 20, and minimum per individual sequencing depth greater than 10. After assessing the tradeoff between low-coverage SNPs and missing genotypes (Supplementary Material Figure 1), the data were further filtered to include 197 individuals, 12 populations, and 75,000 loci; 2 populations, 1 in California and 1 in Idaho (Table 1), were subsequently dropped for the purposes of population genetic analyses as a result of low sample size ($n < 3$). Seven outlier individuals with greater than >40% heterozygosity were also subsequently removed to avoid inclusion of samples potentially subject to contamination (as indicated by a histogram of the distribution of heterozygosity across all individuals which showed individuals above this threshold to be clear outliers, Supplementary Material Figure 2). The final RADseq dataset consisted of 197 individuals and 72,263 SNPs.

### Population Genetic Structure
Significant pairwise $F_{ST}$ between the 10 sampling locations (2 were filtered out because they had fewer than 4 individuals, see above) ranged from 0.0010 to 0.0162 (Table 2). The positive correlation between $F_{ST}$ and geographic distance ($r^2 = 0.123$; $P$-value = 0.01) suggests that isolation by distance contributes to genetic differentiation across the range (Figure 3). Overall, $F_{ST}$ was highest between non-migratory Florida and Texas sampling locations, while the genetic differentiation was the lowest among sampling locations within eastern and western breeding areas (Table 2). PCA based on 72,263 RAD-sequence loci revealed 4 main clusters with eastern, western, and Florida falling out separately, while Alaska and Texas overlapped (Figure 1). Principal component (PC) 1 was strongly influenced by data missingness, while PC2 and PC3 reflected differences in geography. Although there was overlap in PC space between Texas and Alaska along PC2 and PC3, these groups were separated in the subsequent PC axes and in the search for loci representative of the observed population structure based on significant pairwise $F_{STs}$ and geographic distance. Overall, the first 3 PCs explained <3% of the total variation in allele frequencies.

### Genoscape Construction/Structure Results
We successfully genotyped 376 samples collected from 34 breeding locations in the United States and Canada using the final panel of 192 SNP-type assays for population assignment (Table 1). Running STRUCTURE with $K$ ranging from 1 to 10 revealed the strongest support for $K = 5$, where the plateau of delta $K$ (i.e. the greatest change in $K$) supports the uppermost hierarchical level of structure being $K = 3$ (Supplementary Material Figure 3A; Pritchard et al. 2000, Evanno et al. 2005); however, subsequent increases of $K$, until $K = 5$, reduced the log-likelihood of the model and the addition of the Florida ($K = 4$) and Alaskan populations ($K = 5$) were biologically feasible (Supplementary Material Figure 4), suggesting that American Kestrels can be separated into 5 genetically distinct populations. Ultimately, we find genetic distinctness of the residential Texas population, the residential Florida *F. s. paulus* subspecies, as well as distinct clustering of the eastern and western migratory populations, and an Alaskan migratory population (Figure 2).

## DISCUSSION

Identifying distinct units for conservation is an important first step in the management of declining populations (Allendorf and Luikart 2007, Funk et al. 2012). Historically, conservation biologists have used a combination of morphological, behavioral, and genetic variation to define management or conservation units within species and these have sometimes, but not always, corresponded to subspecies boundaries. Here we generate a genoscape for the American Kestrel by sequencing the first American Kestrel genome, assessing population structure at 72,263 SNP markers screened in 12 populations from across the U.S. and Canadian migratory and non-migratory range, and validating patterns of population structure at 192 SNP markers screened in 34 populations. In contrast to previous work based on a more limited number of samples and markers that detected no major signals of population structure across the breeding range (Miller et al. 2012), our genoscape supports the existence of 5 genetically distinct populations within American Kestrels found breeding across Canada and the United States (eastern, western, Alaska, Texas, and Florida), one of which correlates with the previously identified southeastern subspecies (*F. s. paulus*). Overall, the most significant genetic differences occurred between the 2 resident populations (Texas and Florida), followed by differences between resident and migratory populations, and regional separation of eastern and western breeding populations. Here we discuss the utility of the resulting genoscape for clarifying the relationship between previously defined subspecies boundaries and genetically distinct populations identified

**TABLE 2.** Weighted mean pairwise (Weir and Cockerham 1984) $F_{ST}$ values (in bold) and confidence intervals (in roman) between all American kestrel sampling locations with greater than 5 individuals. Sampling locations correspond to those listed in Figure 1 and Table 1 and are organized by approximate migration strategy.

|  |  | Complete migrant | | Partial migrant | | | | | | Resident | |
|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | AK1 | SK1 | OR1 | ID1 | CA1 | WI1 | PA1 | VA1 | TX1 | FL1 |
| Complete migrant | AK1 | – | **0.0031** | **0.0041** | **0.0048** | **0.0053** | **0.0060** | **0.0062** | **0.0086** | **0.0096** | **0.0124** |
|  | SK1 | 0.0028–0.0034 | – | **0.0019** | **0.0016** | **0.0014** | **0.0015** | **0.0024** | **0.0046** | **0.0078** | **0.0103** |
| Partial migrant | OR1 | 0.0036–0.0046 | 0.0015–0.0023 | – | **0.0025** | **0.0031** | **0.0046** | **0.0051** | **0.0073** | **0.0077** | **0.0112** |
|  | ID1 | 0.0044–0.0050 | 0.0014–0.0019 | 0.0021–0.0030 | – | **0.0019** | **0.0040** | **0.0051** | **0.0072** | **0.0088** | **0.0125** |
|  | CA1 | 0.0048–0.0058 | 0.0010–0.0019 | 0.0025–0.0037 | 0.0015–0.0023 | – | **0.0039** | **0.0051** | **0.0070** | **0.0099** | **0.0128** |
|  | WI1 | 0.0056–0.0065 | 0.0012–0.0018 | 0.0040–0.0052 | 0.0037–0.0043 | 0.0035–0.0044 | – | **0.0010** | **0.0029** | **0.0100** | **0.0082** |
|  | PA1 | 0.0058–0.0065 | 0.0021–0.0027 | 0.0046–0.0056 | 0.0048–0.0054 | 0.0047–0.0057 | 0.0006–0.0012 | – | **0.0027** | **0.0103** | **0.0076** |
|  | VA1 | 0.0082–0.0090 | 0.0043–0.0048 | 0.0068–0.0079 | 0.0069–0.0075 | 0.0065–0.0074 | 0.0025–0.0032 | 0.0025–0.0030 | – | **0.0127** | **0.0093** |
| Resident | TX1 | 0.0091–0.0102 | 0.0074–0.0083 | 0.0070–0.0082 | 0.0083–0.0092 | 0.0092–0.0106 | 0.0094–0.0105 | 0.0099–0.0108 | 0.0121–0.0132 | – | **0.0162** |
|  | FL1 | 0.0120–0.0129 | 0.0099–0.0106 | 0.0105–0.0116 | 0.0121–0.0129 | 0.0122–0.0134 | 0.0076–0.0088 | 0.0072–0.0081 | 0.0089–0.0098 | 0.0156–0.0168 | – |
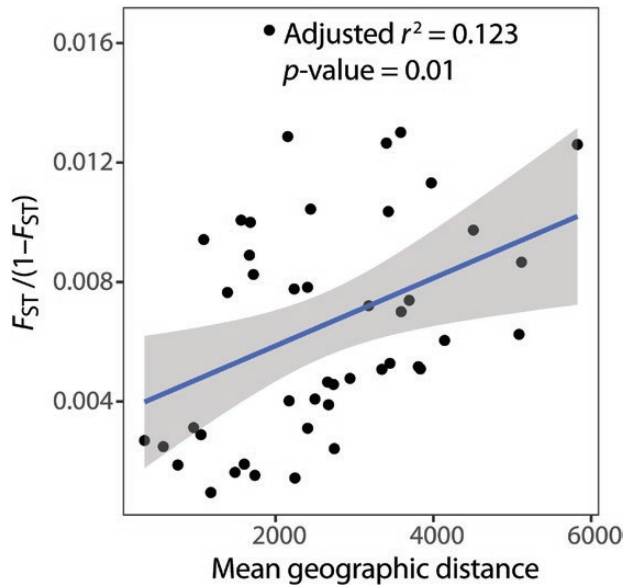
**FIGURE 3.** The relationship between pairwise linearized $F_{ST}$ and geographic distance. The significant $r^2$ supports the idea that genetic divergence across the range is partly a result of isolation by distance.

using genome-wide genetic data, as well as for providing a framework for developing hypotheses regarding drivers of regional variation in demographic trends.

The question of whether subspecies represent defensible taxonomic units has been controversial in the past because some molecular studies have failed to identify subspecies as phylogenetically distinct (Barrowclough 1980, Mayr and Ashlock 1991, O'Brien and Mayr 1991, Ball and Avise 1992, Burbrink et al. 2000). Further, discord in location of subspecies boundaries often arises when there are mismatches in the timescales over which divergence occurs in various datasets; for example, subspecies boundaries based on neutral genetic markers often diverge from subspecies boundaries identified based on genetic or morphological markers that may be under selection (Haig and Winker 2010). Using genome-wide sequencing, we found support for genetic differentiation between the 2 U.S. and Canadian subspecies of American Kestrels breeding North of Mexico, *F. s. paulus* (southeastern) and *F. s. sparverius* (remainder of the U.S. and Canadian breeding range), but also found that divergence between the subspecies is similar in magnitude to the degree of divergence detected between resident *F. s. sparverius* in Texas and their migratory counterparts to the North. In general, levels of genetic differentiation across the range were low, and pairwise genetic distance vs. geographic distance suggest that patterns of divergence are in large part explained by isolation by distance (Figure 3). Such low levels of differentiation are consistent with past studies of American Kestrels based on fewer loci (Miller et al.

2012) and suggest that gene flow may homogenize the diversifying effects of local adaptation and drift in high dispersal species like American Kestrels (Willoughby et al. 2017, Doyle et al. 2018, Medina et al. 2018). While our genome-wide genetic analysis supports the existence of *F. s. paulus* as a genetically distinct subspecies, it also suggests that weak population structure within American Kestrels relates as much to migratory phenotype as it does to subspecies boundaries per se.

One possible explanation for the putative relationship between genetic differentiation and migratory phenotype in American Kestrels is that dispersal is limited between distinct migratory phenotypes, as has been found in other migratory systems (reviewed within Turbek et al. 2018). Like most raptors, Kestrels in eastern and western North America follow a strong north-to-south pattern of migration, with little longitudinal drift (Mueller and Berger 1967, Evans and Rosenfield 1985, Goodrich and Smith 2008). Additionally, the frequency of long-distance migration into Mexico is thought to increase from east to the west (Mueller and Berger 1967, Evans and Rosenfield 1985, Goodrich and Smith 2008), supporting the idea that eastern and western populations have different overwintering locations. One explanation for the observed genetic break in central North America is that separate eastern and western migratory routes and overwintering locations have resulted in the evolution of a weak migratory divide where gene flow is limited as a result of reproductive isolation between distinct migratory phenotypes, as has been documented across migratory divides in other avian taxa (reviewed within Turbek et al. 2018). Alternatively, the observed genetic break between eastern and western populations may have nothing to do with migratory phenotype per se, but instead may result from low population density in central North America (an additive effect of isolation by distance) limiting gene flow between eastern and western groups (Winker 2010). Future work will focus on quantifying migratory phenotypes in American Kestrels across their North American range and assessing the relative contribution of migration and isolation by distance to patterns of genetic divergence in this and other species with similar variation in migratory phenotypes.

An alternative explanation for higher levels of divergence between resident and migratory *F. sparverius* populations in our study is that divergence is not caused by differences in migratory phenotype, but instead results from gene flow between resident forms in Texas and Florida (*F. s. Paulus*) and resident subspecies further to the south (*F. s. peninsularis* from Baja and W. Mexico, and *F. s. sparverioides* from Cuba). Previous work suggests *F. s. sparverius* may hybridize with *F. s. peninsularis* in northern and eastern Arizona (Bond 1943), but here we successfully genotyped 15 samples from breeding migratory *F. s. sparverius* from Arizona and found that they

all assigned clearly to the western migratory group of *F. s. sparverius* rather than to the resident Texas group. This result is opposite of what we would expect if high levels of divergence between Texas residents and their migratory counterparts to the North were due to hybridization between Texas birds and the *F. s. peninsularis* subspecies to the south. While to the best of our knowledge, there are no known records of hybridization between *F. s. sparverioides* from Cuba and *F. s. paulus* from Florida, eBird records of kestrels from extreme southeastern Florida during the summer support the possibility that *F. s. sparverioides* vagrants may occur within the same region. However, gene flow is unlikely, given the infrequent (<1 per year) sightings of kestrels in southeastern Florida and the fact that the nearest breeding population of *F. s. paulus* is >150 km to the north (Florida Fish and Wildlife Conservation Commission 2003). Thus, while on-going gene flow with resident subspecies to the south seems like an unlikely explanation for the observed patterns of divergence between resident and migratory populations north of Mexico, more extensive sampling south of the U.S. border is needed to fully test all alternative hypotheses.

Heterogeneity in patterns of American Kestrel population decline across North America suggest that regional populations are experiencing different threats and/or are responding to the same threats differently (Butcher 1990, Smallwood et al. 2009, Sauer et al. 2014, McClure et al. 2017), but past analyses have been limited by the lack of genetically distinct populations. The results presented herein demonstrate the utility of the genoscape approach for identifying 5 genetically distinct populations of American Kestrels—eastern, western, Alaska, Texas, and Florida—which can serve as the foundation for the development of hypotheses to explain regional variation in demographic trends. For example, while interpreting patterns of population decline from existing datasets is complicated by known northward shifts in distribution (Paprocki et al. 2014), migration count data from the Raptor Population Index project between 2006 and 2016 support the idea that western populations have largely remained stable or are increasing, while eastern populations are largely declining (Supplementary Material Figure 3; Brandes et al. 2016, Crewe et al. 2016). In addition, work focused specifically on understanding responses of American Kestrels to climate change in the last decade supports the hypothesis that western populations are migrating shorter distances and breeding earlier (Heath et al. 2012), while corresponding changes in the east have not been documented. In light of the genoscape results presented herein, one hypothesis that warrants further exploration is that genetically based differences in phenology between eastern and western groups affect population-specific responses to changing climate

conditions, resulting in population decreases in the east, but not the west.

In conclusion, the American Kestrel genoscape reveals previously undetected levels of population structure among eastern, western, Alaska, Texas, and Florida populations. While our data support the existence *F. s. sparverius* and *F. s. paulus* subspecies as genetically distinct groups, it also suggests that genetic differentiation is more closely tied to migratory phenotype (resident, long-distance, and short-distance migrants) than to previously defined subspecies boundaries. Based on our results, we suggest it would be ecologically appropriate to establish 5 management areas corresponding to the 5 genetically unique populations identified by our genoscape. More importantly, when the resulting genetically distinct populations are paired with data from existing long-term monitoring efforts, such as the Raptor Population Index, the results can be used to test hypotheses regarding drivers of observed population-specific responses to climate change and other stressors.

## SUPPLEMENTARY MATERIAL

Supplementary material is available at *Ornithology* online.

## ACKNOWLEDGMENTS

**Ethics statement:** Each author had their own IACUC protocol and corresponding permits for the collection of blood samples.

**Author contributions:** K.C.R., J.A.H., M.D.O., S.S., C.J.W.M., T.B.S., and M.B. conceived of the idea design and experiment. K.C.R., M.B., C.M.B., R.A.B., and E.C.A. collected the genomic data and conducted the bioinformatic and population genetic analyses. K.C.R., M.B., J.A.H., and C.M.B. wrote the manuscript with contributions from all co-authors. M.B., C.W.B., R.D.D., A.E., C.J.W.M., K.E.M., L.M., J.M., M.D.O., B.R., S.S., T.S., J.F.T., and J.A.H. contributed blood samples for the genome-wide genetic analysis.

**Data deposits:** Analyses reported in this article can be reproduced using the data provided by Ruegg et al. (2021).

## LITERATURE CITED

Ali, O. A., S. M. O'Rourke, S. J. Amish, M. H. Meek, G. Luikart, C. Jeffres, and M. R. Miller (2016). RAD capture (Rapture): Flexible and efficient sequence-based genotyping. Genetics 202:389–400.

Allendorf, F. W., and G. Luikart (2007). Management units. In Conservation and the Genetics of Populations. 16.5.3. Blackwell Publishing, Oxford, UK.

Anderson, E. (2019). genoscapeRtools: Tools for building migratory bird genoscapes. R package version 0.1.0. https://rdrr.io/github/eriqande/genoscapeRtools/

Ball, R. M., and J. C. Avise (1992). Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. The Auk 109:626–636.

Barrowclough, G. F. (1980). Genetic and phenotypic differentiation in a Wood Warbler (*Genus Dendroica*) hybrid zone. The Auk 97:655–668.

Benestan, L., T. Gosselin, C. Perrier, B. Sainte-Marie, R. Rochette, and L. Bernatchez (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). Molecular Ecology 24:3299–3315.

Bird, D. M. (2009). The American Kestrel: From common to scarce? Journal of Raptor Research 43:261–262.

Bird, D. M., and R. S. Palmer (1988). American Kestrel. In Handbook of North American Birds (R. S. Palmer, Editor). Yale University Press, New Haven, CT, USA. pp. 253–290.

Bivand, R., T. Keitt, and B. Rowlingson (2017). rgdal: Bindings for the "Geospatial" Data Abstraction Library. R package version 1.4–8. https://CRAN.R-project.org/package=rgdal

Bivand, R. S., E. Pebesma, and V. Gómez-Rubio (2013). Applied Spatial Data Analysis with R. Springer, New York, NY, USA.

Boetzer, M., C. V. Henkel, H. J. Jansen, D. Butler, and W. Pirovano (2011). Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579.

Bond, R. M. (1943). Variation in western sparrow hawks. The Condor 45:168–185.

Brandes, D., D. Oleyar, S. Hoffman, and L. Goodrich (2016). The Raptor Population Index, 2016 Regional Trend Summaries and Conservation Assessments. http://rpi-project.org/2016/assessments2013.php

Burbrink, F. T., R. Lawson, and J. B. Slowinski (2000). Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): A critique of the subspecies concept. Evolution: International Journal of Organic Evolution 54:2107–2118.

Bussche, R. A. V. D., M. E. Judkins, M. J. Montague, and W. C. Warren (2017). A resource of genome-wide single nucleotide polymorphisms (SNPs) for the conservation and management of golden eagles. Journal of Raptor Research 51:368–377.

Butcher, G. S. (1990). Audubon Christmas bird counts. In Survey Designs and Statistical Methods for the Estimation of Avian Population Trends. Biological Report 90(1) (J. R. Sauer and S. Droege, Editors). U.S. Fish and Wildlife Service, Washington, DC, USA. pp. 5–13.

Cantarel, B. L., I. Korf, S. M. Robb, G. Parra, E. Ross, B. Moore, C. Holt, A. Sánchez Alvarado, and M. Yandell (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research 18:188–196.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko (2013). Stacks: An analysis tool set for population genomics. Molecular Ecology 22:3124–3140.

Crewe, T., P. Taylor, D. Lepage, L. Goodrich, J. Brown, and J. Sodergren (2016). The Raptor Population Index, 2016 analysis methods and trend results. http://rpi-project.org/2016/

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Doyle, J. M., D. A. Bell, P. H. Bloom, G. Emmons, A. Fesnock, T. E. Katzner, L. LaPré, K. Leonard, P. SanMiguel, R. Westerman, et al. (2018). New insights into the phylogenetics and population structure of the prairie falcon (*Falco mexicanus*). BMC Genomics 19:233.

Egger, B., M. Roesti, A. Böhne, O. Roth, and W. Salzburger (2017). Demography and genome divergence of lake and stream populations of an East African cichlid fish. Molecular Ecology 26:5016–5030.

Evanno, G., S. Regnaut, and J. Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology 14:2611–2620.

Evans, D. L., and R. N. Rosenfield (1985). Migration and mortality of sharp-shinned Hawks ringed at Duluth, Minnesota, USA. ICBP Technical Publication 5:311–316.

Farmer, C. J., and J. P. Smith (2009). Migration monitoring indicates widespread declines of American Kestrels (*Falco sparverius*) in North America. Journal of Raptor Research 43:263–273.

Ferguson-Lees, J., and D. A. Christie (2006). Raptors of the World. Princeton University Press, Princeton, NJ, USA.

Florida Fish and Wildlife Conservation Commission (2003). Florida's breeding bird atlas: A collaborative study of Florida's birdlife. http://myfwc.com/bba

Francis, R. M. (2017). pophelper: An R package and web app to analyse and visualize population structure. Molecular Ecology Resources 17:27–32.

Fredrickson, R. J., P. W. Hedrick, R. K. Wayne, B. M. vonHoldt, and M. K. Phillips (2015). Mexican wolves are a valid subspecies and

an appropriate conservation target. The Journal of Heredity 106:415–416.

Freer, J. J., B. K. Mable, G. Clucas, A. D. Rogers, M. J. Polito, M. Dunn, R. Naveen, H. Levy, and T. Hart (2015). Limited genetic differentiation among Chinstrap penguin (*Pygoscelis antarctica*) colonies in the Scotia Arc and Western Antarctic Peninsula. Polar Biology 38:1493–1502.

Funk, W. C., J. K. McKay, P. A. Hohenlohe, and F. W. Allendorf (2012). Harnessing genomics for delineating conservation units. Trends in Ecology & Evolution 27:489–496.

Funk, W. C., T. D. Mullins, and S. M. Haig (2007). Conservation genetics of Snowy Plovers (*Charadrius alexandrinus*) in the Western Hemisphere: Population genetic structure and delineation of subspecies. Conservation Genetics 8:1287–1309.

Goodrich, L. J., and J. P. Smith (2008). Raptor migration in North America. In State of North America's Birds of Prey (K. L. Bildstein, J. P. Smith, E. Ruelas Inzunza, and R. R. Veit, Editors). The Nuttall Ornithological Club and the American Ornithologists' Union, Washington, DC, USA. pp. 37–149.

Gosselin, T., E. C. Anderson, and I. Bradbury (2019). assigner: Assignment analysis with GBS/RAD data using R. R package, version 0.5.6. https://rdrr.io/github/eriqande/assigner/

Haig, S. M., and K. Winker (2010). Avian subspecies: Summary and prospectus. In Avian Subspecies (K. Winker and S. M. Haig, Editors). Ornithological Monographs 67:172–175.

Heath, J. A., K. Steenhof, and M. A. Foster (2012). Shorter migration distances associated with higher winter temperatures suggest a mechanism for advancing nesting phenology of American Kestrels *Falco sparverius*. Journal of Avian Biology 43:376–384.

Henny, C. J., and G. L. Brady (1994). Partial migration and wintering localities of American Kestrels nesting in the Pacific northwest. Northwestern Naturalist 75:7.

Hijmans, R. J. (2017). raster: Geographic data analysis and modeling. R package version 3.1-5. https://CRAN.R-project.org/package=raster

Hijmans, R. (2019). geosphere: Spherical trigonometry. R package version 1.5-10. https://CRAN.R-project.org/package=geosphere

Hinnebusch, D. M., J.-F. Therrien, M.-A. Valiquette, B. Robertson, S. Robertson, and K. L. Bildstein (2010). Survival, site fidelity, and population trends of American kestrels wintering in southwestern Florida. The Wilson Journal of Ornithology 122:475–483.

Hoffman, M. L., and M. W. Collopy (1988). Historical status of the American Kestrel (*Falco sparverius paulus*) in Florida. The Wilson Bulletin 100:91–107.

Hunt, M., T. Kikuchi, M. Sanders, C. Newbold, M. Berriman, and T. D. Otto (2013). REAPR: A universal tool for genome assembly evaluation. Genome Biology 14:R47.

Jahner, J. P., D. Gibson, C. L. Weitzman, E. J. Blomberg, J. S. Sedinger, and T. L. Parchman (2016). Fine-scale genetic structure among greater sage-grouse leks in central Nevada. BMC Evolutionary Biology 16:127.

Lane, J. J., and R. A. Fischer (1997). Species profile: Southeastern American Kestrel (*Falco sparverius paulus*) on military installations in the southeastern United States. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, USA.

Langmead, B., and S. L. Salzberg (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357–359.

Larson, W. A., L. W. Seeb, M. V. Everett, R. K. Waples, W. D. Templin, and J. E. Seeb (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). Evolutionary Applications 7:355–369.

Layne, J. M. (1982). Analysis of Florida-related banding data for the American Kestrel. North American Bird Bander 7:94–99.

Li, H., and R. Durbin (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

Mayr, E. (1982). Of what use are subspecies? The Auk 99:593–595.

Mayr, E., and P. D. Ashlock (1991). Principles of Systematic Zoology. McGraw-Hill, New York, NY, USA.

McClure, C. J. W., S. E. Schulwitz, R. V. Buskirk, B. P. Pauli, and J. A. Heath (2017). Commentary: Research recommendations for understanding the decline of American Kestrels (*Falco sparverius*) across much of North America. Journal of Raptor Research 51:455–464.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20:1297–1303.

Medina, I., G. M. Cooke, and T. J. Ord (2018). Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. Ecology Letters 21:638–645.

Miller, M. P., T. D. Mullins, J. W. Parrish, Jr, J. R. Walters, and S. M. Haig (2012). Variation in migratory behavior influences regional genetic diversity and structure among American Kestrel populations (*Falco sparverius*) in North America. The Journal of Heredity 103:503–514.

Morin, P. A., K. K. Martien, and B. L. Taylor (2009). Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources 9:66–73.

Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. Trends in Ecology & Evolution 9:373–375.

Mueller, H. C., and D. D. Berger (1967). Wind drift, leading lines, and diurnal migration. The Wilson Bulletin 79:50–63.

Mura-Jornet, I., C. Pimentel, G. P. M. Dantas, M. V. Petry, D. González-Acuña, A. Barbosa, A. D. Lowther, K. M. Kovacs, E. Poulin, and J. A. Vianna (2018). Chinstrap penguin population genetic structure: One or more populations along the Southern Ocean? BMC Evolutionary Biology 18:90.

NatureServe (2018). NatureServe Explorer. NatureServe, Arlington, VA, USA. https://explorer.natureserve.org/

O'Brien, S. J., and E. Mayr (1991). Bureaucratic mischief: Recognizing endangered species and subspecies. Science 251:1187–1188.

O'Connell, J., O. Schulz-Trieglaff, E. Carlson, M. M. Hims, N. A. Gormley, and A. J. Cox (2015). NxTrim: Optimized trimming of Illumina mate pair reads. Bioinformatics 31:2035–2037.

Paprocki, N., J. A. Heath, and S. J. Novak (2014). Regional distribution shifts help explain local changes in wintering raptor abundance: Implications for interpreting population trends. PLoS One 9:e86814.

Pritchard, J. K., M. Stephens, and P. Donnelly (2000). Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rowe, H. C., S. Renaut, and A. Guggisberg (2011). RAD in the realm of next-generation sequencing technologies. Molecular Ecology 20:3499–3502.

Ruegg, K. C., E. C. Anderson, K. L. Paxton, V. Apkenas, S. Lao, R. B. Siegel, D. F. DeSante, F. Moore, and T. B. Smith (2014). Mapping migration in a songbird using high-resolution genetic markers. Molecular Ecology 23:5726–5739.

Ruegg, K., R. A. Bay, E. C. Anderson, J. F. Saracco, R. J. Harrigan, M. Whitfield, E. H. Paxton, and T. B. Smith (2018). Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. Ecology Letters 21:1085–1096.

Ruegg, K. C., M. Brinkmeyer, C. M. Bossu, R. A. Bay, E. C. Anderson, C. W. Boal, R. D. Dawson, A. Eschenbauch, C. J. W. McClure, K. E. Miller, et al. (2021). Data from: The American Kestrel (*Falco sparverius*) genoscape: Implications for monitoring, management, and subspecies boundaries. Ornithology 138: 1–14. https://doi.org/10.5068/D1H38S

Ryman, N., S. Palm, C. André, G. R. Carvalho, T. G. Dahlgren, P. E. Jorde, L. Laikre, L. C. Larsson, A. Palmé, and D. E. Ruzzante (2006). Power for detecting genetic divergence: Differences between statistical methods and marker loci. Molecular Ecology 15:2031–2045.

Sauer, J. R., J. E. Hines, J. E. Fallon, K. L. Pardieck, D. J. Ziolkowski, Jr., and W. A. Link (2014). The North American Breeding Bird Survey, Results and Analysis, 1966–2012, Version 02.19.2014. USGS Patuxent Wildlife Research Center, Laurel, MD, USA.

Seutin, G., B. N. White, and P. T. Boag (1991). Preservation of avian blood and tissue samples for DNA analyses. Canadian Journal of Zoology 69:82–90.

Smallwood, J. A. (1990). Kestrel and Merlin. In Proceedings of the Southeast Raptor Management Symposium (B. A. Giron Pendelton, Editor). National Wildlife Federation, Washington, DC, USA. pp. 29–37.

Smallwood, J. A., and D. M. Bird (2020). American Kestrel (*Falco sparverius*), version 1.0. In Birds of the World (A. F. Poole and F. B. Gill, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA.

Smallwood, J. A., M. F. Causey, D. H. Mossop, J. R. Klucsarits, B. Robertson, S. Robertson, J. Mason, M. J. Maurer, R. J. Melvin, R. D. Dawson, et al. (2009). Why are American Kestrel (*Falco sparverius*) populations declining in North America? Evidence from nest-box programs. Journal of Raptor Research 43:274–282.

Tarailo-Graovac, M., and N. Chen (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Current Protocols in Bioinformatics 25:4.10.1–4.10.14.

Turbek, S. P., E. S. C. Scordato, and R. J. Safran (2018). The role of seasonal migration in population divergence and reproductive isolation. Trends in Ecology & Evolution 33:164–175.

Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, et al. (2013). From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Current Protocols in Bioinformatics 43:11.10.1–11.10.33.

Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. Journal of Heredity 89:438–450.

Waples, R. S., P. B. Adams, J. Bohnsack, and B. L. Taylor (2007). A biological framework for evaluating whether a species is threatened or endangered in a significant portion of its range. Conservation Biology: The Journal of the Society for Conservation Biology 21:964–974.

Weir, B. S., and C. C. Cockerham (1984). Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Willoughby, J. R., M. Sundaram, B. K. Wijayawardena, M. C. Lamb, S. J. A. Kimble, Y. Ji, N. B. Fernandez, J. D. Antonides, N. J. Marra, and J. A. Dewoody (2017). Biome and migratory behaviour significantly influence vertebrate genetic diversity. Biological Journal of the Linnean Society 121:446–457.

Winker, K. (2010). Subspecies represent geographically partitioned variation, a gold mine of evolutionary biology, and a challenge for conservation. In Avian Subspecies (K. Winker and S. M. Haig, Editors). Ornithological Monographs 67:6–23.

Zdobnov, E. M., and R. Apweiler (2001). InterProScan – An integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847–848.

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328.

Zink, R. M. (2014). Comparison of patterns of genetic variation and demographic history in the Greater Sage-Grouse (*Centrocercus urophasianus*): Relevance for conservation. Open Ornithology Journal 7:19–29.