LANDSCAPE LEVEL PATTERNS OF GENETIC DIVERSITY OF BAT SPECIES OF GREATEST CONSERVATION NEED IN LOUISIANA

JENNA R. GRIMSHAW, JOSEPH D. MANTHEY, NICOLE S. PAULAT, RICHARD D. STEVENS, AND DAVID A. RAY

ABSTRACT

Loss of genetic diversity decreases fitness and increases risk of extinction. Therefore, understanding population genetic structure is important to conserving populations, especially those at risk. Several species of bats in Louisiana recently have been listed as Species of Greatest Conservation Need, including *Eptesicus fuscus*, *Myotis austroriparius*, and the federally threatened *Myotis septentrionalis*. To characterize genetic diversity, heterozygosity and inbreeding coefficients were calculated using double-digest RADseq data obtained from 41 *E. fuscus*, 31 *M. austroriparius*, and 30 *M. septentrionalis*. Nonparametric (k-means and DAPC) and parametric (FastStructure and STRUCTURE) methods were implemented to ascertain the existence of population structure. Both nonparametric methods exhibited no clear population structure. FastStructure found no evidence of population structure and STRUCTURE estimated two genetic clusters for *E. fuscus* and *M. austroriparius* and three genetic clusters for *M. septentrionalis*. Membership assignment to genetic clustering from STRUCTURE was not related to geography. Due to these inconsistencies, there is lack of support for recognizing more than one interbreeding population in Louisiana of any of the species examined. Evidence suggests the presence of inbreeding for all three species of bats, which could increase risk of extinction. Conservation of suitable habitat should be pursued to protect species of bats in Louisiana and further genetic studies should characterize and compare other species of bats not only in Louisiana but also outside the state.

Key words: bats, *Eptesicus*, fastStructure, genetic diversity, heterozygosity, inbreeding, Louisiana, *Myotis*, population structure, STRUCTURE

Supplementary materials are available at https://github.com/jennagrimshaw/LAgenetics.

INTRODUCTION

Although many species worldwide are threatened by climate change, habitat loss, and anthropogenic activity, bats contend with additional challenges such as wind farms, human persecution, and in North America, white-nose syndrome (WNS; Voigt and Kingston 2015).
Many bat species are now facing extinction, or at the very least, declines in population size, with 21% of bat species now listed in threatened categories in comparison to 15% in 2015 (Jones et al. 2009; Voigt and Kingston 2015; IUCN 2021). Therefore, to prevent further decline it is imperative to monitor bat populations and their associated extinction risks (Jones et al. 2009).

An important aspect of evaluating populations is genetic diversity because the loss of such diversity increases effects of inbreeding, decreases sperm quality, and increases susceptibility to disease and parasites, thereby increasing possibility of extinction (Ralls et al. 1988; Coltman et al. 1999; Reed and Frankham 2003; Hedrick and Fredrickson 2010; Pekkala et al. 2014), although the relationship between genetic diversity and extinction is not simple and has been debated (Teixeira and Huber 2021). Genetic data may provide insights into population demographic histories, current population structure, and future extinction risks. For example, low genetic diversity may indicate historical bottlenecks or reduced gene flow among populations resulting in decreased heterozygosity (Hedrick and Miller 1992; Frankham 2018). Understanding population structure (presence of one interbreeding population or multiple isolated populations) is particularly important because as populations become increasingly isolated, gene flow is reduced, leading to higher inbreeding and further loss of genetic variation (Frankham 2005; Ralls et al. 2013).

Louisiana is home to 12 species of bats, including three that are the focus of this study: Eptesicus fuscus (Big Brown Bat), Myotis austroriparius (Southeastern Myotis), and the now-threatened M. septentrionalis (Northern Long-eared Myotis) (Lowery 1974; Crnkovic 2003). Myotis septentrionalis was recently listed as threatened by the United States Fish and Wildlife Service due to high mortality rates from WNS (United States Fish and Wildlife Service 2015; Holcomb et al. 2019). The main objective of this study was to better understand population genetics of bats in Louisiana by characterizing genetic diversity and population structure of E. fuscus, M. austroriparius, and M. septentrionalis. Specifically, our goals were to: 1) determine nucleotide diversity, levels of heterozygosity, and inbreeding for each of the three species; and 2) ascertain if population genetic structure exists for these species in Louisiana.

We expected to find lower heterozygosity and higher levels of inbreeding for M. septentrionalis due to severe population declines experienced by this species. Previous genetic studies found little population structure for E. fuscus in Indiana and Illinois (Vonhof et al. 2008), as well as for M. septentrionalis in Canada (Johnson et al. 2015) and in New York and West Virginia (Johnson et al. 2014). Consequently, we did not expect to identify strong genetic structure for any of the species within Louisiana owing to the high vagility of bats and a lack of conspicuous geographic boundaries.

**Methods**

**Focal taxa.**—Eptesicus fuscus is a relatively large brown bat (12–30 g), that is distributed throughout the United States and forages among treetops primarily on hard-bodied beetles and other insects (Kurta and Baker 1990; Feldhamer et al. 2009). The smaller M. austroriparius (5–12 g) from the southeastern United States prefers to roost in caves but can be found in buildings and hollow trees (Jones and Manning 1989). These bats feed mainly on soft-bodied insects such as caddisflies, moths, and mosquitos (Feldhamer et al. 2009). Once common throughout the central and northeastern United States as well as Canada, M. septentrionalis is a small gleaning insectivore (5–7 g) that roosts in trees and artificial structures and forages on moths, beetles, and spiders (Caceres and Barclay 2000; Brack and Whitaker 2001; Feldhamer et al. 2009). All three species are distributed widely throughout Louisiana although M. septentrionalis is rarer than the other two (Stevens et al. 2017, 2020).

**DNA sequencing.**—DNA samples were analyzed for E. fuscus, M. austroriparius, and M. septentrionalis from Louisiana (Fig. 1). Because M. septentrionalis was the least abundant species (n = 37), this taxon was used to determine the geographic domain of this study. Geographic domain was set to minimize any differences in genetic diversity owing to spatial scale. Geographical coordinates for M. septentrionalis cap-
Species Occurrences

- E. fuscus
- M. austroriparius
- M. septentrionalis

Figure 1. Sampling sites for occurrences of *Eptesicus fuscus*, *Myotis austroriparius*, and *M. septentrionalis* from Louisiana that were used in this study.


tures were entered into ArcMap GIS and buffers (10, 15, 20, 25, 30, 35, and 40 km) were generated around each coordinate. These were then intersected with coordinate data for all *E. fuscus* (*n* = 98) and *M. austroriparius* (*n* = 196). Abundance of *E. fuscus* and *M. austroriparius* within the given buffer was determined. A 30 km buffer provided the most similar abundances and spatial extents for all three species. There were 37 *M. septentrionalis*, 34 *M. austroriparius*, and 50 *E. fuscus* retained for genetic analysis, although several individuals were removed from downstream analysis. For *M. septentrionalis*, wing punches from bats that were caught and released throughout Louisiana served as the source of genetic material following approved protocol (Sikes et al. 2016). Specimens from the Natural Sciences Research Laboratory (NSRL) at the Museum of Texas Tech University that were previously collected (i.e., before *M. septentrionalis* was listed as Threatened under the ESA) from Louisiana served as a secondary source. For the other two species, liver samples were used from specimens collected during an extensive bat survey of Louisiana (Stevens et al. 2017) in addition to samples from the NSRL.
A Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) was used to isolate DNA following the recommended protocol including centrifugation for an additional minute after discarding waste material from the second wash (to avoid ethanol carryover) and repeating the elution step (to maximize DNA yield). After isolation, DNA was quantified with the Qubit 3.0 Fluorometer by ThermoFisher and samples were sent to Admera Health (South Plainfield, NJ) to be processed using a variant of double-digest restriction-site associated DNA (ddRADseq) that identifies polymorphic single nucleotide polymorphisms (SNPs) from across the genome (Peterson et al. 2012). This variation of RAD sequencing uses two restriction enzymes (in this study, EcoRI-MspI). Illumina HiSeqX methodology was used to generate 150 bp paired-end reads. AfterQC (Chen et al. 2017) was used to trim and filter out low quality reads by removing those with Q-scores below Q30 (base call accuracy of 99.9%), more than 5 Ns, with a mononucleotide tract longer than 35 bp, or read lengths less than 35 bp. All reads were uploaded to NCBI’s sequence read archive (BioProject accession PRJNA 785461).

Species-specific reference genomes were used to improve alignment accuracy. A de novo assembly for *E. fuscus* (NCBI accession GCA_000308155.1) and a chromosome-length genome assembly for *M. septentrionalis* from the DNA Zoo Consortium (Dudchenko et al. 2017) were used as reference genomes. For *M. australriparius*, low coverage sequencing was used to generate a reference-based assembly using *M. lucifugus* de novo assembly (Supplementary Materials 1). RAD-seq data were then aligned to their respective assembly using *BWA* (Burrows-Wheeler Alignment tools) and *SAMtools* (Sequence Alignment/Map; Li et al. 2009; Li and Dur 2009). Loci were identified and genotyped with the module *gstacks* (Stacks v 2.52), which was written to analyze RAD sequence data (Catchen et al. 2013), and alignments that were of poor quality or had mapping errors were filtered out and discarded. The genotyped data were analyzed with Stacks’ populations module and only loci found in at least 80% of the samples for each species were retained. A minimum allele count of three was required to remove low frequency SNPs called due to sequencing errors, i.e., an allele had to appear at least three times to be retained (O’Leary et al. 2018). To limit processing linked SNPs and potential biases in downstream analyses, we selected the first SNP from each locus (Raj et al. 2014). After running the populations module, individuals that were missing more than 25% of loci were removed and *gstacks* and populations were rerun with the remaining individuals.

**Genetic analyses.**—Basic population statistics including expected and observed heterozygosity (Hardy 1908; Mayo 2008) and Wright’s inbreeding coefficient (Wright 1951) were calculated for each species. Observed and expected heterozygosity and inbreeding coefficients are based on allele frequencies (proportion of major and minor alleles, p and q, respectively) and the Hardy-Weinberg Equilibrium (HWE) principle which states that in the absence of evolutionary mechanisms (i.e., mutation, migration, nonrandom mating, natural selection, and genetic drift), a population’s allele and genotype frequencies will remain in equilibrium (Hardy 1908; Mayo 2008). Deviations from HWE provide clues to potential mechanisms acting on the population. At equilibrium, observed heterozygosity (ranging from 0 to 0.5) equals the population’s expected heterozygosity, the probability that an individual is heterozygous at a certain locus based on allele frequencies and calculated as 2pq. Alternatively, excessive inbreeding leads to observed heterozygosity being lower than expected and is a sign that a population could be at risk (Wright 1951; Hartl and Clark 1980; Weir and Cockerham 1984). This is quantified using Wright’s inbreeding coefficient (F_€; ranging from 0 to 1), which incorporates observed and expected heterozygosity, whereby higher values indicating higher levels of inbreeding (Wright 1951; Hartl and Clark 1980; Weir and Cockerham 1984). Heterozygosity and inbreeding were calculated on all retained loci (variant and invariant).

**Population structure.**—Nonparametric (k-means clustering and DAPC) and parametric (fastStructure and STRUCTURE) methods were used to determine potential population genetic structure. K-means clustering and discriminant analysis of principal components (DAPC) were performed using the adegenet R package (Jombart and Ahmed 2011). K-means clustering identified the number of groups (K) by comparing multiple values of K and choosing the one with the lowest Bayesian Information Criterion (BIC) (Jombart et al. 2010). DAPC conducts a discriminant analysis (DA) by first transforming genetic data using principal components analysis (PCA) followed by a DA on the PCA scores to optimize among-group variation...
while minimizing within-group variation (Jombart et al. 2010).

FastStructure was then used to detect allele frequency differences and estimate the log-marginal likelihood that an individual belongs to each genetic cluster. The software then assigned individuals to the genetic cluster with the highest likelihood (Raj et al. 2014). FastStructure was run to assess the number of potential clusters (K = 1 through K = 10). Each iteration was repeated ten times and the model with the highest mean likelihood was chosen (Pritchard et al. 2000). Finally, STRUCTURE was run for K = 1 through K = 5 to further investigate potential genetic clusters (Pritchard et al. 2000). STRUCTURE is computationally two orders of magnitude slower, but may detect fine-scaled clustering more accurately than fastStructure (Stift et al. 2019). The admixture model was selected with 50,000 burn-in replicates and 250,000 Markov Chain Monte Carlo (MCMC) iterations (Pritchard 2007; Gilbert et al. 2012). STRUCTURE results were then analyzed using StructureHarvester (Earl and vonHoldt 2012) using both the Pritchard method (Pritchard et al. 2000) and the Evanno method (Evanno et al. 2005). The model with the optimal K was then plotted with Clumpp (Jakobsson and Rosenberg 2007), and DISTRICT (Rosenberg 2004) and membership coefficients were estimated for each individual representing the proportions of alleles originating from each ancestral cluster (Rosenberg 2004).

Results

DNA sequencing.—After removing samples due to low quality Phred scores or high proportion of missing loci, 41 E. fuscus, 31 M. austroriparius, and 30 M. septentrionalis were retained for analysis. Mean reads per individual were 1,168,643 for E. fuscus, 1,474,296 for M. austroriparius, and 1,473,420 for M. septentrionalis. Loci were genotyped and filtered for each species (E. fuscus: 89,284; M. austroriparius: 107,508; M. septentrionalis: 76,380). Additional information (e.g., percentage of reads removed, coverage, mean locus size, etc.) can be found in Supplementary Materials.

Genetic analyses.—Observed heterozygosity was lower than expected heterozygosity for all three species, suggesting the presence of inbreeding in Louisiana (Table 1). Eptesicus fuscus had expected heterozygosity values of 0.00056 and observed heterozygosity of 0.00040. Expected heterozygosity for M. austroriparius was 0.00052, and observed heterozygosity was 0.00043. Similarly, M. septentrionalis had lower observed heterozygosity values (0.00039) than expected heterozygosity (0.00062). Myotis septentrionalis had the highest inbreeding coefficient (0.00114) followed by E. fuscus (0.00090) and M. austroriparius (0.00039).

Population structure.—To examine population structure, 52,012 SNPs for E. fuscus, 41,004 for M. austroriparius, and 28,601 for M. septentrionalis were analyzed using K-means clustering, DAPC, fastStructure, and STRUCTURE. K-means clustering did not clearly indicate the best K for any of the species (Fig. 2A). Similarly, the DAPC exhibited very little structure. The first axis (that explains the greatest variation among groups) for each species accounted for only 4.1% of the variation for E. fuscus, 4.6% for M. austroriparius, and 6.9% for M. septentrionalis.

Parametric results (fastStructure and STRUCTURE) were less obvious and left more to interpretation. For E. fuscus, K = 1 had the highest likelihoods for fastStructure (Fig. 2B) and K = 3 for the Pritchard and Evanno methods (Fig. 2C). FastStructure indicated K = 1 as the best choice of K for M. austroriparius (Fig. 2B) although the Pritchard and Evanno methods selected K = 2 as having the highest likelihoods (Fig. 2C). For M. septentrionalis, the best choice was K = 1 for fastStructure (Fig. 2B) and both methods of STRUCTURE (Fig. 2C). Membership probabilities were plotted for each species for K = 2 and K = 3 with individuals ordered based on geographic location (Fig. 3). Clustering did not appear to be related to geographic location as individuals assigned to the same clusters were often found on opposite sides of Louisiana (see examples in Figures 3C and 3D).
Table 1. Basic population statistics for *Eptesicus fuscus*, *Myotis austroriparius*, and *M. septentrionalis*. Samples and samples/locus refer to samples retained after filtering. $H_E$ = expected heterozygosity, $H_O$ = observed heterozygosity, and $F_{IS}$ = inbreeding coefficient.

<table>
<thead>
<tr>
<th>Species</th>
<th>Samples</th>
<th>Samples/locus</th>
<th>$H_E$</th>
<th>$H_O$</th>
<th>$F_{IS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. fuscus</em></td>
<td>41</td>
<td>36</td>
<td>0.00056</td>
<td>0.00040</td>
<td>0.00090</td>
</tr>
<tr>
<td><em>M. austroriparius</em></td>
<td>31</td>
<td>28</td>
<td>0.00052</td>
<td>0.00043</td>
<td>0.00039</td>
</tr>
<tr>
<td><em>M. septentrionalis</em></td>
<td>30</td>
<td>26</td>
<td>0.00062</td>
<td>0.00039</td>
<td>0.00114</td>
</tr>
</tbody>
</table>

Figure 2. Population structure results for *Eptesicus fuscus*, *Myotis austroriparius*, and *M. septentrionalis*: (A) K-means clustering results; (B) FastStructure mean likelihoods for one to ten estimated genetic clusters ($K = 1$ through $K = 10$); and (C) STRUCTURE results with solid line indicate estimated natural log probability for genetic clusters $K = 1$ through $K = 5$ (Pritchard method). Dashed line indicates Delta $K$ for genetic clusters $K = 2$ through $K = 4$ (Evanno method).
Figure 3. STRUCTURE results processed with Clumpp and Distruct for K = 2 and K = 3 for (A – B) *Eptesicus fuscus*, (C – D) *Myotis austroriparius*, and (E – F) *M. septentrionalis*. Each vertical line indicates estimated membership probabilities of an individual. Individuals are ordered based on geographic proximity.
Discussion

In this study, genetic diversity of *E. fuscus*, *M. austroriparius*, and *M. septentrionalis* in Louisiana was examined using RADseq data. Sequencing depth of coverage was in line with previous low-coverage RADseq studies for bats (Auteri and Knowles 2020; Pinzari et al. 2020) and other mammals (Shafer et al. 2017; Martin Cerezo et al. 2020). For all three species, observed heterozygosity was lower than expected, indicating potential inbreeding.

Nonparametric and parametric methods for detecting population structure were incongruent. There could be several reasons for this. First, nonparametric methods (K-means clustering and DAPC) are more robust than parametric methods to threshold decisions on minor alleles (Linck and Battey 2019) such as minimum number of times an allele must be present to be included (here, minimum allele count was set to three). Second, high levels of inbreeding can lead to an overestimation of K (Pritchard 2007). Third, the presence of very fine-scaled clusters unrelated to geographic location potentially due to bat vagility.

Although the Pritchard and Evanno methods seem to agree based on highest likelihoods, there are two caveats. Pritchard (2010) cautions that choosing the best K is not always as straightforward as calculating the highest likelihood and recommends when the probability of likelihoods plateau and several values of K give similar estimates that the smallest K is “correct.” The caveat is that what is considered “similar estimates” is left up to the researcher. It could be argued that the likelihood values for *E. fuscus* were plateauing at K = 1 through K = 3 and therefore, K = 1 was the better choice. Similarly, the likelihood values for *M. austroriparius* could be plateauing for K = 1 and K = 2. Additionally, the Evanno method is based on the second order rate of change and therefore by default cannot select K = 1, otherwise known as the K = 2 conundrum (Janes et al. 2017), and therefore is unreliable for *M. austroriparius* (especially given the Pritchard method indicated K may equal 1).

When choosing the most appropriate K, species biology also must be considered and the biological interpretation can be problematic (Pritchard 2007). The vagility of bats could effectively dissolve historical population structure as highly mobile species can experience a rapid loss of structure within a few generations (Landguth et al. 2010). Genetic admixture for species with large dispersal abilities has been shown in mountain goats (Wolf et al. 2020), minke whales (Quintela et al. 2014), and other bats (Vonhof and Russell 2015).

Our results suggest each of these species consists of a single panmictic population with genetic remnants of ancestral populations at least within Louisiana. This is supported by previous research on *E. fuscus* (Vonhof et al. 2008) and *M. septentrionalis* (Johnson et al. 2014; Johnson et al. 2015). To our knowledge, no population structure analyses have been conducted on *M. austroriparius*. The previous study on *E. fuscus* used nuclear microsatellite data to identify potential substructuring in Indiana (Vonhof et al. 2008). Their STRUCTURE analysis identified only one cluster (K = 1) and further analyses using AMOVA and pairwise Fst values revealed that less than 1% of genetic differences were among colonies sampled. A 2014 study on *M. septentrionalis* analyzed microsatellite data from captures in New Brunswick and Nova Scotia, Canada (Johnson et al. 2014); although multiple fine-scaled clusters were identified with the STRUCTURE analysis, clustering was not related to geographic location or distance among colonies and indicated that sex-biased dispersal may play a role although further research would be needed. Another *M. septentrionalis* study examined microsatellite data for bats captured in New York and West Virginia (Johnson et al. 2015) and concluded that groups of this species were genetically indistinguishable regardless of spatial scale studied.

While the threatened species *M. septentrionalis* did not have the lowest levels of heterozygosity in this study, it did have higher levels of inbreeding, which may put it at an increased risk for local extinction. One possibility is that the establishment of *M. septentrionalis* in Louisiana is recent (Crnkovic 2003), as recent range expansions can cause a loss of genetic diversity for populations on the leading edge (Pierce et al. 2017; Rougemont et al. 2020). Efforts to conserve suitable habitat should be pursued in Louisiana especially given that the state may become a refuge for this species in the
wake of WNS. Additional studies to examine genetic diversity of other Louisiana bats should be conducted as well as comparing individuals from Louisiana to those at other localities to further increase our understanding of genetic diversity in the state. Evaluating the genetic diversity of at-risk species is not only crucial for biological assessments used to build conservation strategies (Reed and Frankham 2003; Frankham 2005), but also contributes to the overall knowledge of population dynamics and structure.

ACKNOWLEDGMENTS

We would like to thank M. K. Halsey from TTU for her contribution in creating the analysis pipeline, H. Garner (TTU-NSRL) for assistance with specimens, and Louisiana Department of Wildlife and Fisheries for funding this project (#2000330811, LDWF). The High-Performance Computing Center at Texas Tech University provided computational infrastructure and technical support throughout the work. D.A.R. was supported by NSF-DEB 1838283 and NSF-IOS 2032006.

LITERATURE CITED


Addresses of authors:

**Jenna R. Grimshaw**

Department of Biological Sciences  
Texas Tech University  
Lubbock, TX 79409  
jenna.grimshaw@ttu.edu

**Joseph D. Manthey**

Department of Biological Sciences  
Texas Tech University  
Lubbock, TX 79409  
joseph.manthey@ttu.edu
NICOLE S. PAULAT

Department Biological Sciences
Texas Tech University
Lubbock, TX 79409
nicole.paulat@ttu.edu

DAVID A. RAY

Department of Biological Sciences
Texas Tech University
Lubbock, TX 79409
david.4-ray@gmail.com

RICHARD D. STEVENS

Department of Natural Resources and
Natural Science Research Laboratory
Museum of Texas Tech University
Texas Tech University
Lubbock, TX 79409
richard.stevens@ttu.edu