PATTERNS OF GENETIC STRUCTURE IN THE AMERICAN KESTREL (*FALCO SPARVERIUS*): INFLUENCE OF DISTANCE AND MIGRATION AND IMPLICATIONS FOR MONITORING AND MANAGEMENT

by

Michaela Brinkmeyer

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Raptor Biology

Boise State University

December 2018

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Michaela Brinkmeyer

Thesis Title: Patterns of Genetic Structure in The American Kestrel (*Falco Sparverius*): Influence of Distance and Migration and Implications for Monitoring and Management

Date of Final Oral Examination: 03 October 2018

The following individuals read and discussed the thesis submitted by student Michaela Brinkmeyer, and they evaluated her presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Julie Heath, Ph.D.	Co-Chair, Supervisory Committee
Kristen C. Ruegg, Ph.D.	Co-Chair, Supervisory Committee
Stephen Novak, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Julie Heath, Ph.D., Co-Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

I would like to dedicate this work to my family and friends. The past two years have not been easy, but you were always there for me when I needed you most. You inspired me to do my best and encouraged me to push through the difficult times. For that, I am tremendously grateful. Thank you!

ACKNOWLEDGEMENTS

I would like to thank the many people who have helped make this project possible. First, I would like to acknowledge my major advisor and co-chair of my supervisory committee, Dr. Julie A. Heath. Julie is an excellent scientist, mentor and friend. She has shown me endless support, guidance and patience throughout my time at Boise State. I relied heavily on her knowledge of kestrel ecology, and I could not have done it without her help. I also want to thank the other co-chair of my supervisory committee, Dr. Kristen C. Ruegg. Kristen was a major contributor to this project and it would not have been possible without her expert knowledge, insight and guidance. Finally, I would also like to thank my committee member Dr. Stephen J. Novak. Steve taught me how to ask difficult questions; his curiosity and skepticism inspired me to be a better scientist. I am so grateful to have had the pleasure of learning from and working with such experienced academics and wonderful human beings.

This is a large collaborative project and it would not have been possible without my lab members at BSU and everyone associated with UCLA's Center for Tropical Research. First and foremost, I would like to thank Anjolene R. Hunt with Boise State University. Anjolene organized and implemented a large-scale database that enabled us to collaborate with research professionals and citizen scientists across North America. This project would not have been possible without her organizational skills. I would also like to thank UCLA for use of laboratory space and equipment and the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center for their assistance with

sequencing. I would specifically like to thank Winnie Le Underwood and Jasmine Rajbhandary for their help processing and preparing samples for sequencing at UC Davis. I would also like to recognize Rachael Bay and Christen Bossu. These women have shown me endless support and patience. They taught me most of what I know about coding from basic Unix commands to complex modeling in R. Your knowledge of bioinformatics, genomics, ecology and evolution never ceases to impress me. You are all very much appreciated.

This research was supported by a grant to Dr. Kristen C. Ruegg from the California Institute for the Energy and the Environment (POEA01-Z01), a grant to Dr. Julie A. Heath from SERDP (RC-2702), Boise State Raptor Research Center, private donations made to the American Kestrel Genoscape Project through Go-Fund-Me, and a generous donation from HawkWatch International.

This work would not have been possible without a collaborative sampling effort to collect high-quality DNA samples from American kestrels across North America. I would like to thank Karl Miller with Florida Fish and Wildlife Conservation Commission in Tallahassee, FL; Jean-Francois Therrien with Acopian Center for Conservation Learning at Hawk Mountain Sanctuary in Orwigsburg, PA; Ted Swem at the Anchorage Fish and Wildlife Field Office, U.S. Fish and Wildlife Service in Anchorage, AK; Clint W. Boal with the Department of Natural Resources Management at Texas Tech University in Lubbock, TX; Janet Eschenbach with Kestrel Research of Plainfield, WI; Joel Pagel of the U.S. Fish and Wildlife Service in Albuquerque, NM; Russell D. Dawson with the Ecosystem Science and Management Program at University of Northern British Columbia, Prince George BC, Canada; Rich Van Buskirk with the Department of

Environmental Studies at Pacific University in Forest Grove, OR; Jill and Lance Morrow with the Shenandoah Valley Raptor Study Area in Timberville, VA.; and many more partners that contributed samples to the American Kestrel Genoscape Project.

I would also like to thank my collaborators Thomas B. Smith at the Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles; Christopher J.W. McClure and Sarah Schulwitz with the Peregrine Fund's American Kestrel Partnership in Boise, ID; David Oleyar and Neil Paprocki with HawkWatch International in Salt Lake City, UT; and the staff with Research Computing at Boise State University in Boise, ID (doi:10.5281/zenodo.1646597). This research would not have been possible without your collaboration, insight and support. Thank you.

ABSTRACT

Genetic structure is useful for inferring patterns of selection, gene flow and connectivity, and can define management units that aid in interpretation of spatially-specific trends and species management. American kestrels (*Falco sparverius*) are a widespread, generalist species with fully migratory, partial migrant, and resident populations. In many parts of their range, kestrels show evidence of declining population trends; however, it has been difficult to identify threats to kestrels because of differences in regional trends. We used a genome-wide sequencing approach to investigate the genetic structure of American kestrels, test hypotheses about the processes that influence genetic structuring of populations by affecting dispersal and gene flow and suggest new approaches for kestrel management based on genetic information. Specifically, we sequenced the first American kestrel genome and used restriction site associated DNA (RAD) sequencing to assess population structure at 72,263 SNP markers screened in 12 populations from across the migratory and non-migratory range of two subspecies of the American kestrel (*F. s. sparverius* and *F. s. paulus*) in North America. We revealed previously unrecognized amounts of population genetic structure in American kestrels. We found the highest amount of genetic differentiation between resident populations, followed by moderate levels of differentiation between migratory and resident populations, and the lowest amount of genetic differentiation between long-distance migrants. These results suggest that migratory behavior facilitates dispersal, increases gene flow, and therefore reduces the amount of genetic differentiation and structuring

between populations. Further, we suggest that genetically distinct groups of kestrels be monitored and managed separately to identify limiting factors that may affect these groups differently. This information increases our understanding of migrant ecology and evolution and has important implications for management of American kestrels.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

An important application of population genetics is the identification of genetically distinct groups, and the evolutionary forces that have shaped these groups. Population genetic structure is determined by the interplay between gene flow, genetic drift, and natural selection, and is influenced by evolutionary history (patterns of divergence), the historical biogeography of taxa (glaciation) and landscape features (mountains) (Slatkin 1985; Barton & Clark 1990; Bohonak 1999). The identification of genetically distinct groups also provides important information for conservation and management. Such groups are often referred to as management units (MUs), evolutionary significant units (ESU's), conservation units, or distinct population segments, depending on the context (Moritz 1994; Allendorf & Luikart 2007; Funk et al. 2007). The amount of genetic structure among populations has been frequently used to define genetically distinct units for conservation management. However, it is often difficult to identify genetically distinct units in highly mobile species with large population sizes such as migratory birds where gene flow is high and the effect of genetic drift in large populations is expected to be small (Willoughby et al. 2017; Doyle et al. 2018; Medina et al. 2018). As a result, defining genetically distinct conservation units in migratory animals such as birds has remained a challenge (Larson et al. 2013; Zink 2014; Freer et al. 2015; Mura-Jornet et al. 2018).

Traditionally, genetic studies 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, it is often challenging to detect subtle genetic differentiation in high dispersal species with a limited number of markers (Younger et al. 2015). Recent advances in sequencing technology have enabled researchers to obtain genotypic information for non-model organisms at a large number of molecular markers (Rowe et al. 2011). Restriction-site associated DNA (RAD) sequencing is a next-generation sequencing (NGS) technology that uses enzymatic fragmentation of the genome for the discovery of large numbers of genome-wide single-nucleotide polymorphisms (SNPs; Baird et al. 2008). 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 (even cryptic genetic structure) (Ruegg et al. 2014; Benestan et al. 2015; Jahner et al. 2016) and provided opportunities for fine scale investigations of population boundaries (Larson et al. 2013; Fredrickson et al. 2015; Van Den Bussche et al. 2017; Mura-Jornet et al. 2018). High-resolution approaches may be particularly important for widespread species where there may be a continuous gradient in the degree of differentiation between individuals within and among populations (Waples & Gaggiotti 2006), making it difficult to define units that are relevant to conservation management.

Ecological and behavioral differences between populations can also help with defining conservation units in highly mobile species, for which patterns of population structure are difficult to detect (Fraser & Bernatchez 2001; Oyler-McCance et al 2005; Geist 2010; Funk et al. 2012). In migratory birds, differences in migratory behavior –

either the direction of migration or the degree to which individuals within a population migrate – may influence the extent of genetic differentiation (Medina et al. 2018; Pruett et al. 2008), although examples of this remain limited. In particular, partially migratory species provide ideal systems for studying the influence of migratory strategy on gene flow because individuals within the same population and among populations within the same species that exhibit differences in migratory behavior (Lack 1943; Lack 1944; Berthold 2001). When partially migratory species have large geographic distributions, populations are likely to follow a gradient of migratory behavior that ranges from obligatory longer-distance movements in the north, to complete residency in the south, with a range of partial migration in-between (Cohen 1967; Lundberg 1988). Such differences may result in limited gene flow between distinct migratory forms. Here we use NGS sequencing technology to investigate the influence of differences in migratory behavior on population genetic structure in the partially migratory American kestrel (*Falco sparverius*). In doing so, we initiate the process of developing a framework for conservation and management for this and other highly mobile species with differences in migratory behavior.

The American kestrel is a partial migrant that breeds throughout North and South America (Smallwood & Bird 2002) and has upwards of 17 recognized subspecies (Ferguson-Lees & Christie 2001). In North America, kestrels show highly variable migration strategies, including individuals that migrate long distances, short distances, or do not migrate, and populations that are complete migrants, partial migrants, or complete non-migrants (Layne 1982; Henny & Brady 1994). American kestrels in the southeastern United States are year-round residents and have been designated as a separate subspecies

(*F. s. paulus*) from the nominate subspecies (*F. s. sparverius*) that are found throughout the remainder of North America (Hoffman & Collopy 1988; Smallwood 1990). In general, it is thought that populations of *F. s. sparverius* follow a pattern of leap-frog migration, where individuals and populations in the northernmost part of the kestrel's range migrate the furthest distances and follow a gradient of decreasing migratory distance to complete residency in the south (Heath et al. 2012). Furthermore, there is evidence kestrel populations are declining, but estimates of demographic trends differ from various regions and with different metrics (McClure et al. 2017). Furthermore, it is often uncertain which breeding populations are being monitored during migration counts, which makes it difficult to interpret long-term population trends. Therefore, defining genetically distinct units across the geographic distribution of kestrels will improve our ability to interpret demographic trends.

Previous work investigating the role of migratory strategy on population genetic structure in American kestrels found only subtle differences in population structure across the kestrel's North American range, but their results were limited by the number of markers employed, and the narrow range of their sampling design (Miller et al. 2012). Here we used RAD-sequencing to evaluate patterns of genome-wide population structure in the American kestrel. We collected samples throughout the kestrel's North American range and included replicate comparisons of migratory to non-migratory populations in the western and eastern regions. Specifically, we asked: (1) can we detect increased levels of genetic structure using genome-wide molecular data within and between groups with differences in migratory strategy compared to previous genetic studies? (2) How does migration distance influence population genetic structure? If migratory distance

influences gene flow, we would expect to see greater genetic structuring and stronger signals of genetic divergence among non-migratory compared to migratory populations. (3) Can we use these data to improve conservation? We discuss our results in light of their implications for a better understanding of the ecology and evolutionary biology of migratory organisms more generally, as well as specifically for the conservation management of American kestrels in North America.

METHODS

Sample collection and DNA extraction

In 2015 and 2016, we captured and sampled unrelated breeding adult or nestling American kestrels from 12 sites along the boundaries of the American kestrel North American range. This sampling was done in collaboration with several non-profit organizations, state agencies, university researchers, and citizen-scientists (Table 1; Table SI 1). Sampling locations were selected to represent the outer edges of the kestrel's North American breeding range. We collected blood (~ 30 µl) via brachial or jugular venipuncture from 198 individuals across 10 sampling locations and preserved blood in lysis buffer or anticoagulant tubes, then we stored samples at -80^{o}C until analysis (Seutin et al. 1991). Alternatively, we collected 3 - 4 feathers from the breast of one nestling per brood from an additional 89 individuals across five of the original sampling locations, and two additional sites for a total of 12 sampling locations. These feathers contained a small amount of blood in the feather tip because they were collected when the feathers were growing. Feathers were stored in envelopes at room temperature until analysis, then the tips were clipped for DNA extraction. We extracted genomic DNA with Qiagen DNeasy Blood and Tissue Kits and used Qubit dsDNA HS157 Assay kits to quantify and select samples with intact, high molecular weight DNA (Thermo Fisher Scientific) for sequencing.

Genome sequencing, assembly and annotation

We used the Illumina TruSeq DNA PCR-Free LT kit (Illumina), following the adjustments made by Ruegg et al. (2018), to prepare a genomic DNA library from a single individual from Boise, Idaho. We prepared the genomic DNA library for sequencing on two lanes of an Illumnia HiSeq2500 using 250 base-pair (bp) paired-end sequencing at the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center, supported by NIH Shared Instrumentation Grant 1S10OD010786-01. We assembled initial contigs using the 250bp paired end reads with the Discovar DeNovo assembler from the Broad Institute (http://www.broadinstitute.org), discarding contigs less than 1000bp in length. We trimmed paired end reads with NxTrib (O'Connell et al. 2015) and generated scaffolds with SSPACE (overlap requirement k=3) (Boetzer et al. 2010). We repeated the SSPACE scaffolding with k=5 and discarded scaffolds less than 1000bp for the final assembly.

For annotation purposes, we used REPEATMASKER (-species birds) (Tarailo-Graovac & Chen 2009) to replace repetitive regions of the final genome assembly with N's. We used two different ab initio gene predictions within the MAKER pipeline (Cantarel et al. 2008): SNAP and AUGUSTUS. Then, we used Zebra Finch cDNA and protein sequences downloaded from Ensembl to train SNAP, and the available chicken training dataset to train AUGUSTUS. We used ITERPROSCAN (Zdobnov & Apweiler 2001) to add Pfam protein annotation and gene ontology (GO) terms and identified 13,342 genes. Using the software PROMER, part of the MUMMER package (Delcher et al. 2003), we aligned scaffolds to the Zebra Finch genome (version 3.2.4). After alignment, we retained the longest consistent alignment (-q) for each chromosome while filtering for

similarity (-I 50) and alignment length (-l 500). We then determined the location of the longest alignment for each scaffold and ordered scaffolds accordingly for visualization purposes.

SNP discovery and SNP filtering

We conducted genome scans on 287 individuals following a modified version of the BestRAD library preparation protocol (Table SI 1; Ali et al. 2016). We normalized DNA from every sample to a final concentration of 100ng in a 10ul volume, digested DNA by sbfl restriction enzyme (New England Biolabs, NEB) and ligated DNA fragments with SBfI adapters prepared with biotinylated ends. After ligation, we pooled and cleaned samples using 1X Agencourt® AMPure XP beads (Beckman Coulter), and sheared all DNA fragments to an average length of 400bp with 10 cycles on the Bioruptor NGS sonicator (Diagenode). We used the Illumina NEBNext Ultra DNA Library Prep Kit [New England Biolabs (NEB)] to repair blunt ends and ligate NEBNext Adaptors onto the blunt ends of all DNA fragments. Then, we used Agencourt® AMPure XP beads (Beckman Coulter) to select DNA fragments with an average of 500bp, enriched libraries with PCR, and cleaned final libraries with Agencourt® AMPure XP beads (Beckman Coulter). We generated 250bp paired end sequencing reads for 287 individuals on three lanes of an Illumina HiSeq 2500 at the UC Davis Genome Center, and re-sequenced 66 individuals with low coverage on a forth lane.

We used the program STACKS (Catchen et al. 2013) to demultiplex, filter and trim adapters from the data with the process_radtags function and remove duplicate read pairs using the clone_filter function. We used BOWTIE2 to map reads to the genome, and the Haplotype caller in the Genome Analysis Toolkit to identify SNPs. Finally, we used

VCFTOOLS (Danecek et al. 2011) to remove indels, non-biallelic SNPs, and low quality and rare variants (genotype quality 20; coverage depth 10; minor allele frequency 0.05). We determined the final number of SNPs and individuals to be retained in further analyses by visualizing the tradeoff between discarding low coverage SNPs and discarding individuals with missing genotypes using GENOSCAPERTOOLS. In addition, we discarded individuals with >30% heterozygosity and loci with >50% heterozygosity. Finally, we removed individuals that were outliers in our analysis of principal components. We did not remove outlier SNPs from our analyses of genetic structure because we were interested in capturing the full spectrum of genomic differences between geographically distant populations, including SNPs that were under putative selective pressure and potentially involved with local adaptation.

Population genetic structure

We examined genetic structure using all SNPs retained after filtering based on sequencing coverage and missing data per individual and per locus. We identified genetic clusters with principal components analysis (PCA) using the R packages GENOSCAPERTOOLS and SNPRELATE (Anderson 2017, Zheng et al. 2012). Then, we used ADMIXTURE version 1.3.0 (Alexander et al. 2009) and the R package TESS3R (Caye et al. 2015) to further assess patterns of genetic structure across the breeding range. ADMIXTURE uses a maximum likelihood model to estimate individual ancestry proportions from genome-wide SNP datasets. Unlike ADMIXTURE, TESS3R uses a statistical model that incorporates geography to estimate ancestry coefficients among populations with low-levels of genetic divergence (Durand et al. 2009). We performed

each analysis using 5 replicates of each assumed number of clusters (K), where K ranged from 1:10.

Genetic distance correlations

We ran distance correlation analyses using all locations with at least five individuals remaining after filtering; two populations were removed including one from southern California and another from Fairfield, Idaho; and thus, 194 individuals from 10 sampling locations were included in this analysis. We calculated pairwise F_{ST} between each population using 72,263 SNPs in the R package SNPRELATE (Zheng et al. 2012), and pairwise geographic distance from the central longitudinal and latitudinal coordinates of each location using the Vincenty ellipsoid method in the R package GEOSHPERE (Hijmans 2011). We also calculated pairwise F_{ST} and bootstrapped confidence intervals for each pairwise comparison between the 10 sampling locations used in our analysis of genetic distance correlations using the R package ASSIGNER. As a proxy for migration strategy, we calculated the mean migration distance for each location using band and encounter data for American kestrels in North America from the Bird Banding Laboratory (BBL; USGS Bird Banding Laboratory 2017). Because we were interested in estimating migration distance of breeding populations at each sampling location, we filtered the encounter database to include individuals that were either banded during the breeding season (April 15 - August 15), and encountered during winter (November - February), or vice versa. Then, we used the Vincenty ellipsoid method in the R package GEOSHPERE to calculate the true encounter distance for each band encounter record included in our analysis. The Vincenty ellipsoid method accounts for the curvature of the earth by calculating the great-circle distance between two locations (Hijmans 2011),

which is an important dimension for long distance migrant populations. Then we used the average distance to represent each site's migration distance. We classified sites with average migration distance > 2000 km to be long distance, > 75 km to be short distance, and considered Florida, southern California, and Texas sites as nonmigratory. We used Mantel tests and multiple regression of distance matrices (MRM) to test for associations between linearized F_{ST}, which is calculated as (F_{ST} /1- F_{ST}), and both geographic distance and pairwise differences in migration distance.

RESULTS

Genome sequencing, assembly and annotation

The final American kestrel genome assembly is 1.23 Gb in length and consists of 5,096 scaffolds. This genome sequence was successfully used to align RAD sequencing data.

SNP discovery and SNP filtering

We used RAD-sequencing data from 287 individuals to identify 199,705 bi-allelic loci with a minor allele-frequency greater than 5%, minimum quality score greater than 20, and minimum sequencing depth greater than 10. We also filtered based on missing data and assessed the tradeoff between low coverage SNPs and missing genotypes and retained 204 individuals and 75,000 loci (Figure SI 1). From the filtered data set, we removed two individuals with greater than 35% heterozygosity, and an additional five individuals that were PCA outliers. Also, we removed loci with greater than 50% heterozygosity. The final data set used to assess genetic structure and design assays consisted of 197 individuals (Table SI 1) and 72,263 SNPs.

Population genetic structure

A principal components analysis based on RAD-sequence data showed evidence for 3 main clusters with separation of Florida from eastern sampling locations and separation of eastern from western sampling locations (Figure 3). Alaska and Texas separated from other western sampling locations but did not form distinct clusters in this

analysis. Overall, the first three principle components explained less than 3% of the total variation in allele frequencies across 72,263 SNPs, suggesting low genetic structure.

Results from the TESS3R clustering analysis suggested that kestrels separated into five genetically distinct clusters, with separation of Alaska, Texas, and Florida from the eastern and western sampling locations (Figure 1A). Texas and Florida formed the most distinct clusters, followed by Alaska and regional clustering of eastern and western breeding populations. Results from ADMIXTURE showed a clinal pattern of genetic separation within the eastern and western sampling locations, with nonmigratory populations in Florida and Texas exhibiting the most differentiation, which is consistent with the results of PCA and clustering analysis performed with TESS3R.

Results from the clustering analysis that we performed using the R package Tess3r suggested there are five genetic clusters, with separation of Alaska, Texas, and Florida from the eastern and western sampling locations (Figure 1A). Texas and Florida formed the most distinct clusters, followed by Alaska and regional clustering of eastern and western breeding populations. A principal coordinates analysis plot showed evidence for 3 main clusters with separation of Florida from eastern sampling locations and separation of eastern from western sampling locations (Figure 1B). Alaska and Texas are also separated from other western sampling locations, but do not form independent clusters. The first three principle components explained less than 3% of the total variation in allele frequencies across 72,263 SNPs. Results from Admixture showed a clinal pattern of genetic separation within the eastern and western sampling locations, with nonmigratory populations in Florida and Texas exhibiting the most differentiation, which

is consistent with the results of PCA and clustering analysis performed with Tess3r (Figure SI 3).

Pairwise F_{ST} between all 10 sampling locations ranged from 0-0.0177 (Table 2). None of the confidence intervals associated with these values overlapped zero (Table SI 2), indicating that levels of genetic differentiation between sites are low, yet significant. Kestrels in Florida and Texas were the most genetically different from each other, and from kestrels from other sites. We found the highest estimates of pairwise F_{ST} between non-migratory (Texas and Florida) and long-distance migratory populations in Alaska and Saskatchewan. The lowest estimates of pairwise F_{ST} were between kestrels from migratory populations within eastern and western breeding areas.

Genetic distance correlations

We included both nonmigratory and migratory populations in a MRM analysis and found a positive correlation between F_{ST} and geographic distance ($\beta = 0.506$; $P = 0.002$) that suggests that isolation by distance contributes to patterns of genetic differentiation. We also found a tendency for a negative correlation between F_{ST} and migratory distance ($\beta = -0.355$; $P = 0.074$) that suggests that migratory populations had lower genetic structure than nonmigratory populations (Figure 2). Overall, the MRM accounted for about 25% of the variation in F_{ST} ($R^2 = 0.251$ $P = 0.033$) and suggests that the correlation between F_{ST} and migratory distance remains informative even after accounting for the underlying relationship between genetic and geographic distance in the model.

DISCUSSION

Identifying genetically distinct conservation units is an important first step in the management of declining populations (Allendorf & Luikart 2007; Funk et al. 2012). For decades, conservation biologists have used population genetic techniques to quantify genetic distinctiveness of populations, but this remains a challenge for highly mobile species with high dispersal capabilities. In the past, the ability to detect subtle patterns of population genetic structure in animals with a high capacity for dispersal has been limited by a lack of markers with sufficient resolution (Ruegg et al. 2014; Benestan et al. 2015; Younger et al. 2015; Jahner et al. 2016). Here we sequence the first American kestrel genome and assess population structure at 72,263 SNP markers screened in 12 populations from across the migratory and non-migratory range of the American kestrel and find previously unrecognized amounts of population genetic structure. We provide evidence for five genetically-distinct breeding populations of kestrels in North America, with the most significant differences between non-migratory and migratory populations, followed by regional separation of eastern and western breeding populations. We suggest that migratory behavior influences regional patterns of genetic structure by increasing dispersal and gene flow between geographically isolated populations. This information has important implications for understanding and interpreting demographic trends by providing evidence for genetic divergence between populations from eastern and western North America, Alaska, Texas and Florida.

Overall, levels of genetic differentiation in American kestrels were low as indicated by the low percentage of variation explained by the first two axes of PCA, and by the small estimates of genetic differentiation among populations. Such low levels of differentiation are consistent with past studies based on fewer loci that also generated small estimates of genetic differentiation (Miller et al. 2012). Low genetic structure suggests that there is gene flow throughout the range of the kestrel that is facilitated by migratory behavior. Further, low levels of genetic structure between *F. s. sparverius* and *F. s. paulus* may suggest there has been recent gene flow between the two putative subspecies, or recent divergence. This result is consistent with recent research showing that long-distance dispersal (> 30 km) is relatively common in American kestrels (McCaslin et al. *in prep*). Other raptor species have shown low levels of genetic structure, and high amounts of gene flow across their North American range. Doyle et al. (2018) suggested that relatively few long-distance dispersal events were enough to result in genetic homogenization of the prairie falcon across western North America. Additionally, understanding how populations are connected via gene flow has important implications for conservation management of the species.

Although genetic structure was low, we detected patterns of genetic differentiation that indicate geography influences genome-wide patterns of genomic structure among populations of American kestrels in North America. Consistent with a pattern of isolation by distance, geography explained a large proportion of among population genetic variation, estimates of population level genetic distance were strongly correlated to geographic distance, and regional groups of sampling locations clustered together in ordination space. The effect of geographic distance may be compounded by

migration routes of kestrels. Like most raptors, kestrels follow a strong north-to-south pattern of migration, with little longitudinal drift, and the probability of long-distance migration into Mexico increases from east-to-west (Mueller & Berger 1967; Evans & Rosenfield 1985; Goodrich & Smith 2008). Spatial separation may result in divides between kestrels along the coastal flyways with mixing in the central flyway. This would be consistent with the pattern of genetic divergence that we observed between eastern and western regional breeding populations. In addition to the separation of east and west, Alaska and Texas individuals appeared distinct. This is a novel finding and not consistent with previous studies that found structure between subspecies using mtDNA sequences, but did not detect genetic differences between eastern and western breeding populations of kestrels in North America. Perhaps previous genetic studies were limited by the approach or they did not sample the complete range of genetic variation and therefore, were not able to detect subtle patterns of genetic structure between eastern and western breeding populations.

We provide evidence for genetic differentiation between breeding populations in eastern and western North America, Alaska, Texas and Florida. These results demonstrate regional differences in genetic structure that are consistent with regional variation in migratory behavior of American kestrels in North America. By including samples from more than one nonmigratory, and two completely migratory breeding populations, our results support and expand upon previous findings that suggest migratory behavior influences patterns of population structure in the American kestrel (Miller et al. 2012). We found the highest amount of genetic differentiation between resident populations (Texas and Florida), followed by moderate levels of differentiation

between migratory and resident populations, and the lowest amount of genetic differentiation between long-distance migrants. We also found that resident populations in Texas and Florida are well differentiated within ordination space. In addition, matrix correlation analyses suggest average population-level estimates of migratory distance are a good predictor of genetic distance, even after accounting for geographic distance between populations. Our findings suggest that migratory behavior facilitates dispersal and increases gene flow among migratory and partially migratory populations, which reduces the amount of genetic differentiation and structuring between populations with similar migratory strategies (populations within the west and within the east). Long-distance movements including both dispersal and seasonal migratory movements are controlled by similar intrinsic (e.g. genetic, hormonal, etc.) and extrinsic (environmental, social, etc.) factors (Dingle & Drake 2007; Dufty & Belthoff 2001). These shared mechanisms for dispersal and migration are one potential reason we see a positive correlation between migration strategy and dispersal or gene flow, demonstrated here as the relationship between migratory distance and genetic differentiation between geographically distant populations.

The novel finding of five distinct groups may aid in American kestrel monitoring and management. Kestrels are declining at differential rates across their distributional range and for unknown reasons (Butcher 1990; Smallwood et al. 2009; Sauer et al. 2014). For example, migration count data indicate that kestrels in the western U.S. began declining in the late 1990s (Farmer & Smith 2009). Iin contrast, data from the CBC suggest kestrels were declining from 1975 to the late 1990s, but have experienced recent population stabilization (Paprocki et al. 2014). Heterogeneity in rates of population

declines across North America suggest that regional populations are experiencing different threats, or are responding to them differentially (McClure et al. 2017). Conflicting trends among datasets make it difficult to discern potential causes of decline. Considering kestrels in eastern and western North America, Alaska, Texas and Florida as belonging to separate groups for conservation purposes will improve our ability to interpret regional differences in population demographic trends and facilitate studies of potential causes of decline within each genetically-distinct group.

Given the regional differences in demographic trends, it is likely that eastern and western populations are experiencing different stressors, and designating these regional groups as separate entities will improve our ability to interpret estimates of demographic trends and to determine underlying causes of regional declines. In addition, focus should be paid to breeding populations of kestrels in Texas and Alaska because those populations have not been monitored under the knowledge that they are distinct from other kestrels in North America. It is possible that resident breeding kestrels in Texas are experiencing different population constraints in comparison to breeding kestrels in eastern and western North America. Similarly, it is important that we monitor kestrels in Alaska as an independent conservation unit and determine migratory patterns (including distance and direction of migration), so that we can start to understand where those populations may be limited most.

Further, considering eastern and western regional groups as independent units is important because responses to climate change vary regionally across the kestrel's geographic range. In western North America, American kestrel migration distances have significantly decreased (Heath et al. 2012), wintering distributions have shifted northward

(Paprocki et al. 2014), migration phenology has changed (Oleyar & Hawks *in prep*.), breeding phenology has changed (Heath et al. 2012), and population trends have remained stable (Steenhof & Peterson 2009). However, migration distances of American kestrels in eastern North America have not decreased over time, poleward shifts in wintering distributions have not been as drastic, nesting phenology has not changed (Smallwood, Therrin, Miller, unpub data), and populations are declining (Smallwood et al. 2009; Brandes et al. 2016; Crewe et al. 2016; Figure SI 4). Given the regional pattern of genetic structure and differences in response to climate change, including regional differences in population decline, it is likely that kestrels are experiencing different stressors in different parts of their range. Therefore, our identification of five distinct genetic groups, which should be considered five conservation units, will improve our ability to interpret estimates of population trends and enhance our understanding of how populations are shifting their migratory movements; information that is crucial for effective conservation management of the American kestrel across their North American range.

REFERENCES

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

Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C et al. (2016) RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics*, 202, 389-400.

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

Anderson EC (2017) eriqande/genoscapeRtools. GitHub repository, doi: 10.5281/zenodo.848279.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA & Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3(10), e3376.

Barton N & Clark A (1990) Population structure and processes in evolution. *Population Biology*, 115-173.

Benestan L, Gosselin T, Perrier C, Sainte-Marie B, Rochette R & Bernatchez L (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.

Berthold P (2001) Bird migration. A general survey, 2nd ed. Oxford Univiversity Press.

BirdLife International, NatureServe, "Bird species distributions maps of the world" (BirdLife International, Cambridge, UK).

Boetzer M, Henkel CV, Jansen HJ, Butler D & Pirovano W (2010). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, 27, 578-579.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. Quarterly Review of Biology, 74 (1), 21–45.

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

Butcher GS (1990) Audubon Christmas Bird Counts. Pages 5–13 in Sauer JR & Droege S [EDS.], Survey designs and statistical methods for the estimation of avian population trends. Biological Report 90(1). U.S.D.I. Fish and Wildlife Service, Washington, DC U.S.A.

Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B et al. (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Resources*, 496 18, 188-196.

Catchen J, Hohenlohe PA, Bassham S, Amores A & Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22, 3124-3140.

Caye K, Deist TM, Martins H, Michel O & Francois O (2015) TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, n/a–/a. doi:10.1111/1755- 0998.12471.

Cohen D (1967) Optimization of seasonal migratory behavior. *American Naturalist*, 101, 5–17.

Crewe T, Taylor P, Lepage D, Goodrich L, Brown J, & Sodergren J (2016). The Raptor Population Index, 2016 Analysis Methods and Trend Results. Available at http://rpi-project.org/2016/

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156-2158.

Delcher AL, Salzberg SL & Phillippy AM (2003) Using MUMmer to identify similar regions in large sequence sets. *Current Protocols in Bioinformatics*, 10.13. 11-10.13. 18.

Dingle H & Drake VA (2007) What is Migration? *BioScience*, 57(2), 113-121.

Doyle JM, Bell DA, Bloom PH, Emmons G, Fesnock A, Katzner TE, La Pre L, Leonard K, San Miguel P, Westerman R & DeWoody JA (2018) New insights into the phylogenetics and population structure of the prairie falcon (*Falco mexicanus*). *BMC Genomics*, 19(233), doi: 10.1186/s12864-018-4615-z.

Dufty AM Jr., JR Belthoff (2001). "Proximate Mechanisms of Natal Dispersal: The Role of Body Condition and Hormones". Dispersal, 217-229.

Durand E, Jay F, Gaggiotti OE & Francois O (2009) Spatial inference of admixture proportions and secondary contact zones. *Molecular Biology and Evolution*, 26, 1963–1973.

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

Evans DL & Rosenfield RN (1985) Migration and mortality of sharp-shinned hawks ringed at Duluth, Minnesota, USA. *ICBP Technical Publication*, 5, 311-316.

Farmer C & Smith J (2009) Migration Monitoring Indicates Widespread Declines of American kestrels (*Falco sparverius*) in North America. *Journal of Raptor Research*, 43(4), 263-273.

Ferguson-Lees J & Christie DA (2001) Raptors of the world. Houghton Mifflin Co., Boston, MA U.S.A

Fraser DJ & Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, 10, 2741-2752.

Fredrickson RJ, Hedrick PW, Wayne RK, Von Holdt BM & Phillips MK (2015) Mexican wolves are a valid subspecies and an appropriate conservation target. *Journal of Heredity*, 1-2.

Freer JJ, Mable BK, Clucas G, Rogers AD, Polito MJ, Dunn M, Naveen R, Levy H & Hart T (2015) Limited genetic differentiation among chinstrap penguin (*Pygoscelis antarctica*) colonies in the scotia arc and western Antarctic peninsula. *Polar Biology*, 38:1493–502.

Funk CW, Mullins TD & Haig SM (2007) Conservation Genetics of Snowy Plovers (*Charadrius alexandrines*) in the Western Hemisphere: Population Genetic Structure and Delineation of Subspecies. *Conservation Genetics*, 8, 1287-1309.

Funk CW, McKay JK, Hohenlohe PA & Allendorf FW (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, 27(9), 489-496, doi:10.1016/j.tree.2012.05.012.

Geist J (2010) Strategies for the conservation of endangered freshwater pearl mussels (*Margaritifera margaritifera* L.): a synthesis of Conservation Genetics and Ecology. *Hydrobiologia*, 644(1), 69-88.

Goodrich LJ & Smith JP (2008) Raptor migration in North America. In K.L. Bildstein KL, Smith JP, Ruelas Inzunza E & Veit RR [Eds.], State of North America's Birds of Prey. Series in Ornithology, No. 3. Nuttall Ornithological Club, Cambridge, MA U.S.A. and American Ornithologists' Union, Washington, DC U.S.A.

Heath, JA, Steenhof K & Foster MA (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 CJ & Brady GL (1994) Partial migration and wintering localities of American kestrels nesting in the Pacific Northwest. *Northwest Naturalist*, 75, 37-43.

Hijmans RJ (2011) Introduction to the "geosphere" package (Version 1.2-19).

Hoffman ML & Collopy MW (1988) Historical Status of the American kestrel (*Falco sparverius paulus*) in Florida. *The Wilson Bulletin*, 100(1), 91-107.

Jahner JP, Gibon D, Weitzman CL, Blomberg EJ, Sedinger JS & Parchman TL (2016) *BMC Evolutionary Biology*, 16:127. doi 10.1186/s12862-016-0702-4.

Lack D (1943) The problem of partial migrations. *British Birds*, 37, 143-150.

Lack D (1945) The Galapagos finches (Geospizinae), a study in variation. *Occassional Papers in California Academy of Science*, 21, 1-152.

Larson WA, Seeb LW, Everett MV, Waples RK, Templin WD & Seeb JE (2013) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, 7, 355-369.

Layne JN (1982) Analysis of Florida-related banding data for the American kestrel. *North American Bird Bander*, 7:94-99.

Lundberg P (1988) The evolution of partial migration in Birds. *Trends in Ecology and Evolution*, 3(7), 172-175.

McCaslin HM, Caughlin TT, & Heath JA. *in prep*. Frequency and correlates of long-distance dispersal in a widespread raptor. *Ecology*.

McClure CJW, Schulwitz SE, Van Buskirk R, Pauli BP & Heath JA (2017) Commentary: research recommendations for understanding the decline of American kestrels (*Falco sparverius*) across much of North America. *Journal of Raptor Research*, 51(4), 455-464.

Medina I, Cooke GM & Ord TJ (2018) Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecology Letters*, 21(5).

Miller MP, Mullins TD, Parrish Jr. JW, Walters JR & Haig SM (2012) Variation in Migratory Behavior Influences Regional Genetic Diversity and Structure among American kestrel Populations (*Falco sparverius*) in North America. *Journal of Heredity*, doi:10.1093/jhered/ess024.

Morin PA, Martien KK & Taylor BL (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 and Evolution*, 9, 373-375.

Mueller HC & Berger DD (1967) Wind drift, leading lines, and diurnal migration. *The Wilson Bulletin*, 79, 50-63.

Mura-Jornet I, Pimentel C, Dantas GPM, Petry MV, Gonzalez-Acuna D, Barbosa A, Lowther AD, Kovacs KW, Poulin E & Vianna JA (2018) Chinstrap penguin population genetic structure: one or more populations along the Southern Ocean? *BMC Evolutionary Biology*, 18:90.

O'Connell J, Schulz-Trieglaff O, Carlson E, Hims MM, Gormley NA & Cox AJ (2015) NxTrim: optimized trimming of Illumina mate pair reads. *Bioinformatics*, 31, 2035-2037.

Oleyar MD, & Hawks S. *In prep*. Phenology shifts in migration of diurnal raptors in western North America. *The Auk*.

Oyler-McCance SJ, St. John J, Taylor SE, Apa AD & Quinn TW (2005) Population genetics of Gunnison sage-grouse: implications for management. *The Journal of Wildlife Management*, 69(2), 630-637.

Paprocki N, Heath JA & Novak SJ (2014) Regional Distribution Shifts Help Explain Local Changes in Wintering Raptor Abundance: Implications for Interpreting Population Trends. *PLOS One*, 9(1), e86814

Pruett CL, Arcese P, Chan YL, Wilson AG, Patten MA, Keller LF & Winker K (2008) The effects of contemporary processes in maintaining the genetic structure of western song sparrows (*Melospiza melodia*). *Heredity*, 101, 67-74.

Rowe HC, Renaut S & Guggisberg A (2011) RAD in the realm of next-generation sequencing technologies. *Molecular Ecology*, 20, 3499-3502.

Ruegg KC, Anderson EC, Paxton KL, Apkenas V, Lao S, Siegel RB, Desante DF, Moore F & Smith TB (2014) Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23, 5726-5739.

Ruegg KC, Bay RA, Anderson EC, Saracco JF, Harrigan RJ, Whitfield M, Paxton EH & Smith TB (2018) Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters*, 21(7), 1085-1096.

Ryman N, Palm S, André C, Carvalho GR, Dahlgren TG, Jorde PE, Laikre L, Palme A & Ruzzante DE (2006) Power for detecting genetic divergence: differences between statistical methods and marker loci. *Molecular Ecology*, 15, 2031–2045.

Sauer JR, Hines JE, Fallon JE, Pardieck KL, Zoilkowski DJ & Link WA (2014) The North American Breeding Bird Survey, results and analysis 1966–2012. Version 02.19.2014. U.S.G.S. Patuxent Wildlife Research Center, Laurel, MD U.S.A.

Seutin G, White BN & Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, 69, 82-90.

Slatkin M (1985) Gene Flow in Natural Populations. *Annual Review of Ecology and Systematics*, 16, 393-430.

Smallwood JA (1990) Kestrel and Merlin. In: Giron-Pendleton BA, editor. Proceedings of the Southeast Raptor Management Symposium and Workshop. Washington (DC): National Wildlife Federation. p. 29–37.

Smallwood JA & Bird DM (2002) American kestrel (*Falco sparverius*). In A. Poole and F. Gill [EDS.], The birds of North America, No. 602. The Academy of Natural Sciences, Philadelphia, PA and the American Ornithologists' Union, Washington, DC U.S.A.

Smallwood JA, Causey MF, Mossop DH, Klucsarits JR, Robertson B, Robertson S, Mason J, Maurer MJ, Melvin RJ, Dawson RD, Bortolotti GR, Parrish JW, Breen TF & Boyd K (2009) Why are American kestrel (*Falco sparverius*) populations declining in North America? Evidence from nest-box programs. *Journal of Raptor Research*, 43, 274-282.

Steenhof K & Peterson BE (2009) American kestrel reproduction in southwestern Idaho: annual variation and long-term trends. *Journal of Raptor Research*, 43, 283-290.

Tarailo-Graovac M & Chen N (2009) Using RepeatMasker to identify repetitive elements in genomic sequences. *Current protocols in bioinformatics*, 4.10. 11-14.10. 14.

USGS Bird Banding Laboratory (2017) Unpublished data retrieved {month, day}. Patuxent Wildlife Research Center, Laurel, MD.

Van Den Bussche RA, Judkins ME, Montague MJ & Warren WC (2017) A resource of genome-wide single nucleotide polymorphisms (SNPs) for the conservation and management of Golden Eagles. *Journal of Raptor Research*, 51(3), 368-377.

Waples RS & Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15, 1419-1439.

Willoughby JR, Sundaram M, Wijayawardena BK, Lamb MC, Kimble SJA, Ji YZ, Fernandez NB, Antonides JD, Marra NJ, Dewoody JA (2017) Biome and migratory behaviour significantly influence vertebrate genetic diversity. *Biological Journal of the Linnean Society*, 121(2), 446-457.

Younger JL, Clucas GV, Koa D, Rogers AD, Gharbi K, Hart T & Miller KJ (2015) The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Molecular Ecology*, 1-15. doi: 10.1111/mec.14172.

Zdobnov EM & Apweiler R (2001) InterProScan–an integration platform for the signature recognition methods in InterPro. *Bioinformatics*, 17, 847-848.

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C & Weir BS (2012) A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. *Bioinformatics*, doi: 10.1093/bioinformatics/bts606.

Zink RM (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.

TABLES

Table 1. Location table organized by migration strategy and an estimate of mean migration distance for each study site. Summary of the number of individuals used to generate RAD-sequencing data (NRAD_nofilter), and the number of samples retained in the analyses of population genetic structure after SNP filtering based on the amount of missing data per individual and per locus (NRAD_filtered). For each of the 12 study sites, the central longitudinal and latitudinal coordinates, migration strategy and mean migration distance estimated from the bird banding laboratory band-recovery database is listed. Study sites are organized by Mean Migration Distance (km) and shaded according to their Migration Strategy.

Study Site	Latitude	Longitude	Migration Strategy	Mean Migration Distance (km)	$N_{RAD\_nofilter}$	$N_{RAD\_filtered}$
AK1	64.12286	-145.644	Long-distance	2164	25	18
SK1	55.04513	-106.017	Long-distance	2085	40	28
OR1	45.53687	-123.043	Short-distance	1082	14	11
WI1	43.70895	-89.3888	Short-distance	654.2	30	18
PA1	40.58988	-75.7492	Short-distance	331.5	37	23
ID1	43.50649	-116.386	Short-distance	259.4	33	29
CA1	37.22091	-119.917	Short-distance	129.6	26	10
VA1	38.64525	-78.7657	Short-distance	75.89	38	29
ID2	43.28057	-114.999	Short-distance	-	9	2
TX1	33.59943	-101.936	Nonmigratory	37.24	13	12
FL1	29.2975	-82.3363	Nonmigratory	18.92	18	16
CA2	33.78779	-117.853	Nonmigratory	-	4	1
Total	-	-	-	-	287	197

Table 2. Pairwise estimates of genetic differentiation (FST) between all sampling locations with greater than five individuals. The weighted mean value of pairwise FST using 72,263 SNPs with the R package SNPRelate. We calculated pairwise FST between 10 sampling locations; two study sites had fewer than five individuals and were removed from our analysis. Sampling locations correspond to those listed in Table 1 and are organized by migration strategy. Color coding of values indicates the degree of differentiation between locations with greatest values of genetic divergence between nonmigratory populations in Florida and Texas. These results are consistent with our prediction that nonmigratory populations are more genetically differentiated compared to migrant populations.

		Long-distance		Short-distance						Nonmigratory
		AK1	SK1	OR1	ID1	CA1	WI1	PA1	VA1	TX1
Long-distance	SK1	0.00398	-	-	-	-	-	-	-	-
Short-distance	OR1	0.00535	0.00345	-	-	-	-	-	-	-
	ID1	0.00554	0.00238	0.00387	-	-	-	-	-	-
	CA1	0.00665	0.00301	0.00535	0.00325	-	-	-	-	-
	WI1	0.00701	0.00257	0.00630	0.00494	0.00559	-	-	-	-
	PA1	0.00713	0.00330	0.00664	0.00585	0.00678	0.00205	-	-	-
	VA1	0.00945	0.00533	0.00883	0.00788	0.00850	0.00390	0.00356	-	-
Nonmigratory	TX1	0.01079	0.00900	0.00954	0.00993	0.01187	0.01136	0.01164	0.01394	-
	FL1	0.01359	0.01133	0.01302	0.01346	0.01473	0.00947	0.00877	0.01040	0.01770

Key
$F_{ST} < 0.005$
$0.005 < F_{ST} < 0.01$
$0.01 < F_{ST} < 0.015$
$F_{ST} < 0.015$

FIGURES

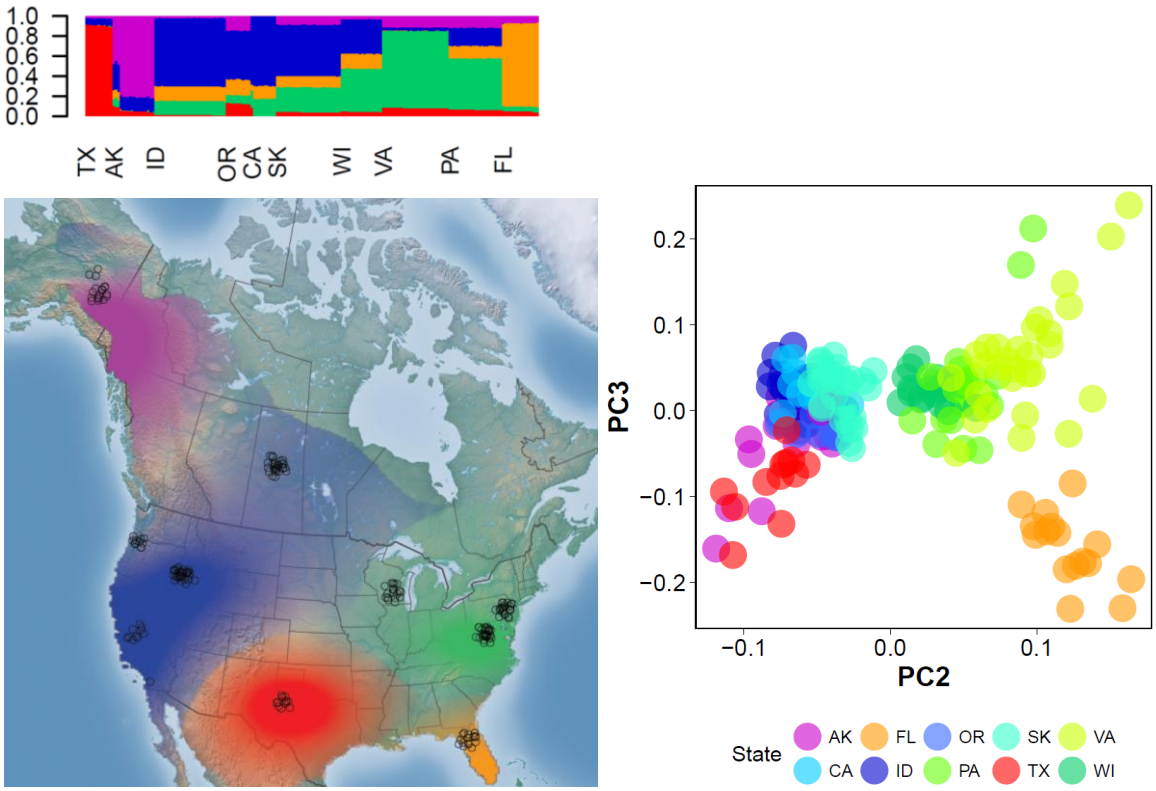

Figure 1. Tess map of interpolated ancestry across the breeding range with PCA using the complete data set. (A) Results from TESS3R demonstrate an isolation by distance pattern of genetic structure with Florida and Texas populations exhibiting the greatest assignment probability to distinct clusters. Each bar represents an individual, and individuals are organized by geographic location. (B) Results from PCA suggest kestrels can be separated into three genetically distinct populations with separation of Florida from the eastern and western sampling locations. The PCA also suggests separation of Alaska and Texas from the western sampling locations. These results indicate that nonmigratory populations of kestrels are more genetically structured compared to migratory populations.

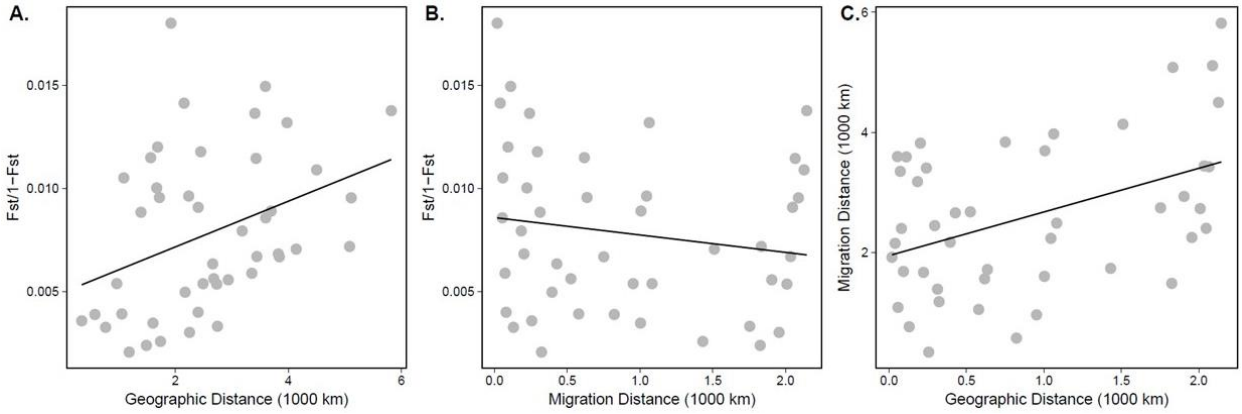

Figure 2. Relationship between genetic, geographic and migration distance.Multiple regression on distance matrices using all populations with greater than five individuals indicate that geographic and migration distance influences genetic distance. The direction of the effect is consistent with our prediction that populations with the shortest mean migration distance are the most genetically differentiated from one another.

SUPPORTING INFORMATION

Table SI 1. Individual information organized by sampling location. We used RAD-sequencing to obtain genome-wide sequence data for 287 kestrels that were sampled across 12 sampling locations. Of the 287 individuals presented in this table, only 197 samples were retained in the analyses of population genetic structure after filtering based sequencing coverage (RAD_filtered). This table provides a summary of sampling location (Latitude, Longitude, Study_site, Location), collection date, collector, sample identification (Sample_ID), and RAD-sequencing information for each individual included in our analysis. Most individuals were banded using rings from the BBL (Band_number; USGS Bird Banding Laboratory), and we assigned a unique identifier to individuals that did not have BBL band numbers (Sample_ID).

Sample_ID	Band_number	Latitude	Longitude	Study_site	Location	Sample_type	Collector	Collection_date	RAD_filtered
1523-92422	1523-92422	63.957	-145.421	AK	Delta Junction, AK	Blood	Ted Swem	12 July, 2016	Yes
1523-92430	1523-92430	63.990	-145.369	AK	Delta Junction, AK	Blood	Ted Swem	12 July, 2016	Yes
1523-92433	1523-92433	63.955	-145.075	AK	Delta Junction, AK	Blood	Ted Swem	12 July, 2016	Yes
1523-92439	1523-92439	63.948	-145.107	AK	Delta Junction, AK	Blood	Ted Swem	12 July, 2016	Yes
1593-31251	1593-31251	63.980	-145.348	AK	Delta Junction, AK	Blood	Ted Swem	4 July, 2015	Yes
1593-31256	1593-31256	63.938	-145.252	AK	Delta Junction, AK	Blood	Ted Swem	4 July, 2015	Yes
1593-31261	1593-31261	64.016	-145.125	AK	Delta Junction, AK	Blood	Ted Swem	4 July, 2015	Yes
1593-31266	1593-31266	63.957	-145.421	AK	Delta Junction, AK	Blood	Ted Swem	4 July, 2015	Yes
1593-31271	1593-31271	63.939	-145.127	AK	Delta Junction, AK	Blood	Ted Swem	7 July, 2015	Yes
1593-31274	1593-31274	63.955	-145.092	AK	Delta Junction, AK	Blood	Ted Swem	7 July, 2015	Yes
1593-31277	1593-31277	63.985	-145.127	AK	Delta Junction, AK	Blood	Ted Swem	7 July, 2015	Yes
1593-31281	1593-31281	63.993	-145.127	AK	Delta Junction, AK	Blood	Ted Swem	7 July, 2015	Yes
1593-31282	1593-31282	63.922	-145.143	AK	Delta Junction, AK	Blood	Ted Swem	26 July, 2015	Yes
1593-31288	1593-31288	63.980	-145.127	AK	Delta Junction, AK	Blood	Ted Swem	26 July, 2015	Yes
17N02286	1523-92462`	63.961	-145.090	AK	Delta Junction, AK	Blood	Ted Swem	9 July, 2017	-
17N02287	1523-92456	63.928	-145.151	AK	Delta Junction, AK	Blood	Ted Swem	9 July, 2017	-
17N02288	1523-92451	63.939	-145.235	AK	Delta Junction, AK	Blood	Ted Swem	9 July, 2017	-
17N02289	1523-92465	63.980	-145.127	AK	Delta Junction, AK	Blood	Ted Swem	19 July, 2017	-
17N02290	1523-92445	64.019	-145.136	AK	Delta Junction, AK	Blood	Ted Swem	29 June, 2017	-
17N02284	1523-92450	63.947	-145.273	AK	Delta Junction, AK	Feather	Ted Swem	9 July, 2017	Yes
16N0980	1523-92409	63.948	-145.107	AK	Delta Junction, AK	Feather	Ted Swem	4 July, 2016	-

16N0981	1523-92413	63.948	-145.107	AK	Delta Junction, AK	Feather	Ted Swem	4 July, 2016	-
1593-31242	1593-31242	64.902	-147.804	AK	Fairbanks, AK	Blood	Ted Swem	3 July, 2015	Yes
1593-31243	1593-31243	64.945	-147.846	AK	Fairbanks, AK	Blood	Ted Swem	3 July, 2015	Yes
17N02285	1523-92446	64.902	-147.804	AK	Fairbanks, AK	Feather	Ted Swem	18 July, 2017	Yes
17N02198	1833-45558	37.360	-119.798	CA1	Anwahnee, CA	Feather	Bill Ralph	19 May, 2017	Yes
17N02182	1833-45591	37.246	-120.156	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	Yes
17N02183	1833-45596	37.250	-120.152	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	Yes
17N02186	1823-51512	37.272	-120.149	CA1	Le Grand, CA	Feather	Bill Ralph	15 May, 2017	Yes
17N02190	1833-45524	37.228	-120.001	CA1	Le Grand, CA	Feather	Bill Ralph	15 May, 2017	Yes
17N02200	1833-45575	37.232	-119.104	CA1	Le Grand, CA	Feather	Bill Ralph	29 May, 2017	Yes
17N02194	1833-45538	37.048	-119.952	CA1	Madera, CA	Feather	Bill Ralph	18 May, 2017	Yes
17N02197	1833-45533	37.228	-120.001	CA1	Madera, CA	Feather	Bill Ralph	18 May, 2017	Yes
17N02192	1833-45544	37.104	-119.876	CA1	Raymond, CA	Feather	Bill Ralph	17 May, 2017	Yes
17N02195	1823-51520	37.242	-119.983	CA1	Raymond, CA	Feather	Bill Ralph	18 May, 2017	Yes
17N02203	1823-51531	37.362	-119.794	CA1	Anwahnee, CA	Feather	Bill Ralph	31 May, 2017	-
17N02177	1833-45561	37.229	-120.098	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	-
17N02178	1833-45563	37.228	-120.085	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	-
17N02179	1833-45577	37.235	-120.106	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	-
17N02180	1833-45581	37.236	-120.113	CA1	Le Grand, CA	Feather	Bill Ralph	14 May, 2017	-
17N02185	1823-51509	37.254	-120.131	CA1	Le Grand, CA	Feather	Bill Ralph	15 May, 2017	-
17N02187	1823-51514	37.265	-120.152	CA1	Le Grand, CA	Feather	Bill Ralph	15 May, 2017	-
17N02199	1833-45568	37.230	-120.106	CA1	Le Grand, CA	Feather	Bill Ralph	29 May, 2017	-
17N02168	1823-51504	37.051	-120.023	CA1	Madera, CA	Feather	Bill Ralph	3 May, 2017	-
17N02169	1823-51505	37.069	-119.976	CA1	Madera, CA	Feather	Bill Ralph	3 May, 2017	-
17N02173	1833-45551	36.858	-119.998	CA1	Madera, CA	Feather	Bill Ralph	5 May, 2017	-
17N02154	1857-45763	37.106	-119.881	CA1	Raymond, CA	Feather	Bill Ralph	19 April, 2017	-
17N02184	1803-28501	37.217	-120.004	CA1	Raymond, CA	Feather	Bill Ralph	15 May, 2017	-
17N02193	1833-45545	37.106	-119.881	CA1	Raymond, CA	Feather	Bill Ralph	17 May, 2017	-
17N02207	1803-28518	37.216	-119.993	CA1	Raymond, CA	Feather	Bill Ralph	13 June, 2017	-
17N02208	1823-51534	37.206	-119.999	CA1	Raymond, CA	Feather	Bill Ralph	13 June, 2017	-

OCBPC-2	-	33.788	-117.853	CA2	Orange County, CA	Blood	Joel Pagel & Scott Thomas	-	Yes
OCBPC-3	-	33.788	-117.853	CA2	Orange County, CA	Blood	Joel Pagel & Scott Thomas	-	-
OCBPC-4	-	33.788	-117.853	CA2	Orange County, CA	Blood	Joel Pagel & Scott Thomas	-	-
16N1025	-	33.619	-117.929	CA2	Newport Beach, CA	Feather	Krysta Rogers	21 June, 2016	-
1783-83757	1783-83757	29.285	-82.173	FL	Anthony, FL	Blood	Karl Miller	15 May, 2015	Yes
1783-83769	1783-83769	29.272	-82.132	FL	Anthony, FL	Blood	Karl Miller	19 May, 2015	Yes
1783-83760	1783-83760	29.305	-82.547	FL	Bronson, FL	Blood	Karl Miller	18 May, 2015	Yes
1783-83745	1783-83745	29.330	-82.225	FL	Lowell, FL	Blood	Karl Miller	8 May, 2015	Yes
1783-83756	1783-83756	29.346	-82.198	FL	Lowell, FL	Blood	Karl Miller	15 May, 2015	Yes
1783-83770	1783-83770	29.296	-82.170	FL	Ocala, FL	Blood	Karl Miller	21 May, 2015	Yes
1783-83750	1783-83750	29.367	-82.203	FL	Reddick, FL	Blood	Karl Miller	11 May, 2015	Yes
1783-83764	1783-83764	29.389	-82.175	FL	Reddick, FL	Blood	Karl Miller	19 May, 2015	Yes
1783-83768	1783-83768	29.376	-82.175	FL	Reddick, FL	Blood	Karl Miller	19 May, 2015	Yes
1783-83772	1783-83772	29.367	-82.122	FL	Sparr, FL	Blood	Karl Miller	21 May, 2015	Yes
1783-83761	1783-83761	29.291	-82.558	FL	Williston, FL	Blood	Karl Miller	18 May, 2015	Yes
1783-83762	1783-83762	29.234	-82.539	FL	Williston, FL	Blood	Karl Miller	18 May, 2015	Yes
1783-83763	1783-83763	29.225	-82.536	FL	Williston, FL	Blood	Karl Miller	18 May, 2015	Yes
1783-83775	1783-83775	29.218	-82.555	FL	Williston, FL	Blood	Karl Miller	4 June, 2015	Yes
1783-83776	1783-83776	29.215	-82.557	FL	Williston, FL	Blood	Karl Miller	4 June, 2015	Yes
1783-83777	1783-83777	29.244	-82.516	FL	Williston, FL	Blood	Karl Miller	4 June, 2015	Yes
1783-83741	1783-83741	29.387	-82.187	FL	Reddick, FL	Blood	Karl Miller	7 May, 2015	-
1783-83753	1783-83753	29.357	-82.163	FL	Reddick, FL	Blood	Karl Miller	11 May, 2015	-
1523-75227	1523-75227	43.522	-116.315	ID1	Boise, ID	Blood	Julie Heath	9 May, 2014	Yes
1593-10594	1593-10594	43.486	-116.254	ID1	Boise, ID	Blood	Julie Heath	22 June, 2016	Yes
1593-10595	1593-10595	43.445	-116.353	ID1	Boise, ID	Blood	Julie Heath	22 June, 2016	Yes
1783-54602	1783-54602	43.525	-116.526	ID1	Boise, ID	Blood	Julie Heath	29 April, 2014	Yes
1783-54641	1783-54641	43.493	-116.274	ID1	Boise, ID	Blood	Julie Heath	29 April, 2014	Yes

1783-54677	1783-54677	43.517	-116.339	ID1	Boise, ID	Blood	Julie Heath	6 May, 2014	Yes
1783-54817	1783-54817	43.453	-116.137	ID1	Boise, ID	Blood	Julie Heath	2 May, 2014	Yes
1783-54818	1783-54818	43.474	-116.274	ID1	Boise, ID	Blood	Julie Heath	5 May, 2014	Yes
1783-54819	1783-54819	43.522	-116.315	ID1	Boise, ID	Blood	Julie Heath	5 May, 2014	Yes
1783-54820	1783-54820	43.458	-116.214	ID1	Boise, ID	Blood	Julie Heath	16 May, 2014	Yes
1783-54821	1783-54821	43.490	-116.264	ID1	Boise, ID	Blood	Julie Heath	19 May, 2014	Yes
1783-54880	1783-54880	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	18 April, 2014	Yes
1783-54881	1783-54881	43.548	-116.488	ID1	Boise, ID	Blood	Julie Heath	21 April, 2014	Yes
1783-54883	1783-54883	43.550	-116.474	ID1	Boise, ID	Blood	Julie Heath	6 May, 2014	Yes
1783-54884	1783-54884	43.539	-116.454	ID1	Boise, ID	Blood	Julie Heath	6 May, 2014	Yes
1783-54886	1783-54886	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	6 May, 2014	Yes
1783-54887	1783-54887	43.512	-116.523	ID1	Boise, ID	Blood	Julie Heath	6 May, 2014	Yes
1783-54889	1783-54889	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	9 May, 2014	Yes
1783-54892	1783-54892	43.503	-116.476	ID1	Boise, ID	Blood	Julie Heath	19 May, 2014	Yes
1783-54894	1783-54894	43.517	-116.520	ID1	Boise, ID	Blood	Julie Heath	20 May, 2014	Yes
1783-80516	1783-80516	43.444	-116.245	ID1	Boise, ID	Blood	Julie Heath	29 April, 2014	Yes
1783-80522	1783-80522	43.487	-116.510	ID1	Boise, ID	Blood	Julie Heath	23 April, 2014	Yes
1783-80529	1783-80529	43.445	-116.353	ID1	Boise, ID	Blood	Julie Heath	5 May, 2014	Yes
1783-80539	1783-80539	43.444	-116.245	ID1	Boise, ID	Blood	Julie Heath	16 May, 2014	Yes
1783-80599	1783-80599	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	18 April, 2014	Yes
1593-03535	1593-03535	43.483	-116.454	ID1	Kuna, ID	Blood	Julie Heath	3 May, 2016	Yes
1593-10600	1593-10600	43.474	-116.368	ID1	Kuna, ID	Blood	Julie Heath	22 June, 2016	Yes
1593-90204	1593-90204	43.563	-116.484	ID1	Meridian, ID	Blood	Julie Heath	22 June, 2016	Yes
1593-90205	1593-90205	43.539	-116.454	ID1	Meridian, ID	Blood	Julie Heath	29 June, 2016	Yes
1783-54780	1783-54780	43.464	-116.341	ID1	Boise, ID	Blood	Julie Heath	29 April, 2014	-
1783-54890	1783-54890	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	16 May, 2014	-
1783-54891	1783-54891	43.564	-116.474	ID1	Boise, ID	Blood	Julie Heath	18 May, 2014	-
1593-03524	1593-03524	43.503	-116.522	ID1	Meridian, ID	Blood	Julie Heath	22 April, 2016	-
1783-22363	1783-22363	43.256	-115.030	ID2	Fairfield, ID	Blood	Julie Heath	-	Yes
1783-22397	1783-22397	43.305	-114.969	ID2	Fairfield, ID	Blood	Julie Heath	-	Yes
1783-22365	1783-22365	43.256	-115.044	ID2	Fairfield, ID	Blood	Julie Heath	-	-
1783-22376	1783-22376	43.259	-115.047	ID2	Fairfield, ID	Blood	Julie Heath	-	-

1783-22400	1783-22400	43.290	-114.969	ID2	Fairfield, ID	Blood	Julie Heath	-	-
1783-65910	1783-65910	43.293	-114.969	ID2	Fairfield, ID	Blood	Julie Heath	13 July, 2016	-
1783-65912	1783-65912	43.303	-114.969	ID2	Fairfield, ID	Blood	Julie Heath	13 July, 2016	-
1783-65914	1783-65914	43.256	-115.033	ID2	Fairfield, ID	Blood	Julie Heath	13 July, 2016	-
1783-65916	1783-65916	43.256	-115.033	ID2	Fairfield, ID	Blood	Julie Heath	13 July, 2016	-
1783-97718	1783-97718	45.485	-123.052	OR	Cornelius, OR	Blood	Rich Van Buskirk	28 July, 2015	Yes
1783-97713	1783-97713	45.564	-123.084	OR	Forest Grove, OR	Blood	Rich Van Buskirk	14 July, 2015	Yes
1783-97714	1783-97714	45.556	-123.073	OR	Forest Grove, OR	Blood	Rich Van Buskirk	15 July, 2015	Yes
1783-97715	1783-97715	45.466	-123.149	OR	Forest Grove, OR	Blood	Rich Van Buskirk	17 July, 2015	Yes
1783-97719	1783-97719	45.564	-123.093	OR	Forest Grove, OR	Blood	Rich Van Buskirk	6 August, 2015	Yes
1783-97721	1783-97721	45.586	-123.087	OR	Forest Grove, OR	Blood	Rich Van Buskirk	14 August, 2015	Yes
1783-97722	1783-97722	45.555	-123.074	OR	Forest Grove, OR	Blood	Rich Van Buskirk	14 August, 2015	Yes
1783-97710	1783-97710	45.405	-123.103	OR	Gaston, OR	Blood	Rich Van Buskirk	20 May, 2015	Yes
1623-57120	1623-57120	45.557	-122.893	OR	Hillsboro, OR	Blood	Rich Van Buskirk	18 June, 2015	Yes
1783-97716	1783-97716	45.577	-122.938	OR	Hillsboro, OR	Blood	Rich Van Buskirk	25 July, 2015	Yes
1783-97717	1783-97717	45.590	-122.932	OR	Hillsboro, OR	Blood	Rich Van Buskirk	25 July, 2015	Yes
1783-97720	1783-97720	45.536	-123.103	OR	Forest Grove, OR	Blood	Rich Van Buskirk	13 August, 2015	-
1783-97712	1783-97712	45.587	-123.017	OR	North Plains, OR	Blood	Rich Van Buskirk	3 June, 2015	-
1783-97723	1783-97723	45.598	-123.024	OR	North Plains, OR	Blood	Rich Van Buskirk	31 August, 2015	-
1703-98047	1703-98047	40.608	-75.836	PA	Albany, PA	Blood	Robertson	11 June, 2015	Yes
1703-98089	1703-98089	40.516	-76.051	PA	Hamburg, PA	Blood	Robertson	18 June, 2015	Yes
1703-98093	1703-98093	40.526	-76.066	PA	Hamburg, PA	Blood	Robertson	18 June, 2015	Yes
1703-98062	1703-98062	40.626	-75.895	PA	Kempton, PA	Blood	Robertson	17 June, 2015	Yes
1703-98070	1703-98070	40.623	-75.869	PA	Kempton, PA	Blood	Robertson	17 June, 2015	Yes

1703-98074	1703-98074	40.632	-75.812	PA	Kempton, PA	Blood	Robertson	17 June, 2015	Yes
1703-98078	1703-98078	40.638	-75.794	PA	Kempton, PA	Blood	Robertson	17 June, 2015	Yes
1703-98083	1703-98083	40.589	-75.842	PA	Klinesville, PA	Blood	Robertson	17 June, 2015	Yes
1703-98088	1703-98088	40.550	-75.731	PA	Kutztown, PA	Blood	Robertson	18 June, 2015	Yes
1703-98044	1703-98044	40.679	-75.811	PA	Lynnport, PA	Blood	Robertson	8 June, 2015	Yes
1703-98046	1703-98046	40.679	-75.811	PA	Lynnport, PA	Blood	Robertson	8 June, 2015	Yes
1703-98056	1703-98056	40.723	-76.021	PA	New Ringgold, PA	Blood	Robertson	15 June, 2015	Yes
1703-98055	1703-98055	40.705	-75.779	PA	New Tripoli, PA	Blood	Robertson	11 June, 2015	Yes
1703-98058	1703-98058	40.651	-76.073	PA	Orwigsburg, PA	Blood	Robertson	17 June, 2015	Yes
1703-98095	1703-98095	40.620	-75.896	PA	Albany, PA	Blood	Robertson	19 June, 2015	-
1783-98201	1783-98201	40.606	-75.853	PA	Albany, PA	Blood	Robertson	19 June, 2015	-
1703-98073	1703-98073	40.656	-75.860	PA	Steinsville, PA	Blood	Robertson	17 June, 2015	-
16N0988	1143-38019	40.628	-75.853	PA	Kempton, PA	Feather	Jean-Francois Therrien	22 June, 2016	Yes
16N0990	1143-36204	40.628	-75.853	PA	Kempton, PA	Feather	Jean-Francois Therrien	30 June, 2016	Yes
16N0554	1833-17553	40.740	-75.310	PA	Nazareth, PA	Feather	Tim Kita	13 June, 2016	Yes
16N0582	1833-17586	40.686	-75.497	PA	Northampton, PA	Feather	Tim Kita	22 June, 2016	Yes
16N0586	1833-17611	40.635	-75.585	PA	Orefield, PA	Feather	Tim Kita	21 June, 2016	Yes
16N0588	1833-17619	40.635	-75.585	PA	Orefield, PA	Feather	Tim Kita	21 June, 2016	Yes
16N0593	1833-17633	40.336	-75.927	PA	Reading, PA	Feather	Tim Kita	24 June, 2016	Yes
16N0549	1833-17493	40.267	-75.116	PA	Reinholds, PA	Feather	Tim Kita	13 June, 2016	Yes
16N0592	1833-17638	40.267	-75.116	PA	Reinholds, PA	Feather	Tim Kita	20 July, 2016	Yes
16N0986	1143-38003	40.628	-75.853	PA	Kempton, PA	Feather	Jean-Francois Therrien	14 June, 2016	-
16N0989	1143-38022	40.628	-75.853	PA	Kempton, PA	Feather	Jean-Francois Therrien	22 June, 2016	-
16N0991	1833-36207	40.628	-75.853	PA	Kempton, PA	Feather	Jean-Francois Therrien	30 June, 2016	-
16N0556	1833-17560	40.726	-75.394	PA	BATH, PA	Feather	Tim Kita	13 June, 2016	-

16N0585	1833-17575	40.788	-75.476	PA	DANIELSVILLE, PA	Feather	Tim Kita	15 June, 2016	-
16N0555	1833-17557	40.740	-75.310	PA	Nazareth, PA	Feather	Tim Kita	13 June, 2016	-
16N0587	1833-17614	40.635	-75.585	PA	Orefield, PA	Feather	Tim Kita	21 June, 2016	-
16N0589	1833-17624	40.635	-75.585	PA	Orefield, PA	Feather	Tim Kita	21 June, 2016	-
16N0547	1833-17488	40.267	-75.116	PA	Reinholds, PA	Feather	Tim Kita	13 June, 2016	-
16N0550	1833-17495	40.267	-75.116	PA	Reinholds, PA	Feather	Tim Kita	13 June, 2016	-
16N0591	1833-17629	40.675	-75.618	PA	Schnecksville, PA	Feather	Tim Kita	21 June, 2016	-
2003-40199	2003-40199	55.131	-106.031	SK	Besnard Lake, SK	Blood	Russ Dawson	7 June, 2008	Yes
2003-40469	2003-40469	55.104	-105.942	SK	Besnard Lake, SK	Blood	Russ Dawson	3 June, 2008	Yes
2003-40475	2003-40475	54.773	-105.641	SK	Besnard Lake, SK	Blood	Russ Dawson	30 May, 2008	Yes
2003-40479	2003-40479	55.165	-106.222	SK	Besnard Lake, SK	Blood	Russ Dawson	1 June, 2008	Yes
2003-40486	2003-40486	54.845	-105.703	SK	Besnard Lake, SK	Blood	Russ Dawson	30 April, 2008	Yes
2003-40497	2003-40497	54.838	-105.696	SK	Besnard Lake, SK	Blood	Russ Dawson	5 June, 2008	Yes
2003-40498	2003-40498	54.838	-105.696	SK	Besnard Lake, SK	Blood	Russ Dawson	6 June, 2008	Yes
2003-40506	2003-40506	55.155	-106.038	SK	Besnard Lake, SK	Blood	Russ Dawson	11 June, 2008	Yes
2003-40519	2003-40519	55.157	-106.353	SK	Besnard Lake, SK	Blood	Russ Dawson	10 June, 2008	Yes
2003-40525	2003-40525	55.154	-106.434	SK	Besnard Lake, SK	Blood	Russ Dawson	10 June, 2008	Yes
2003-40526	2003-40526	55.155	-106.425	SK	Besnard Lake, SK	Blood	Russ Dawson	4 June, 2008	Yes
2003-40530	2003-40530	54.996	-105.741	SK	Besnard Lake, SK	Blood	Russ Dawson	12 June, 2008	Yes
2003-40533	2003-40533	55.419	-106.068	SK	Besnard Lake, SK	Blood	Russ Dawson	28 May, 2008	Yes
2003-40535	2003-40535	55.160	-106.185	SK	Besnard Lake, SK	Blood	Russ Dawson	29 May, 2008	Yes
2003-40536	2003-40536	55.146	-106.490	SK	Besnard Lake, SK	Blood	Russ Dawson	29 May, 2008	Yes

2003-40544	2003-40544	54.749	-105.635	SK	Besnard Lake, SK	Blood	Russ Dawson	2 June, 2008	Yes
2003-40549	2003-40549	55.167	-106.287	SK	Besnard Lake, SK	Blood	Russ Dawson	4 June, 2008	Yes
2003-40550	2003-40550	55.148	-106.468	SK	Besnard Lake, SK	Blood	Russ Dawson	4 June, 2008	Yes
2003-40553	2003-40553	54.749	-105.635	SK	Besnard Lake, SK	Blood	Russ Dawson	5 June, 2008	Yes
2003-40556	2003-40556	54.986	-105.733	SK	Besnard Lake, SK	Blood	Russ Dawson	5 June, 2008	Yes
2003-40559	2003-40559	54.824	-105.682	SK	Besnard Lake, SK	Blood	Russ Dawson	6 June, 2008	Yes
2003-40565	2003-40565	55.155	-106.425	SK	Besnard Lake, SK	Blood	Russ Dawson	7 June, 2008	Yes
2003-40566	2003-40566	55.129	-106.039	SK	Besnard Lake, SK	Blood	Russ Dawson	8 June, 2008	Yes
2003-40567	2003-40567	54.855	-105.716	SK	Besnard Lake, SK	Blood	Russ Dawson	8 June, 2008	Yes
2003-40569	2003-40569	55.385	-106.091	SK	Besnard Lake, SK	Blood	Russ Dawson	10 June, 2008	Yes
2003-40575	2003-40575	54.820	-105.718	SK	Besnard Lake, SK	Blood	Russ Dawson	11 June, 2008	Yes
2003-40581	2003-40581	55.158	-106.407	SK	Besnard Lake, SK	Blood	Russ Dawson	14 June, 2008	Yes
2003-40583	2003-40583	55.104	-105.965	SK	Besnard Lake, SK	Blood	Russ Dawson	15 June, 2008	Yes
2003-40483	2003-40483	55.161	-106.145	SK	Besnard Lake, SK	Blood	Russ Dawson	29 May, 2008	-
2003-40485	2003-40485	54.855	-105.716	SK	Besnard Lake, SK	Blood	Russ Dawson	3 June, 2008	-
2003-40494	2003-40494	54.794	-105.655	SK	Besnard Lake, SK	Blood	Russ Dawson	30 May, 2008	-
2003-40503	2003-40503	55.167	-106.257	SK	Besnard Lake, SK	Blood	Russ Dawson	16 June, 2008	-
2003-40521	2003-40521	55.104	-105.965	SK	Besnard Lake, SK	Blood	Russ Dawson	14 May, 2008	-
2003-40539	2003-40539	55.344	-106.102	SK	Besnard Lake, SK	Blood	Russ Dawson	31 May, 2008	-
2003-40540	2003-40540	55.356	-106.099	SK	Besnard Lake, SK	Blood	Russ Dawson	31 May, 2008	-

2003-40543	2003-40543	54.760	-105.641	SK	Besnard Lake, SK	Blood	Russ Dawson	2 June, 2008	-
2003-40558	2003-40558	54.729	-105.607	SK	Besnard Lake, SK	Blood	Russ Dawson	6 June, 2008	-
2003-40568	2003-40568	54.767	-105.642	SK	Besnard Lake, SK	Blood	Russ Dawson	8 June, 2008	-
2003-40572	2003-40572	55.145	-106.087	SK	Besnard Lake, SK	Blood	Russ Dawson	10 June, 2008	-
2003-40584	2003-40584	55.054	-105.803	SK	Besnard Lake, SK	Blood	Russ Dawson	15 June, 2008	-
1623-32299	1623-32299	33.594	-102.038	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14406	1833-14406	33.613	-102.052	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14411	1833-14411	33.603	-102.051	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14416	1833-14416	33.594	-102.052	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14417	1833-14417	33.585	-102.052	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14422	1833-14422	33.586	-102.038	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14430	1833-14430	33.581	-101.047	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14435	1833-14435	33.632	-101.889	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14444	1833-14444	33.602	-101.903	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14455	1833-14455	33.586	-102.027	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14458	1833-14458	33.608	-102.035	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14463	1833-14463	33.610	-102.044	TX	Lubbock, TX	Blood	Clint Boal	2016	Yes
1833-14450	1833-14450	33.634	-101.887	TX	Lubbock, TX	Blood	Clint Boal	2016	-
1833-14520	1833-14520	38.558	-78.861	VA	Rockingham, VA	Blood	Jill Morrow	12 May, 2016	Yes
1833-14525	1833-14525	38.544	-78.844	VA	Rockingham, VA	Blood	Jill Morrow	12 May, 2016	Yes
1833-14540	1833-14540	38.643	-78.742	VA	Rockingham, VA	Blood	Jill Morrow	12 May, 2016	Yes
1833-14555	1833-14555	38.678	-78.766	VA	Rockingham, VA	Blood	Jill Morrow	13 May, 2016	Yes
1833-14560	1833-14560	38.605	-78.743	VA	Rockingham, VA	Blood	Jill Morrow	13 May, 2016	Yes
1833-14565	1833-14565	38.580	-78.820	VA	Rockingham, VA	Blood	Jill Morrow	16 May, 2016	Yes
1833-14581	1833-14581	38.608	-78.830	VA	Rockingham, VA	Blood	Jill Morrow	20 May, 2016	Yes
1833-14584	1833-14584	38.585	-78.738	VA	Rockingham, VA	Blood	Jill Morrow	23 May, 2016	Yes
1833-14597	1833-14597	38.578	-78.899	VA	Rockingham, VA	Blood	Jill Morrow	23 May, 2016	Yes
1833-14602	1833-14602	38.654	-78.765	VA	Rockingham, VA	Blood	Jill Morrow	23 May, 2016	Yes
1833-14606	1833-14606	38.676	-78.752	VA	Rockingham, VA	Blood	Jill Morrow	24 May, 2016	Yes

1833-14620	1833-14620	38.662	-78.791	VA	Rockingham, VA	Blood	Jill Morrow	28 May, 2016	Yes
1833-14629	1833-14629	38.678	-78.816	VA	Rockingham, VA	Blood	Jill Morrow	29 May, 2016	Yes
1833-14650	1833-14650	38.628	-78.857	VA	Rockingham, VA	Blood	Jill Morrow	4 June, 2016	Yes
1833-14666	1833-14666	38.638	-78.723	VA	Rockingham, VA	Blood	Jill Morrow	6 June, 2016	Yes
1833-14530	1833-14530	38.712	-78.580	VA	Shenandoah, VA	Blood	Jill Morrow	12 May, 2016	Yes
1833-14535	1833-14535	38.690	-78.694	VA	Shenandoah, VA	Blood	Jill Morrow	12 May, 2016	Yes
1833-14551	1833-14551	38.702	-78.740	VA	Shenandoah, VA	Blood	Jill Morrow	13 May, 2016	Yes
1833-14576	1833-14576	38.705	-78.753	VA	Shenandoah, VA	Blood	Jill Morrow	20 May, 2016	Yes
1833-14608	1833-14608	38.719	-78.783	VA	Shenandoah, VA	Blood	Jill Morrow	6 June, 2008	Yes
1833-14615	1833-14615	38.714	-78.779	VA	Shenandoah, VA	Blood	Jill Morrow	28 May, 2016	Yes
1833-14640	1833-14640	38.703	-78.686	VA	Shenandoah, VA	Blood	Jill Morrow	30 May, 2016	Yes
1833-14661	1833-14661	38.713	-78.790	VA	Shenandoah, VA	Blood	Jill Morrow	6 June, 2016	Yes
1833-14570	1833-14570	38.633	-78.741	VA	Rockingham, VA	Blood	Jill Morrow	20 May, 2016	-
1833-14592	1833-14592	38.574	-78.886	VA	Rockingham, VA	Blood	Jill Morrow	23 May, 2016	-
1833-14625	1833-14625	38.662	-78.759	VA	Rockingham, VA	Blood	Jill Morrow	29 May, 2016	-
1833-14637	1833-14637	38.595	-78.800	VA	Rockingham, VA	Blood	Jill Morrow	29 May, 2016	-
1833-14645	1833-14645	38.582	-78.831	VA	Rockingham, VA	Blood	Jill Morrow	4 June, 2016	-
1833-14546	1833-14546	38.686	-78.650	VA	Shenandoah, VA	Blood	Jill Morrow	13 May, 2016	-
1833-14575	1833-14575	38.706	-78.701	VA	Shenandoah, VA	Blood	Jill Morrow	20 May, 2016	-
17N02003	1833-14816	38.651	-78.740	VA	Rockingham, VA	Feather	Jill Morrow	14 May, 2017	Yes
17N02004	1833-14824	38.610	-78.706	VA	Rockingham, VA	Feather	Jill Morrow	17 May, 2017	Yes
17N02008	1833-14843	38.612	-78.768	VA	Rockingham, VA	Feather	Jill Morrow	19 May, 2017	Yes
17N02009	1833-14844	38.595	-78.807	VA	Rockingham, VA	Feather	Jill Morrow	19 May, 2017	Yes
17N02010	1833-14852	38.558	-78.856	VA	Rockingham, VA	Feather	Jill Morrow	19 May, 2017	Yes
17N02007	1833-14838	38.712	-78.580	VA	Shenandoah, VA	Feather	Jill Morrow	17 May, 2017	Yes
17N02001	1833-14806	38.568	-78.774	VA	Rockingham, VA	Feather	Jill Morrow	14 May, 2017	-
17N02006	1833-14829	38.702	-78.740	VA	Shenandoah, VA	Feather	Jill Morrow	17 May, 2017	-
16N1825	1833-34704	43.192	-89.223	WI	Plainfield, WI	Feather	Janet Eschenbauch	12 June, 2016	Yes
16N1827	1833-34719	43.192	-89.232	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1831	1833-34748	43.255	-89.252	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes

16N1832	1833-34752	43.293	-89.171	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1833	1833-34763	43.295	-89.154	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1836	1833-34778	43.252	-89.111	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1838	1833-34785	43.252	-89.194	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1839	1833-35758	43.291	-89.160	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	Yes
16N1849	1833-34729	43.166	-89.215	WI	Plainfield, WI	Feather	Janet Eschenbach	24 June, 2016	Yes
16N1850	1833-34839	43.193	-89.211	WI	Plainfield, WI	Feather	Janet Eschenbach	24 June, 2016	Yes
16N1826	1833-34709	43.193	-89.221	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1828	1833-34725	43.172	-89.254	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1829	1833-34735	43.196	-89.212	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1830	1833-34743	43.250	-89.245	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1834	1833-34768	43.300	-89.120	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1835	1833-34773	43.305	-89.111	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1837	1833-34780	43.254	-89.165	WI	Plainfield, WI	Feather	Janet Eschenbach	12 June, 2016	-
16N1824	1783-84383	44.305	-89.645	WI	Plainfield, WI	Feather	Janet Eschenbach	9 June, 2016	Yes
16N1840	1783-84389	44.324	-89.647	WI	Plainfield, WI	Feather	Janet Eschenbach	17 June, 2016	Yes
16N1841	1833-34790	44.198	-89.638	WI	Plainfield, WI	Feather	Janet Eschenbach	17 June, 2016	Yes
16N1843	1833-34813	44.367	-89.589	WI	Plainfield, WI	Feather	Janet Eschenbach	18 June, 2016	Yes
16N1845	1833-34824	44.345	-89.607	WI	Plainfield, WI	Feather	Janet Eschenbach	19 June, 2016	Yes
16N1846	1833-34829	44.352	-89.626	WI	Plainfield, WI	Feather	Janet Eschenbach	19 June, 2016	Yes

16N1853	1833-34857	44.259	-89.647	WI	Plainfield, WI	Feather	Janet Eschenbach	25 June, 2016	Yes
16N1854	1833-34858	44.232	-89.678	WI	Plainfield, WI	Feather	Janet Eschenbach	25 June, 2016	Yes
16N1842	1833-34800	44.288	-89.577	WI	Plainfield, WI	Feather	Janet Eschenbach	17 June, 2016	-
16N1844	1833-34819	44.290	-89.667	WI	Plainfield, WI	Feather	Janet Eschenbach	19 June, 2016	-
16N1847	1833-34830	44.352	-89.626	WI	Plainfield, WI	Feather	Janet Eschenbach	19 June, 2016	-
16N1848	1833-34803	44.381	-89.594	WI	Plainfield, WI	Feather	Janet Eschenbach	23 June, 2016	-
16N1855	1833-34860	44.317	-89.569	WI	Plainfield, WI	Feather	Janet Eschenbach	25 June, 2016	-

Table SI 2. Pairwise estimates of genetic differentiation (FST) between 10 sampling locations with an upper and lower confidence interval. We calculated the weighted mean value of pairwise FST using 72,263 SNPs with the R package Assigner. Sampling locations correspond to those listed in Table 1, excluding two populations with fewer than five individuals. All of the values in this table are reported in 10-1. The top half of the matrix reports the pairwise FST values, and the bottom half represents the bootstrapped upper and lower 95% confidence interval for each estimate of FST calculated with the fst_WC84() function in the R package assigner using 100 iterations. These results are complimentary to the FST values that we used to perform our analyses of genetic distance correlations. The confidence intervals reported here demonstrate that overall levels of genetic differentiation between sites are low, yet significantly different from zero.

	AK	CA1	FL	ID1	OR	PA	SK	TX	VA	WI
AK		0.053	0.125	0.048	0.041	0.062	0.031	0.096	0.086	0.060
CA1	0.047 - 0.058		0.128	0.019	0.031	0.051	0.014	0.099	0.070	0.039
FL	0.121 - 0.129	0.121 - 0.134		0.125	0.112	0.076	0.103	0.162	0.093	0.082
ID1	0.045 - 0.051	0.014 - 0.023	0.121 - 0.130		0.025	0.051	0.016	0.088	0.072	0.040
OR	0.035 - 0.046	0.026 - 0.037	0.106 - 0.118	0.020 - 0.029		0.051	0.019	0.077	0.073	0.046
PA	0.058 - 0.066	0.047 - 0.057	0.072 - 0.081	0.048 - 0.054	0.046 - 0.056		0.024	0.103	0.027	0.010
SK	0.028 - 0.033	0.010 - 0.018	0.098 - 0.106	0.014 - 0.019	0.015 - 0.023	0.022 - 0.027		0.078	0.046	0.015
TX	0.093 - 0.102	0.090 - 0.105	0.156 - 0.167	0.084 - 0.093	0.069 - 0.084	0.098 - 0.108	0.073 - 0.084		0.127	0.100
VA	0.081 - 0.090	0.065 - 0.074	0.089 - 0.097	0.068 - 0.074	0.069 - 0.078	0.024 - 0.030	0.043 - 0.048	0.123 - 0.132		0.029
WI	0.055 - 0.064	0.031 - 0.044	0.077 - 0.087	0.038 - 0.044	0.042 - 0.052	0.007 - 0.013	0.012 - 0.019	0.094 - 0.105	0.025 - 0.032	

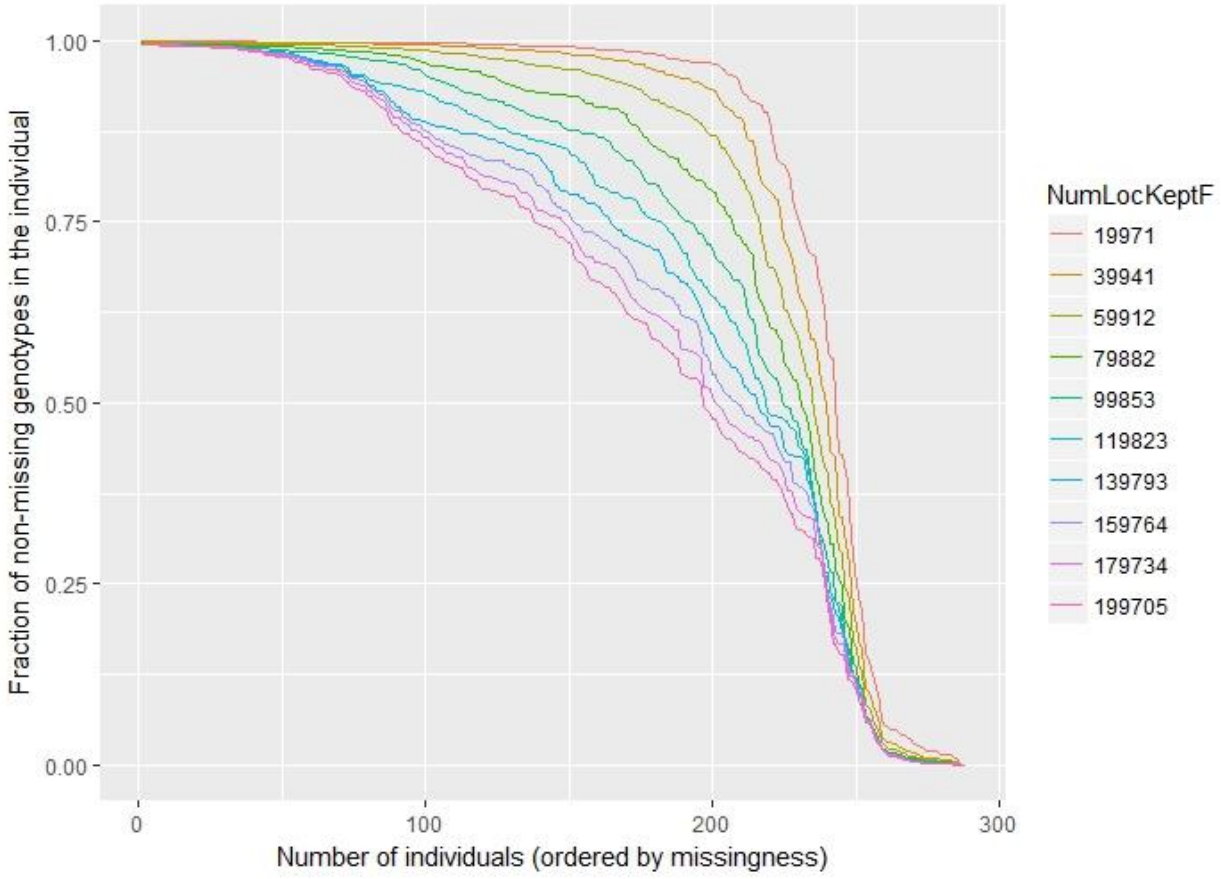

Figure SI 1. Missingness data per individual and per locus. We determined the final number of SNPs and individuals to be retained in analyses of population genetic structure by visualizing the tradeoff between discarding low coverage SNPs (y-axis) and discarding individuals with missing genotypes (x-axis). Based on this analysis, we retained 75,000 SNPs and 204 individuals for further filtering using percent heterozygosity per individual and per locus.

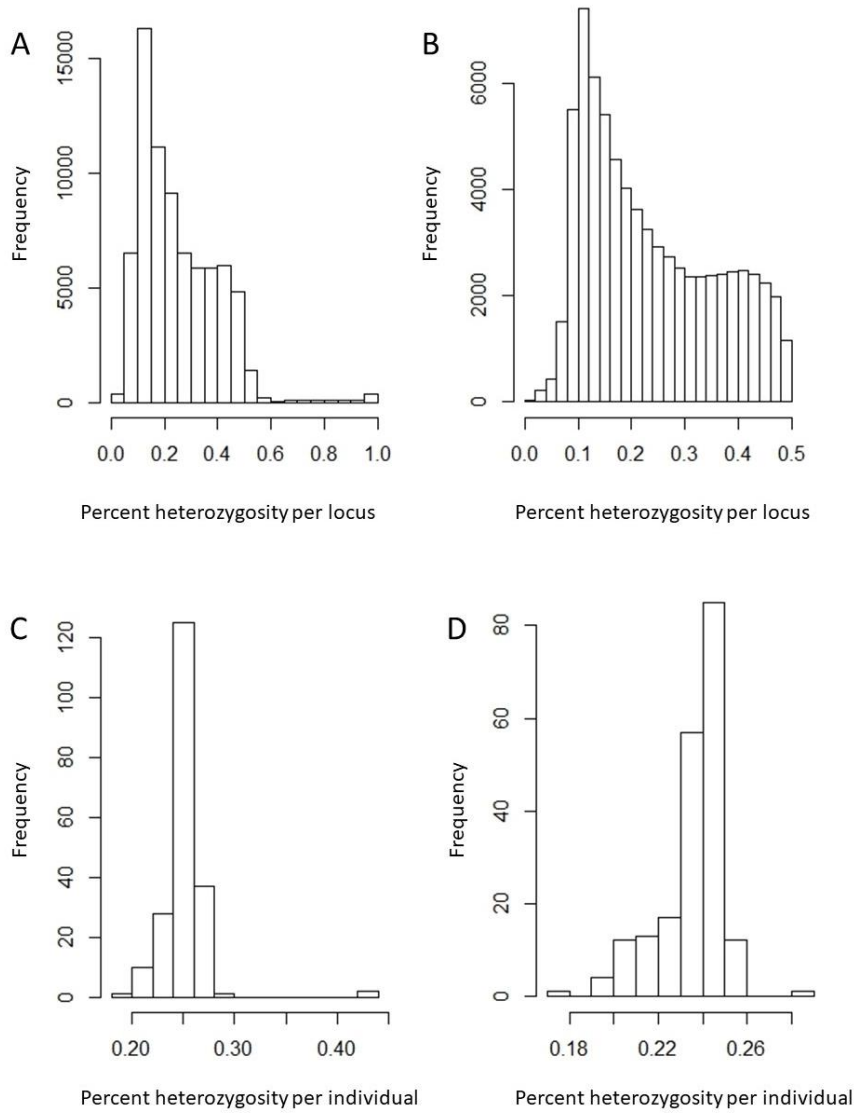

Figure SI 2. Percent heterozygosity per individual and per locus. We discarded individuals with >30% heterozygosity and loci with >50% heterozygosity. We visualized percent heterozygosity per locus and per individual before filtering (Panels A and C) and after filtering (Panels B and D). We retained 72,263 SNPs and 197 individuals in our analyses of population genetic structure.

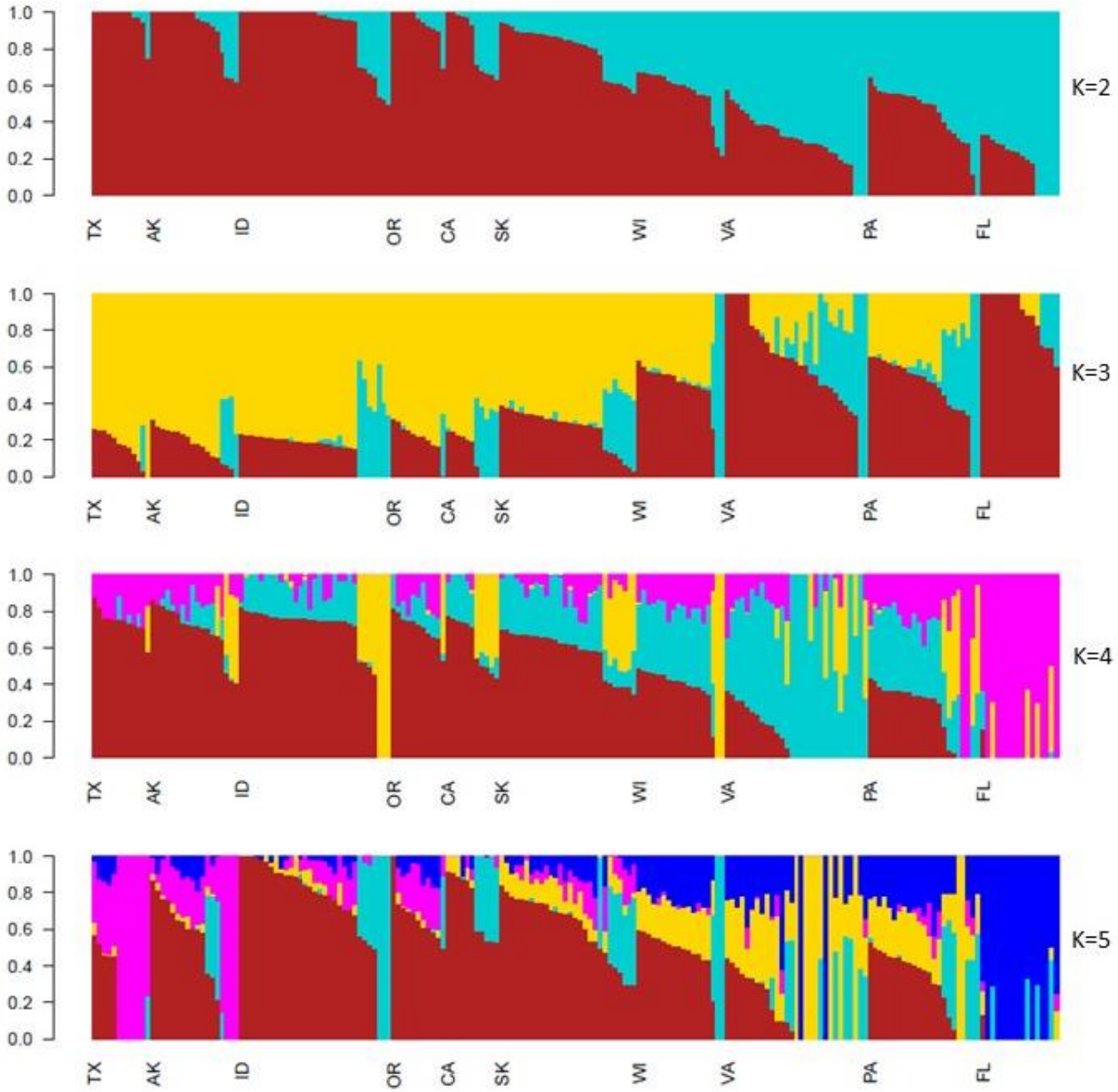

Figure SI 3. Ancestry plots using results from Admixture for 197 individuals sequenced at 72,263 SNPs.

Results from Admixture demonstrate an isolation by distance pattern of genetic structure with Florida and Texas populations exhibiting the greatest assignment probability to distinct clusters; however, we see the most support for K=2. Each panel represents results from ADMIXTURE assuming a different K value. Each bar represents an individual, and individuals are organized by geographic regions (states and one province).

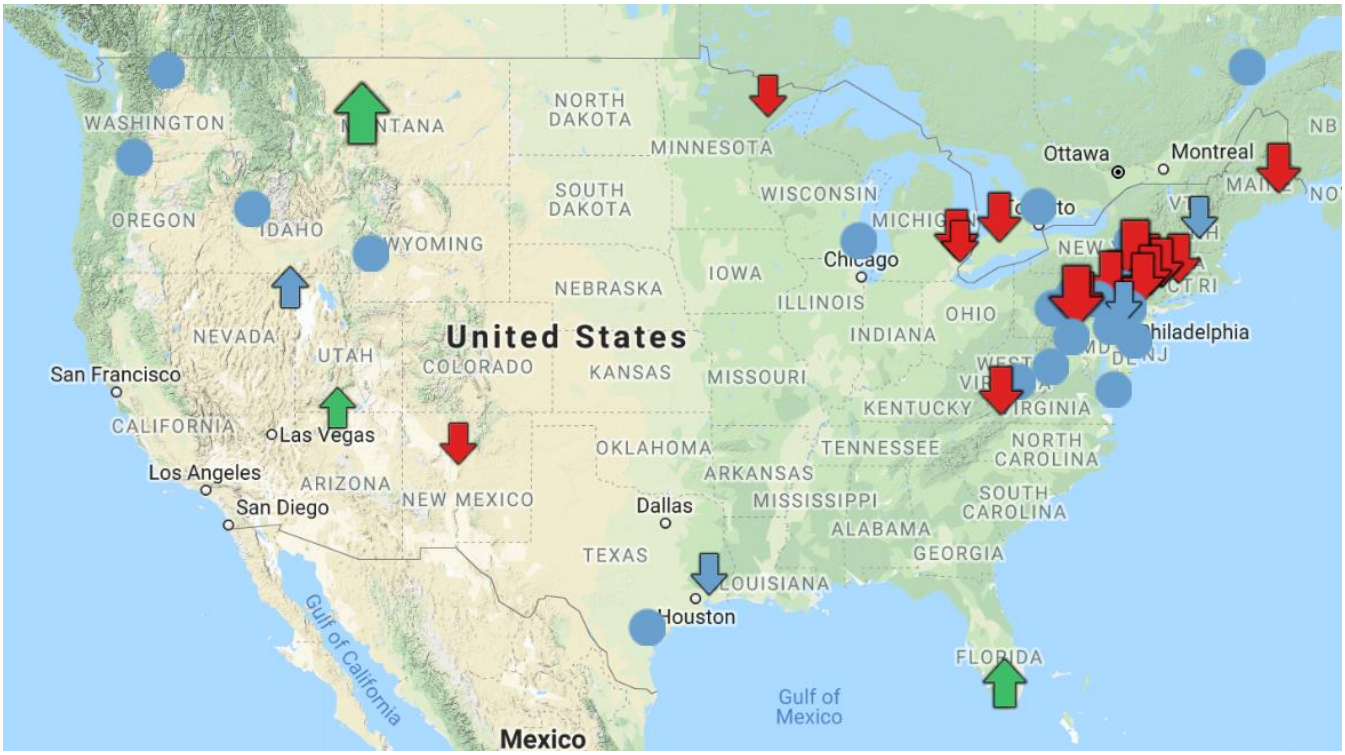

Figure SI 4. Population trends for American kestrels from Raptor Population Index.

A summary of population trend estimates based on migration count data from the Raptor Population Index (RPI) project (Crewe et al. 2016; Brandes et al. 2016). Green arrows represent areas of significant increase, blue arrows represent areas of slight increase (pointing up) or decrease (pointing down), red arrows represent areas of significant decrease, and blue circles signify areas with stable population trends. These data demonstrate well the variability of population trend estimates of American kestrels in North America, with the most decline occurring in the east.

APPENDIX A

The American Kestrel Genoscape

Introduction

American kestrels (*Falco sparverius*) are declining (Butcher et al. 1990; Smallwood et al. 2009; Farmer & Smith 2009; Sauer et al. 2014) and effective management requires identifying which populations are most vulnerable, where they are most limited, and how climate change will impact patterns of decline. Kestrels are monitored during migration by counting individuals at watch-sites that are stationed along major migratory flyways and these data are used to estimate population trends (Rich et al. 2005). However, kestrels are shifting the timing and patterns of their migratory movements (Heath et al. 2012; Paprocki et al. 2014), which could make inferring population trends from migration-counts difficult (McClure et al. 2017). If we want to maintain the utility of migration-counts into the future, we need to develop methods for understanding which populations of kestrels are being monitored at specific watch-sites during migration, and how migratory movements of those populations are changing over time.

Historically, we have studied migratory connectivity, or the link between specific breeding, wintering, and migrating populations, with long-term banding programs and small tracking devices. However, these methods are limited by low recapture rates, and in some cases can be labor intensive or cost prohibitive. An alternative method is to develop high-resolution molecular markers that allow the sampling of birds on the wintering grounds, or during migration, and use the DNA from a single feather to map that individual back to its breeding population of origin (Ruegg et al. 2014). The American kestrel Genoscape Project is a collaboration between Boise State, HawkWatch

International, The Peregrine Fund, St. Mary's University, and UCLA with partners from Hawk Mountain, USFWS, Florida Fish and Wildlife Conservation Commission, Texas Tech University, University of Northern British Columbia, Pacific University, and the Shenandoah Valley Raptor Study Area to develop high-resolution molecular markers for creating a spatially explicit map of American kestrel breeding populations.

The first step for developing molecular tools for population assignment is to identify biologically meaningful populations at spatial scales that are relevant to conservation management. The next step is to develop a panel of SNPs that will allow us to gain information about an individual's breeding population of origin from the DNA of a single feather. Development of a SNP panel for population assignment will allow us to sample individuals during migration or on the wintering grounds and determine where that individual originated.

In 2015-2016 we collected high-quality genetic samples from nine locations on the peripheral edges of the kestrel's breeding range. We used restriction-site associated DNA sequencing (RAD-seq) to discover single-nucleotide polymorphisms (SNPs) across the genome of the American kestrel. Then, we used population genetic analyses to identify genetically distinct populations of kestrels across North America. As described in this thesis, our results suggest that kestrels can be split into five genetically distinct units for conservation management.

Here, we have outlined the methods for developing a panel of rapid SNP-type assays (Fluidigm Inc.) for population assignment. We used the genome and RAD-sequencing data generated from our investigation of genome-wide patterns of population structure. Then we genotyped an additional 413 samples that were collected across 32

breeding populations in the United States and Canada (Table A1) at the 192 SNP assays we developed for population assignment. We used those genotypes to validate our results that suggest kestrels can be split into five genetically distinct populations across North America.

Methods

Development of assays for genetic population assignment

We were most interested in separating five major groups, including Alaska, Texas, the west, the east and Florida. To identify SNPs useful for distinguishing genetically distinct breeding groups within the American kestrel, we used VCFTOOLS (Danecek et al. 2011) to calculate site-wise Fst between the five clusters identified from our analyses of population structure to identify SNPs useful for distinguishing genetically distinct breeding groups within the American kestrel. We used the R package SNPS2ASSAYSRE to evaluate which of our top-ranking SNPs would generate designable assays. We considered SNPs to be designable if GC content was less than 0.65, there were no insertions or deletions (indels) within 30bp, or ambiguous codes within 20bp of the variable site. Additionally, we used bwa (Burrows-Wheeler Aligner; Li & Durbin 2009) to determine which of our designable SNPs mapped uniquely to the reference genome. We developed Fluidigm SNPtype assays (Fluidigm Inc.) for 182 SNPs that were considered designable and had the largest site-wise Fst values from each of our comparisons. In addition to the SNPs that distinguished between groups, we developed Fluidigm SNPtype assays for 34 SNPs that were associated with candidate migration genes, candidate circadian rhythm genes, or climate variables (see Ruegg et al. 2018 for methods).

Genetic screening and building the genoscape

We screened 93 kestrel feather samples and three non-template controls on the Fluidigm Corporation EP1 Genotyping System (Fluidigm Inc.) to evaluate each SNP assay and identify the most reliable 192 SNP assays of 216 that were designed. Then we used the 192 SNP assay set to screen 413 kestrel feather samples from 32 breeding locations in the United States and Canada (Table 2). We followed the extraction methods described above and used the methods in Ruegg et al. (2014) to amplify PCR products using fluorescently labelled allele-specific primers. Then, we used an EP1 Array Reader to take an image of the results, and Fluidigm's automated Genotyping Analysis Software (Fluidigm Inc.) to call alleles with a confidence threshold of 90%. In addition, we visually inspected each genotype call that did not fall clearly into one of three clusters, heterozygote or either homozygote, and removed uncertain genotype calls from the analysis. Samples with missing genotypes at more than 50 SNP assays were removed from our analyses of spatially-explicit population structure. After removing samples with missing data, we retained a total of 375 individuals for analysis (Table 2). We used an admixture model (program STRUCTURE 2.3.4) with uncorrelated allele frequencies, a burn-in period of 50,000 iterations, and run length of 150 000 to run 10 iterations of each assumed number of genetic clusters (K), where K ranged from 1:10 to assess population structure across the breeding grounds (Pritchard et al. 2000). We used the information from STRUCTURE to create a spatially-explicit figure of genetic groups.

Assignment accuracy of the SNP panel

We assessed the assignment accuracy of the SNP panel using two approaches. First, we examined whether known-origin American kestrels from the 5 genetically-

distinct groups could be correctly assigned to their "group origin" using the R package RUBIAS. We tallied the proportion of correctly-assigned individuals for 1) the samples used to design the assay and a test set of 375400 individuals, and 2) the test set only. Then, we assessed assignment accuracy by building a spatially-interpolated genetic map using the posterior probability of assignment from individuals of known origin, and predicting the latitude and longitude of each individual, using the R package ORIGIN. We created separate maps for 1) western North American that consisted of the Alaska, western, and Texas kestrel groups, and 2) eastern North America that consisted of Texas, eastern, and Florida groups. We measured the distance between sample site and predicted assignment site to represent measurement error. Finally, we used RUBIAS to assign a breeding group to unknown-origin kestrels captured during migration or on wintering areas in the non-breeding season.

Results

Development of assays for genetic population assignment

We developed 192 SNP-type assays for population assignment of American kestrel's in North America. We successfully genotyped 413 samples collected from 32 breeding populations in the United States and Canada. Of the 182 designable SNPs, 162 provided reliable genotype information for population assignment. In addition, 15 SNPs associated with candidate genes (Table 4) and 15 SNPs associated with climate were included on the 192 SNP panel used to genotype kestrels.

Genetic screening and building the genoscape

Combined, all of the 192 SNPs reveal 5 genetically distinct groups of American Kestrels in North America (Figure 4). This pattern of 5 groups is the same level of

structure revealed in an analysis of 72,263 SNPs, thus we are confident that our SNP assay adequately captures existing genetic structure of American Kestrels.

Assignment Accuracy of the SNP panel

We found that most kestrels were assigned to the correct genetic group using results from the SNP panel (Table 5). The accuracy of assignment based on ORIGEN interpolated maps ranged from 0 – 4000 km with a median of ~ 400 km (Figure 7) and birds were not assigned to their sampling location (Figures 8 and 9). Assignment of migrating and wintering kestrels to genetically distinct groups was successful (Figure 9). Kestrels in the western portion of North American assigned the "West" group and kestrels in Florida assigned to the "East" group. We would predict that these kestrels are wintering birds from the northeast because they were captured in December-February outside of the range of "Florida" kestrels.

Discussion

Assignment of the 375 samples that we screened at 192 high-resolution SNP assays validates our claim that kestrels form five genetically distinct units in North America: samples collected from Texas, Alaska and Florida assigned with the highest probability back to their population of origin; samples collected from populations in Arizona, California, Oregon, Idaho, Utah, Wyoming, Montana, Alberta and Saskatchewan assigned with the highest probability to the western reporting unit; and samples collected from Nebraska, Kansas, Wisconsin, Illinois, Missouri, Tennessee, Michigan, Wisconsin, Pennsylvania, Connecticut, New York, New Jersey, Delaware, Virginia and North Carolina assigned with the highest probability to the eastern reporting unit (Figure A1). These results support our claim that kestrels should be split regionally

into eastern and western management units that are independent from Alaska, Texas and Florida.

Conclusions

Kestrels are declining for unknown reasons across North American, but patterns of decline vary regionally, which has made it difficult to interpret estimates of population trends. Many causes of the observed declines have been suggested, but we lack empirical evidence to support any hypothesized cause. Given the regional differences in population trends, and the genetic distinctiveness of regional groups, it is likely that kestrels in the eastern and western regions of North America are experiencing different constraints. Considering eastern and western kestrels as independent units for management purposes will improve our ability to interpret regional estimates of population trends and help us understand which of the five regional groups are being monitored at migration sites and how those groups might be responding to various constraints (e.g., loss of habitat, climate change, etc.); information that will be critical for migration-monitoring in the future.

REFERENCES

Butcher GS (1990) Audubon Christmas Bird Counts. Pages 5–13 in Sauer JR & Droege S [EDS.], Survey designs and statistical methods for the estimation of avian population trends. Biological Report 90(1). U.S.D.I. Fish and Wildlife Service, Washington, DC U.S.A.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156-2158.

Farmer C & Smith J (2009) Migration Monitoring Indicates Widespread Declines of American kestrels (*Falco sparverius*) in North America. *Journal of Raptor Research*, 43(4), 263-273.

Heath, JA, Steenhof K & Foster MA (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.

Li H & Durbin R (2009) Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*, 25, 1754–60.

McClure CJW, Schulwitz SE, Van Buskirk R, Pauli BP & Heath JA (2017) Commentary: research recommendations for understanding the decline of American kestrels (*Falco sparverius*) across much of North America. *Journal of Raptor Research*, 51(4), 455-464.

Paprocki N, Heath JA & Novak SJ (2014) Regional Distribution Shifts Help Explain Local Changes in Wintering Raptor Abundance: Implications for Interpreting Population Trends. *PLOS One*, 9(1), e86814

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

Rich TD, Beardmore CJ, Berlanga H, Blancher PJ, Bradstreet MSW, Butcher GS, Demarest DW, Dunn EH, Hunter WC, Iñigo-Elias EE, Kennedy JA, Martell AM, Panjabi AO, Pashley DN, Rosenberg KV, Rustay CM, Wendt JS & Will TC. 2004. Partners in Flight North American Landbird Conservation Plan. Cornell Lab of Ornithology. Ithaca, NY. Partners in Flight website. http://www.partnersinflight.org/cont_plan/ (VERSION: March 2005).

Ruegg KC, Anderson EC, Paxton KL, Apkenas V, Lao S, Siegel RB, Desante DF, Moore F & Smith TB (2014) Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23, 5726-5739.

Sauer JR, Hines JE, Fallon JE, Pardieck KL, Zoilkowski DJ & Link WA (2014) The North American Breeding Bird Survey, results and analysis 1966–2012. Version 02.19.2014. U.S.G.S. Patuxent Wildlife Research Center, Laurel, MD U.S.A.

Smallwood JA, Causey MF, Mossop DH, Klucsarits JR, Robertson B, Robertson S, Mason J, Maurer MJ, Melvin RJ, Dawson RD, Bortolotti GR, Parrish JW, Breen TF & Boyd K (2009) Why are American kestrel (*Falco sparverius*) populations declining in North America? Evidence from nest-box programs. *Journal of Raptor Research*, 43, 274-282.

TABLES

Table A1. Table of the 413 individuals genotyped for 192 SNPs selected for population assignment, their origin location (site), centroid latitude and longitude, and total sample size (N_total). Samples with missing genotypes at more than 50 SNPs were removed from our analyses of population structure. After removing samples (N_post-filtering), we retained a total of 375 individuals. States with numbers have more than one sampling area.

Study_site	Latitude	Longitude	$N_{\_total}$	$N_{\_post\text{-}filtering}$
Alberta, Canada	56.16262	-117.34	12	12
Alaska, USA	63.96041	-145.136	7	7
Arizona, USA	33.45007	-111.943	20	15
California (1), USA	37.23551	-120.244	22	22
California (2), USA	33.7315	-117.878	3	3
Connecticut (1), USA	41.95001	-73.364	10	10
Connecticut (2), USA	41.94374	-72.5072	10	7
Delaware, USA	39.75047	-75.6969	4	4
Florida, USA	29.54151	-82.2334	20	20
Idaho (1), USA	43.51209	-116.377	18	11
Idaho (2), USA	43.27597	-115.003	6	5
Illinois, USA	38.30563	-89.7969	5	3
Kansas, USA	38.77533	-95.2486	4	4
Massachusetts, USA	42.15064	-71.8496	20	18
Mississippi, USA	44.15112	-84.9633	20	19
Missouri (1), USA	39.04824	-94.3663	8	5
Missouri (2), USA	38.3804	-90.925	1	1
Montana, USA	45.76764	-111.154	19	15
North Carolina, USA	35.73285	-82.6776	17	16
Nebraska, USA	40.99115	-96.8957	6	6
New Jersey, USA	40.4297	-74.9271	30	28
New York, USA	41.8773	-74.7309	4	3
Oregon (1), USA	45.57362	-123.048	3	3
Oregon (2), USA	44.36466	-121.412	25	24
Pennsylvania, USA	40.60183	-75.5961	13	13
Saskatchewan, Canada	55.01968	-105.893	12	11
Tennessee, USA	35.14944	-90.0489	5	5
Texas, USA	33.63184	-101.889	1	1
Utah, USA	40.62255	-111.987	30	28
Virginia, USA	38.6457	-78.7522	17	17

Wisconsin, USA	43.77114	-89.4059	24	24
Wyoming, USA	43.60488	-110.658	17	15

FIGURES

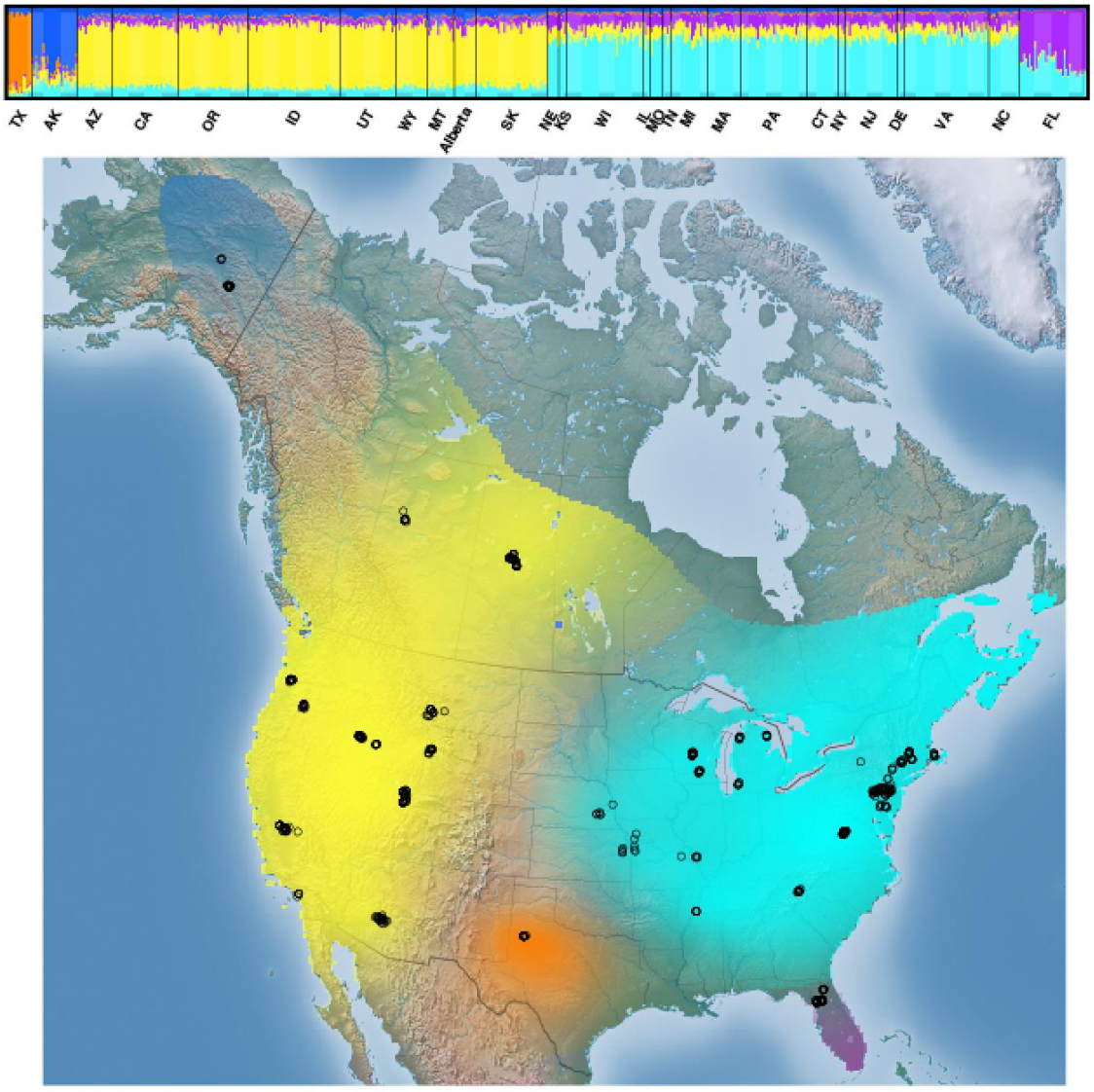

Figure A1. The American kestrel Genoscape reveals five genetically distinct groups (K = 5, represented by colored polygons) within North America.

A. Each bar represents an individual, and individuals are organized by geographic location. B. Black circles show sites where American kestrel DNA was collected via

blood or feather collection. Samples collected from Texas, Alaska and Florida have the greatest assignment probability to distinct clusters. Samples collected in western North America have the greatest assignment probability to the western reporting unit. Similarly, samples collected in eastern North America have the greatest assignment to the western reporting unit. These results support our claim that kestrels should be considered as five genetically distinct units for management in the United States and Canada.