MATING SYSTEM AND GENETIC STRUCTURE OF BROWN-HEADED NUTHATCHES (*Sitta pusilla*) AND THEIR SISTER, PYGMY NUTHATCHES (*Sitta pygmaea*)

By

KIN – LAN HAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2016
To my family and friends for their support and encouragement
ACKNOWLEDGMENTS

Thank you to my advisor and dissertation chair Rebecca Kimball for her patience and encouragement. I thank my committee members Ed Braun, Scott Robinson, Dave Steadman, and Katie Sieving for their advice and support.

I thank Jim Cox who taught me to think like a nuthatch and obtained permits, and helped to collect many of my samples. Tall Timbers Research Station provided funding and assistance. I also thank the U.S. Forest Service, U.S. National Park Service, Florida Forest Service, Florida Fish and Wildlife Commission, Florida Park Service, and the Southwest Florida Water Management District for providing permits, and in some instances, housing and additional help.

I thank Rosana Zenil-Ferguson for statistics and R advice. George Tiley for running some of my analyses. Without his help increasing the number of computation resources available, the analyses would not have finished in time. I would also like to acknowledge the University of Florida High Performance Computing Center for providing computational resources and support, particularly Matt Gitzendanner, who answered endless questions and helped troubleshoot my scripts.

Thanks to the Department of Biology staff during my time in the department: Susan Spaulding, Karen Patterson, Tangelyn Mitchell, Amy Dechow, Leila Long, Kyle Arola, Boyd Westerman, and Dave Reddig.

Thanks to Carol Chaffee, Melania Lopez-Castro, Tasha Hook, Brian Tierney, Nancy Gonzalez, Matt Salomon, Rosana Zenil-Ferguson, Ping Huang and Akanksha Pandey for their friendship and support through the years. I thank my family for their love and support. Without them, I would not have made it this far.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>7</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>8</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>9</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. BACKGROUND INFORMATION</td>
<td>13</td>
</tr>
<tr>
<td>2. UNCOMMON LEVELS OF RELATEDNESS AND PARENAGE IN A COOPERATIVELY BREEDING BIRD, THE BROWN-HEADED NUTHATCH (<em>SITTA PUSILLA</em>)</td>
<td>17</td>
</tr>
<tr>
<td>Introduction</td>
<td>17</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>18</td>
</tr>
<tr>
<td>Study Site and Field Procedures</td>
<td>18</td>
</tr>
<tr>
<td>DNA Sampling</td>
<td>20</td>
</tr>
<tr>
<td>Laboratory Methods</td>
<td>20</td>
</tr>
<tr>
<td>Microsatellite Variation</td>
<td>22</td>
</tr>
<tr>
<td>Analyses of Kinship, Parentage, and Group Characteristics</td>
<td>22</td>
</tr>
<tr>
<td>Results</td>
<td>23</td>
</tr>
<tr>
<td>Group Characteristics</td>
<td>23</td>
</tr>
<tr>
<td>Helper Relatedness</td>
<td>24</td>
</tr>
<tr>
<td>Parentage Analysis</td>
<td>25</td>
</tr>
<tr>
<td>Discussion</td>
<td>26</td>
</tr>
<tr>
<td>Group Structure</td>
<td>26</td>
</tr>
<tr>
<td>Extra-pair Paternity</td>
<td>28</td>
</tr>
<tr>
<td>3. THE EFFECT OF HABITAT FRAGMENTATION ON POPULATION GENETIC STRUCTURE OF BROWN-HEADED NUTHATCHES (<em>SITTA PUSILLA</em>)</td>
<td>33</td>
</tr>
<tr>
<td>Introduction</td>
<td>33</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>37</td>
</tr>
<tr>
<td>Sampling</td>
<td>37</td>
</tr>
<tr>
<td>Laboratory Methods</td>
<td>37</td>
</tr>
<tr>
<td>Genetic Variation and Population Differentiation</td>
<td>38</td>
</tr>
<tr>
<td>Results</td>
<td>39</td>
</tr>
<tr>
<td>Population Structure</td>
<td>40</td>
</tr>
<tr>
<td>Discussion</td>
<td>41</td>
</tr>
</tbody>
</table>
LIMITED POPULATION STRUCTURE AMONG MOUNTAIN POPULATIONS
OF THE PYGMY NUTHATCH (SITTA PYGMAEA) ..................................................52

Introduction .................................................................................................................52
Methods ..........................................................................................................................55
  Laboratory methods .................................................................................................55
  Genetic variation and population differentiation .......................................................55
Results ............................................................................................................................57
Discussion ......................................................................................................................58

THE EFFECTS OF LIMITED SAMPLE SIZES ON THE PROGRAM STRUCTURE ..................68

Introduction .................................................................................................................68
Methods ..........................................................................................................................70
Results ............................................................................................................................70
Discussion ......................................................................................................................71

CONCLUSIONS ...........................................................................................................74

LIST OF REFERENCES ..................................................................................................78

BIOGRAPHICAL SKETCH ..............................................................................................90
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Number of pair and cooperative nests sampled by year and cooperative group size.</td>
<td>31</td>
</tr>
<tr>
<td>3-1</td>
<td>Sample sizes (N), average observed (H₀) and expected (Hₑ) heterozygosities, number of alleles (Nₐ), allelic richness (Aᵣ) and private allelic richness (PAr) for 12 Florida brown-headed nuthatch populations.</td>
<td>46</td>
</tr>
<tr>
<td>3-2</td>
<td>Pairwise estimates of Fₛₜ (below diagonal) and Jost’s D (above diagonal) for 12 Florida brown-headed nuthatch populations.</td>
<td>47</td>
</tr>
<tr>
<td>4-1</td>
<td>Diversity statistics for 9 microsatellite loci.</td>
<td>62</td>
</tr>
<tr>
<td>4-2</td>
<td>Sample sizes (N), average observed (Ho) and expected heterozygosities (He), number of alleles (Na) average allelic richness (Ar) and private allelic richness (PAr) for 11 pygmy nuthatch (Sitta pygmaea) populations.</td>
<td>62</td>
</tr>
<tr>
<td>4-3</td>
<td>Pairwise estimates of Fₛₜ (below diagonal) and Jost's D (above diagonal) for 11 pygmy nuthatch (Sitta pygmaea) populations.</td>
<td>63</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2-1</td>
<td>The proportion of nests with extra-pair young in pair (n = 40) and cooperative (n = 19) groups.</td>
<td>32</td>
</tr>
<tr>
<td>3-1</td>
<td>Sampling localities of brown-headed nuthatches.</td>
<td>48</td>
</tr>
<tr>
<td>3-2</td>
<td>Average allelic richness for 12 sampled brown-headed nuthatch populations.</td>
<td>49</td>
</tr>
<tr>
<td>3-3</td>
<td>Structure plots for 12 sampled brown-headed nuthatch populations.</td>
<td>50</td>
</tr>
<tr>
<td>3-4</td>
<td>PCoA analysis</td>
<td>51</td>
</tr>
<tr>
<td>4-1</td>
<td>Sampling localities of 11 pygmy nuthatch populations.</td>
<td>64</td>
</tr>
<tr>
<td>4-2</td>
<td>STRUCTURE results of 11 pygmy nuthatch populations with no priors.</td>
<td>65</td>
</tr>
<tr>
<td>4-3</td>
<td>STRUCTURE results for 11 sampled pygmy nuthatch populations analyzed using prior information (USELOCPRIOR).</td>
<td>66</td>
</tr>
<tr>
<td>4-4</td>
<td>Mantel’s test for correlation between Rousset’s genetic distance and geographic distance (r = 0.4398, p = 0.083) from 1000 randomizations.</td>
<td>67</td>
</tr>
<tr>
<td>5-1</td>
<td>Proportion of jackknifed datasets for each value of K analyzed without prior information (green) and with priors (blue).</td>
<td>73</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>AMOVA</td>
<td>Analysis of molecular variance</td>
<td></td>
</tr>
<tr>
<td>AP</td>
<td>Apalachicola</td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Allelic richness</td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>Big Cypress</td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>British Columbia</td>
<td></td>
</tr>
<tr>
<td>BW</td>
<td>Blackwater</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>Chiricahua</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>Coconino</td>
<td></td>
</tr>
<tr>
<td>FWS</td>
<td>Fish and Wildlife Service</td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>Goethe</td>
<td></td>
</tr>
<tr>
<td>HE</td>
<td>Expected heterozygosity</td>
<td></td>
</tr>
<tr>
<td>HO</td>
<td>Observed heterozygosity</td>
<td></td>
</tr>
<tr>
<td>IBD</td>
<td>isolation-by-distance</td>
<td></td>
</tr>
<tr>
<td>LOD</td>
<td>Log of odds</td>
<td></td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
<td></td>
</tr>
<tr>
<td>MN</td>
<td>Mono</td>
<td></td>
</tr>
<tr>
<td>MR</td>
<td>Morelos</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Number of individuals</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>Number of alleles</td>
<td></td>
</tr>
<tr>
<td>NL</td>
<td>Nuevo Léon</td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>Ocala</td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td>Osceola</td>
<td></td>
</tr>
<tr>
<td>PAR</td>
<td>Private allelic richness</td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>Pacific Coast</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
<td></td>
</tr>
<tr>
<td>PCoA</td>
<td>Principal coordinate analysis</td>
<td></td>
</tr>
<tr>
<td>RL</td>
<td>Richloam</td>
<td></td>
</tr>
<tr>
<td>SA</td>
<td>Sacramento</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>San Diego</td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>Siskiyou</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>St. Sebastian</td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>Starkey</td>
<td></td>
</tr>
<tr>
<td>TL</td>
<td>Three Lakes</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>Transverse Ranges</td>
<td></td>
</tr>
<tr>
<td>TTRS or TT</td>
<td>Tall Timbers Research Station</td>
<td></td>
</tr>
<tr>
<td>WE</td>
<td>Webb</td>
<td></td>
</tr>
</tbody>
</table>
Understanding the factors that affect population genetic structure and gene flow is important to understanding evolution, as well as being an important consideration in conservation and management decisions. Population connectivity is dependent on factors ranging from the spatial distribution of habitats to species' behavioral responses to that distribution. My dissertation explores the genetic mating system of brown-headed nuthatches to understand mating system dynamics, looks at the effect of habitat fragmentation, both recent and historical, on population genetic structure, and explores how sampling affects our estimates of population genetic structure.

Cooperative breeding, characterized by reproduction that occurs in groups with more than two adults, is a complex mating system that can influence genetic structure of populations. Since many adults forego reproduction, this can result in low effective population sizes and may lead to greater losses of genetic variation than a similar-sized population where reproduction is more evenly distributed among adults. I determined the relatedness of group members of brown-headed nuthatches, whether extra-pair paternity was present, and whether auxiliary adults within a group contributed
genetically to offspring at the nest. I found that most auxiliary adults were males related to one or both of members of the breeding pair. Furthermore, extra-pair paternity occurred in nests of both cooperative groups and pair groups, although auxiliary adults at the nest were not contributing genetically.

I then determined genetic structure among 12 populations of brown-headed nuthatches in Florida. I compared populations in habitats ranging from large, contiguous forests in northern Florida to regions now fragmented by human activities in southern Florida. Populations in northern Florida formed a large, panmictic group with high allelic diversity. Although I found the presence of weak population structure in southern Florida, and a reduction in the amount of genetic variation, it is not clear whether this is due to recent anthropogenic changes, or a more historical signature (e.g., due to natural barriers).

I then examined the effects of long-term habitat fragmentation by examining 11 populations of pygmy nuthatches. Pygmy nuthatches are distributed at high elevations on mountains of western North America, ranging from southern British Columbia to Mexico. I found very little genetic structure among populations, although I identify a unique population, Nuevo Léon in the Sierra Madre Oriental in Mexico. This suggests that nuthatches may be better at dispersing than previously suggested, though results were based on limited sampling per population. Finally, I explored the effects of numbers of individuals sampled on population structure by sub-sampling the more extensive brown-headed nuthatch dataset. These results suggest that limited sampling may reduce the estimates of population structure, and may be one contributing factor to the pygmy nuthatch results.
CHAPTER 1
BACKGROUND INFORMATION

Since genetic variation provides the raw material for natural selection to act upon, understanding the processes and patterns responsible for genetic variation and structure in populations is fundamental to our understanding of ecology, evolution and conservation of biodiversity. Current patterns of genetic diversity and population structure reflect a combination of historical factors, demographic factors, and current factors (e.g., Robin et al. 2010; Row et al. 2011; Zellmer and Knowles 2009). Ultimately, these patterns are determined by the interaction of gene flow with other factors such as genetic drift, selection, and/or mutation, and are driven by the degree of landscape connectivity for a species (Hartl and Clark 1997; Ricketts 2001; Charlesworth et al. 2003; Burney and Brumfield 2009).

Behavioral dynamics are important for understanding the genetic structure of populations (Altmann 1996; Ross 2001). Cooperative breeding is one mating system that could influence genetic structure (Double et al. 2005; Woxvold et al. 2006). Cooperative breeding is typically characterized by reproduction that occurs in groups comprised of two breeding adults with one or more auxiliary adults. Auxiliary adults are typically male offspring from previous years, resulting in family groups where auxiliary adults do not contribute genetically to offspring. However, relatedness among group members, and whether or not auxiliary adults contribute to reproduction at the nest is not well characterized for most cooperatively breeding birds (Hatchwell 2009). Although incest avoidance is thought to be typical, incest may still occur in some cooperatively breeding species (Koenig and Haydock 2004). Genetic studies have also shown that the extent of extra-group fertilizations can vary from rare to very common (e.g., Haig et
al. 1994; Dickinson et al. 1995; Hughes et al. 2003). These different dynamics can all affect genetic structure. For example, if auxiliary adults do not reproduce, this lowers effective population size, which could decrease genetic variation. Alternatively, if auxiliary adults do reproduce (and incest is avoided), this could increase genetic variation. Finally, if incest occurs, this may increase homozygosity and reduce genetic variation. In Chapter 2, I characterize the genetic mating system of brown-headed nuthatches, to determine relatedness of auxiliary adults, whether there is extra-pair paternity, and whether helpers contributed genetically to the nest. These results help both to understand how the mating system of this species may affect genetic structure, as well as to provide natural history information about this poorly studied species.

The spatial configuration of habitat patches in the landscape can also affect population structure. There are various ways the landscape may be structured, each of which may have different effects on population genetic structure. For example, there could be largely contiguous habitat with periodic barriers to gene flow such as rivers or mountains. Alternatively, habitats could be naturally fragmented into habitat islands (e.g., mountain tops), presumably with little gene flow among islands. Finally, human activities are converting formerly contiguous habitats into small, fragmented habitat islands. Depending on patterns of dispersal and historical connectivity, the population genetic structure of species found in these differently structured habitats would be expected to differ. For even limited levels of dispersal, populations in contiguous habitat could effectively be a single panmictic population (even if the distances are large), whereas populations found on habitat islands could be considered a metapopulation (Hanski 1999), and greater genetic structure would be expected in the latter case.
Anthropogenic habitat loss and fragmentation has resulted in a decline in the size of suitable habitat patches, as well as increased spatial isolation between patches (Fahrig 2003; Keyghobadi et al. 2005). As gene flow decreases, populations begin to diverge and differentiate due to factors such as local adaptation and genetic drift. Some populations may be small enough that, in the absence of gene flow, they suffer the negative effects of genetic drift and inbreeding, experiencing a great enough loss in genetic variability that they have the potential to go extinct. However, the degree of population genetic structure due to anthropogenic changes will depend on the length of time since fragmentation began, as well as the degree of fragmentation (e.g., size of remaining habitat patches and distance between patches). In Chapter 3 I examine the effect of habitat fragmentation on population genetic structure in brown-headed nuthatch populations in Florida. Brown-headed nuthatches are endemic to mature pine forests of the southeastern United States. Populations have been declining throughout much of their range presumably due to logging, land conversion, and fire suppression (Lloyd and Slater 2007; Slater et al. 2013; Wilson and Watts 1999), all of which should have led to fragmented patches of suitable habitat surrounded by regions of unsuitable habitat. The level of habitat fragmentation in Florida varies from larger, more extensive pine forests in northern Florida (very little fragmentation), to smaller fragments in southern Florida, where development has been much more extensive.

Populations can also be fragmented naturally over evolutionary timescales. Habitats found at high elevations on mountains are often surrounded by valleys of differing habitats (Mayr and Diamond 1976). Species in these high elevation habitats are isolated from each other, driving differentiation among populations. Because the
processes of isolation, reduced gene flow, and genetic drift are similar to those faced by populations in anthropogenically fragmented habitats, studying populations in high elevation habitats can inform on the long-term effects of habitat fragmentation. In Chapter 4, I examine the genetic structure of high elevation populations of pygmy nuthatches. Pygmy nuthatches are very closely related to brown-headed nuthatches. Like their sister species, they are cooperatively breeding birds found in close association with long needled pines, though distributed in western North America. They are sedentary birds with limited dispersal (Kingery and Ghalambour 2001), and so I would expect to see strong evidence of genetic structure. My sampling for this study was limited to very few individuals per population. Thus, for Chapter 5, I explore the effect of limited sampling on genetic structure by subsampling data from brown-headed nuthatches to match those of the pygmy nuthatch data and explore how this affects conclusions of population structure.
CHAPTER 2
UNCOMMON LEVELS OF RELATEDNESS AND PARENTAGE IN A COOPERATIVELY BREEDING BIRD, THE BROWN-HEADED NUTHATCH (SITTA PUSILLA)

Introduction

Cooperative breeding in birds is defined as a social group where more than two adults contribute to care of the nestlings at a single nest. Typically, these are family groups formed by a reproductive pair and non-breeding ‘auxiliaries’ or ‘helpers’. Among birds, helpers are usually adult male offspring from previous years (Ekman et al. 2004, see table 6.1 in Komdeur 2004), although cooperative groups may also include helpers related to just one breeding individual, helpers unrelated to either breeder, or a mixture of related and unrelated helpers (e.g., Reyer 1980, Haydock et al. 2001, Gienapp and Merilä 2010, Nam et al. 2010). Cooperative breeding is a complex social system that has been the subject of extensive study, in part, because of the diversity of breeding strategies that exists among cooperatively breeding birds and the effects that delayed breeding has on fitness (Brown 1987, Emlen 1997, Ligon and Burt 2004).

Genetic studies indicate that the extent of extra-group fertilizations observed among cooperatively breeding birds varies from very rare (e.g., Haig et al. 1994, Dickinson et al. 1995) to quite common (e.g., Mulder et al. 1994, Hughes et al. 2003). Variation in extra-pair fertilization and kinship can affect behaviors related to group affiliations (Du and Lu 2009), parental investment strategies (Green et al. 1995), and mate selection (Ferree and Dickinson 2011), suggesting that understanding relatedness is important to understanding the evolution of cooperative breeding behavior (Riehl 2013).
The brown-headed nuthatch (*Sitta pusilla*) was featured in some of the early inquiries into cooperative breeding (Skutch 1961), but has since received scant attention. Most helpers appear to be male offspring from previous years, and thus are often related to the breeding pair (Cox and Slater 2007, Haas et al. 2010). However, this characterization is based on small sample sizes and much remains unknown about group dynamics. I used molecular techniques and five years of field observations to characterize the social and genetic structure of brown-headed nuthatch breeding groups in Florida. Specifically, I ask (1) what is the relationship among adults at the nest; (2) is there evidence of multiple paternity in pairs and/or cooperative groups; and (3) if there is multiple paternity, is paternity shared among adults in cooperative groups, extra-group, or a combination of both? Addressing such questions could help to understand the evolution and maintenance of cooperative breeding in this species and also lead to improvements in future field studies.

**Materials and Methods**

**Study Site and Field Procedures**

Brown-headed nuthatches are primary cavity-nesting birds that occupy year-round territories in open pine forests of the southeastern United States. I studied the population at Tall Timbers Research Station (TTRS; 30.66° N, 84.22° W) in northern Florida from 2006 to 2010. TTRS encompasses 1,630 ha and is dominated by upland pine habitats consisting primarily of loblolly (*Pinus taeda*) and shortleaf (*P. echinata*) pines. Dominant pine trees are >75 yrs old, and the pine forests have an open vegetation structure that is maintained by frequent (≤2 yrs interval) prescribed fires. The study area is also surrounded by multiple large (>1,000 ha) properties that contain extensive areas of fire-maintained upland pines and provide suitable habitat for brown-
headed nuthatches. Nests are placed in well-decayed snags, and groups generally excavated fresh cavities each breeding season (Slater et al. 2013).

Methods for monitoring nests and banding adults and nestlings (federal permit 22446, state permit WB04060a, IACUC# VE-2004-01) follow Cox and Slater (2007). Nests were located using area searches initiated in February of each year. I banded nestlings \( n = 406 \) with numbered, metal USGS bands 10 – 14 days post hatching. I accessed nestlings following the technique of Ibarzabal and Tremblay (2006). I used a drill and small say to remove the face of the nest cavity and, once nestlings were banded, replaced the face using wood putty, staples, and duct tape. The number of nestlings banded per nest was used to estimate nest productivity. Re-nesting after a nest failure is uncommon in this species (<10%; Lloyd et al. 2009).

Adults \( n = 248 \) were captured using mist nets placed near nests or by luring individuals into nets using recorded vocalizations of conspecifics or Eastern Screech-Owl \( (Otus asio) \). In addition to a metal band, adults were also marked with a unique combination of colored bands. I assigned sex to those individuals that were not genotyped using adult behavior and sex-specific vocalizations (Norris 1958) as well as the presence/absence of a brood patch, and wing measurements (Cox and Slater 2007). Field procedures used to determine sex provided high accuracy (≥97%; \( n = 34 \)) when compared to sex determinations based on molecular procedures (Haas et al. 2010).

I estimated the size of breeding groups each year by observing the number of adults engaged in nesting activities (i.e., territory defense, cavity excavation, nest building, incubation, or providing food for nestlings or incubating females) on at least
four occasions during the breeding season. Each visit lasted at least 15 min, and recorded conspecific vocalizations were played infrequently near nests to solicit territorial responses. I determined color-band combinations via observation using spotting scopes and binoculars, and I used vocalizations, copulation, incubation, and dominance behaviors (Norris 1958) to assign status either as a breeding adult or as a helper (Cox and Slater 2007). Based on video recordings collected when adults were feeding nestlings in 2006 and 2007 \((n = 35\) nests; 6 hrs per nest; L. Loke, unpubl. data), these procedures provided an accurate estimate of group composition. Kinship relationships inferred from field observations were based on the putative breeding pair and young marked at each nest.

**DNA Sampling**

I sampled family groups from 2006 to 2010 \((n = 59\) clutches representing 50 family groups, 142 adults and 237 nestlings). Group sizes for these samples ranged from two (i.e., pairs) to five (i.e., cooperative groups) adults. I obtained a blood sample \((20 – 40 \, \mu L)\) from the brachial vein of adult birds when they were captured for banding and stored the samples in 1mL of lysis buffer \((0.1\) M Tris-HCl, pH 8.0, 0.1 M EDTA, 0.01 M NaCl, 1% SDS) at room temperature. At the time nestlings were banded, I pulled 4 – 9 feathers for genetic analyses before returning nestlings to the nest. This was less invasive than taking blood samples and growing feathers had a more than sufficient amount of DNA. Feathers were stored in 1 mL of 100% EtOH at -20 °C.

**Laboratory Methods**

I extracted genomic DNA from all blood and feather samples following the standard protocol outlined for tissues in the PUREGENE® DNA Purification Kit (Gentra
Systems, Minneapolis, MN, USA). All samples were quantified using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

I determined the sexes of all individuals via PCR using the primers described in Fridolfsson and Ellegren (1999). All PCR amplifications were performed in 10 µL volumes on a MasterCycler® thermal cycler (Eppendorf AG, Hamburg, Germany). I amplified nine di-nucleotide microsatellite markers specific to the brown-headed nuthatch: Spu4-C6, Spu4-E7, SpuA6, SpuE19, SpuL4-3, SpuL4-30, SpuL5-22, and SpuL6-16 (Haas et al. 2009). Final concentrations were: 1x PCR buffer (New England BioLabs Inc., Ipswich, MA, USA), 1.5 – 3.0 mM Mg++, 0.2 mM each dNTP, 0.2 µM each primer, 0.02 U/µL Taq DNA polymerase (New England BioLabs Inc., Ipswich, MA, USA) and 0.8 ng/µL genomic DNA. Reaction conditions followed those found in Haas et al. (2009) with the exception of SpuE19 and SpuL5-22. These two loci were amplified using a primer with an M13 tail and a universal fluorescently labeled M13 primer (Schuelke 2000). For these two loci, final concentrations were 1x PCR buffer, 2.5 mM Mg++, 0.2 mM each dNTP, 0.02 µM forward primer with M13 tail at the 5’ end, 0.2 µM each reverse primer and labeled M13 primer, 0.02 U/µL Taq DNA polymerase, and 0.8 ng/µL genomic DNA. The loci were then amplified via touchdown PCR. The reaction conditions were: 5 cycles of 94 °C for 30 sec, 62 °C (SpuE19) or 52 °C (SpuL5-22) for 90 sec, 72 °C for 30 sec, followed by 5 cycles with annealing temperature 60 °C (SpuE19) or 50 °C (SpuL5-22), and finally 25 cycles at annealing temperature 48 °C. There was a final 30 min extension at 72 °C. I diluted the final PCR products and multiplexed them into three groups as follows: Pool A (SpuA6 + SpuE19 + Spu4-E7), Pool B (SpuL5-22 + SpuL5-6 + SpuL6-16), and Pool C (Spu4C6 + SpuL4-30 + SpuL4-
Allele sizes were determined using an ABI Prism® 3100 Genetic Analyzer (Applied Biosystems, Grand Island, NY, USA) and the GeneScan™ 500 ROX size standard (Applied Biosystems, Grand Island, NY, USA), and analyzed with GeneMarker version 1.6 (SoftGenetics LLC, State College, PA, USA).

**Microsatellite Variation**

Nine microsatellite loci were characterized using 98 breeding adults (excludes helpers). I evaluated microsatellite data for the presence of null alleles and genotyping errors using MICRO-CHECKER (Van Oosterhout et al. 2004). I tested deviations from Hardy-Weinberg expectations and linkage disequilibrium between pairs of loci using Fisher’s exact tests in GENEPOP 4.0.10 (Raymond and Rousset 1995, Rousset 2008). There was no evidence for null alleles, although exact tests suggested one locus was out of Hardy-Weinberg equilibrium (HWE) due to heterozygote deficiency ($P < 0.0056$) after a Bonferroni adjustment of $\alpha$ for multiple comparisons. I ran our analyses with that locus included as well as excluded and found only minor differences (all associated with a loss of power that resulted from excluding this locus). Thus, I report results using all nine loci. The loci were highly polymorphic with a mean allele frequency of 21.56 alleles/locus (20 alleles/locus for the eight loci in HWE).

**Analyses of Kinship, Parentage, and Group Characteristics**

To estimate relatedness among adults, I used ML-RELATE (Kalinowski et al. 2006), which calculates a maximum likelihood estimate of relatedness. Paternity analysis was performed using CERVUS 3.0 (Kalinowski et al. 2007). CERVUS uses a likelihood approach to calculate the log of odds (LOD) scores of likely parents. It then provides a $\Delta$LOD score, which is defined as the difference in LOD scores between the top two most likely candidates, to assess confidence on parentage assignments. The
combined exclusion probability was >0.99 when neither parent was known, as well as when one parent was known.

In most cases, I knew the mother because there was only one female present at the nest and egg dumping (brood parasitism) was not observed. Therefore, females were specified as “known parents”. I then ran the CERVUS analysis with all males present at the nest as candidate fathers. Trio mismatches that occur among mother-father-offspring trios ranged from 0 – 3 mismatches. I was conservative, and in cases where there were trio mismatches, if the putative father was identified with 95% confidence, then I accepted the putative father as the actual father (i.e., not a case of extra-pair paternity) of the nestling. For all nestlings that failed to match a male at the nest with 95% confidence (n = 66), I carried out an additional analysis using all adult males sampled in the population as candidate fathers.

Sample sizes from field observations were larger than those used for genetic analysis and therefore can help to describe annual variation in characteristics such as group number, helper sex and age, and the putative relatedness of helpers to breeders. Accordingly, I report some features using mean (± SD) that were based on the proportions observed each year (e.g., proportion of cooperative groups). These values help to show central tendencies and annual variation but were not assessed statistically. Other summary statistics are also presented as mean (± SDs). All genetic samples were included in the field observations.

Results

Group Characteristics

I gathered field observations for an average of 47.0 (± 8.3) nests annually, and marked an average of 41.3 (± 9.6) adults and 71.0 (± 26.9) nestlings during each year
of study. The proportion of cooperative groups averaged 0.23 (± 0.07) among all territories monitored each year and ranged from 0.12 (observed in 2006) to 0.29 (observed in 2008 and 2009). While one cooperative group contained five adults, groups containing a single helper (0.87 ± 0.16) were most common (Table 2-1). The broods averaged 4.3 (± 1.4) nestlings (ranging from one to six), and there was no significant difference between the number of young banded in nests tended by pairs (4.3 ± 1.3) versus the number banded in nests tended by larger groups (4.3 ± 1.4) using a t-test. Similarly, there was no significant difference between the average proportions of successful nests (fledged young) for pairs (0.45 ± 0.22) versus larger groups (0.48 ± 0.14) using a chi-square test (chi-square = 0.078, $P = 0.78$).

**Helper Relatedness**

The expectation for a typical cooperative breeder with helpers at the nest is that helpers are related to at least one member of the breeding pair (Riehl 2013). This was generally what I observed with brown-headed nuthatches. Of the 23 helpers genotyped from 19 nests, nine were related to both parents, seven were related to the father only, three were related to the mother only, and four were unrelated to either parent. All helpers were male with the exception of one female helper who was unrelated to either parent. When more than one helper was at the nest, the helpers were either related to both parents or to the father only, with one nest having helpers in each category.

I initially inferred parent-helper relationships based on field observations of marked individuals for 43 groups ($n = 57$ helpers). Most helpers were male descendants of at least one member of the breeding pair (0.58 ± 0.22); however, other groups consisted of: (1) male helpers not directly related to the breeding pair (0.34 ± 0.21); (2) brothers associated with a female ($n = 1$); and (3) female helpers not directly related to
the breeding pair \((n = 2;\) one was part of a genotyped group). All males that were not
directly related to the breeding pair were assisting neighboring nests after failure of their
own nest. However, recent monitoring has turned up genetically sexed males that
dispersed farther and also assisted other (putatively unrelated) adults (JAC, unpubl.
data). One of the female helpers was observed assisting unrelated adults at a nest 1.6
km from her natal territory.

Based on field observations of marked individuals, most helpers were in their first
year and resulted from delayed dispersal \((n = 35; 0.68 \pm 0.25)\). Of the older helpers, six
were individuals that attempted to breed independently, lost their nests, and afterwards
provided assistance at a neighboring nest, while seven individuals assisted breeders
throughout the breeding season (one of these males appeared to be assisting adults
that were not not his parents).

**Parentage Analysis**

Of the 59 nests I genotyped, 24 (41\%) showed evidence of extra-pair young
based on mismatches with the putative father. I was unable to identify the putative sires
of any of these nestlings. Six of these extra-pair nests were from cooperative groups
while the remaining 18 were from nests that belonged to pairs. Extra-pair young (Figure
2-1) were not more likely to occur in cooperative groups (32\%) than in pairs (45\%)
(Fisher’s exact \(P = 0.40\)). Although I attempted to sample all individuals at a nest, it is
possible that some adults may have been missed. In two cases, nests that were
identified as pairs were likely to be cooperative groups where additional adults were not
observed or missed, since all nestling appeared to have the same father, but none of
the nestlings matched the putative father. The most likely explanation for this is that the
resident breeding male was not observed or sampled. When I exclude these nests from
analyses, I still find no difference in frequency of extra-pair young observed between cooperative and pair nests ($P = 0.40$).

Of the six cooperative groups with extra pair young, helpers did not sire extra-pair young in four of them, as would be expected if helpers were related to the female. Of the remaining two nests, the helpers may have sired one nestling. The helpers in these cases were related to both the attending male ($r = 0.34$ and 0.56) and female ($r = 0.4$ and 0.6). The relatedness of the chick to the helper ($r = 0.72$ and 0.79, respectively) was greater than that expected if the chick was a full-sibling of the helper (expected $r = 0.5$), suggesting that incest may have occurred. In the first case, the attending male could be excluded as the sire with high confidence ($\Delta \text{LOD} = 6.61$), suggesting incest is the most likely explanation. However, in the second case, the confidence is much lower ($\Delta \text{LOD} = 0.687$), and it is not clear if this is actually a case of incest or not. Regardless, the maximum percentage of incestuous nestlings was <1% of all chicks genotyped.

**Discussion**

**Group Structure**

Although group composition varied, brown-headed nuthatches basically fit the pattern of many other cooperatively breeding birds (see table 6.1 in Komdeur 2004); most helpers were male offspring that delayed dispersal and were related to at least one parent. One female helper was a likely case of redirected helping (i.e., individuals that help at a nest after their own nest fails) and not delayed dispersal (based on field observations), though the second female helper observed did appear to be a case of delayed dispersal. Thus, while in general females disperse from their natal territory in their first year, this is not universal. Delayed dispersal of females, though uncommon
among cooperatively breeding birds, has been observed (e.g., Caffrey 1992, Richardson et al. 2002).

Relationship to only one parent could occur in a number of ways. Field observations suggested a small number of helpers were siblings of the breeding male. Siblings may inherit the natal territory when both parents are gone or they may disperse together to form new groups (e.g., Heinsohn et al. 2000). Alternatively, one parent could die or move and the surviving adult could acquire a new mate. The fact that more helpers were related to the breeding male than to the breeding female is consistent with the observation of higher female mortality in brown-headed nuthatches (Cox and Slater 2007), though the same pattern could also arise if females that experience mate loss do not retain the territory as frequently as males that experience mate loss. In the cooperatively breeding red-cockaded woodpecker (*Picoides borealis*), breeding females disperse after death of a mate to avoid inbreeding when their sons inherit the territory (Daniels and Walters 2000). A similar phenomenon may occur in brown-headed nuthatches since instances of breeding females dispersing from established territories are more common than instances where breeding males disperse once achieving breeding status (JAC, unpubl. data). Helpers related only to the breeding female may also be the result of extra-pair copulations. In at least one instance, a helper was unrelated to the male although the helper was banded as a chick in the nest of this male and female the previous year. Nonetheless, most of the patterns of relatedness that we observed within groups were consistent with the hypothesis that groups formed due to delayed dispersal of offspring.
Cases in which helpers were unrelated to either parent were rare. Riehl (2013) proposed three models by which non-kin helping in cooperatively breeding birds might arise. First, unrelated helpers might result from high mortality. Death of one parent, then after re-mating the other original parent, could result in complete turnover of the breeding pair occupying a site but retention of the original helpers. Second, individuals may disperse and join an unrelated breeding pair. Third, dispersing individuals may form a coalition with other unrelated individuals. I propose a fourth scenario, that such cases might also be the result of undetected redirected helping. Redirected helping is known to occur in other cooperatively breeding birds (reviewed in Dickinson and Hatchwell 2004). Typically, these helpers are thought to return to nests of close relatives (Hatchwell et al. 2001), but more analyses will be needed to determine whether redirected helping involves non-relatives in species such as the brown-headed nuthatch. Regardless of how unrelated helpers occur, there can be a range of benefits to group living that may favor unrelated individuals forming groups (reviewed by Riehl 2013).

**Extra-pair Paternity**

Although nearly exclusive monogamy is typical of some cooperatively breeding species (e.g., Quinn et al. 1999, Gienapp and Merilä 2010), I found evidence for extra-pair offspring in brown-headed nuthatch nests in pairs and cooperative groups. Extra-pair offspring were in clutches alongside offspring sired by the dominant male and represented a minority of nestlings in those nests, consistent with other studies of cooperative breeders where extra-pair young have been found (reviewed by Cockburn 2004). If helpers better defended a territory from intruding males, a higher level of extra-pair paternity would be expected in pairs than in cooperative groups, while if helpers
reproduced within groups, then cooperative groups would be expected to have higher extra-pair paternity than pairs. Neither expectation was met in brown-headed nuthatches, suggesting that helpers neither limit nor promote extra-pair copulations.

The proportion of nests with extra-pair fertilizations detected in this study population (0.41) was high compared to the average proportion (0.21 ± 0.13) observed for passerines (see table S1 in Kleven et al. 2009) and most cooperatively breeding birds (≤0.10 but some with >0.8; Westneat and Stewart 2003, Koenig et al. 2009, Cornwallis et al. 2010). I have observed one instance of extra-pair copulation involving marked individuals, though shortly after the copulation the dominant male flew in and briefly chased the extra male. Mate guarding, such as observed here, can limit the number of extra-pair copulations (Dickinson 1997). However, even if mate guarding is an important mechanism in this species, there are times when this may be limited. Females incubate eggs and brood young exclusively, providing males with opportunities for extra-pair copulations while females may have opportunities for extra-pair copulations when the breeding male is busy excavating the nest cavity.

In cases where I sampled breeding pairs in multiple years, I did not always observe extra-pair young, suggesting that different breeding strategies may have been employed in different years. While most nests were either monogamous across years or consistently had extra-pair young across years, in the two cases where females subsequently obtained a new mate, extra-pair young were observed where none had been present in preceding years. In one case, the helper and breeding male switched roles. Whether this was due to shifts in female preference, or due to aspects of the male
that limited his ability to control the situation (e.g., the health declined in one male so the helper took over), is not clear.

Although most nests I sampled were monogamous, the degree of extra-pair fertilization I find suggests that it could be possible to increase or decrease extra-pair fertilizations if factors affecting the behavior were manipulated. Rates of extra-pair fertilizations are thought to vary based on individual traits and the social and environmental conditions that individuals encounter (Westneat and Stewart 2003). Manipulations of social (e.g., sex ratios; Rowe and Arnqvist 2002) and environmental conditions (e.g., food availability; Ortigosa and Rowe 2002) performed on insect taxa have yielded new insights on mating behaviors. Similar manipulations attempted with nuthatches could provide novel information on factors affecting extra-pair fertilizations in complex mating systems.

If extra-pair fertilizations occur in cooperative groups, mechanisms to avoid inbreeding should lead to cases of helpers mating only with unrelated breeders. However, the one or (possibly) two cases where helpers might have sired offspring within the group were cases in which the helpers were related to both members of the breeding pair. While incest overall appears to be rare in this species, it could become a greater concern in smaller, more isolated populations, where the likelihood of encountering related individuals increases. Overall, the breeding system of brown-headed nuthatches is much more complex, with higher levels of extra-pair young, than field studies alone have suggested and warrants increased use of molecular data in combination with expanded experimental studies.
Table 2-1. Number of pair and cooperative nests sampled by year and cooperative group size.

<table>
<thead>
<tr>
<th>Year</th>
<th>No. pair (No. genotyped)</th>
<th>No. cooperative breeding</th>
<th>Size range of cooperative groups (mean group size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>37 (16)</td>
<td>5 (5)</td>
<td>3 – 4 (3.3)</td>
</tr>
<tr>
<td>2007</td>
<td>36 (10)</td>
<td>7 (3)</td>
<td>3 – 5 (3.6)</td>
</tr>
<tr>
<td>2008</td>
<td>39 (6)</td>
<td>16 (4)</td>
<td>3 – 4 (3.2)</td>
</tr>
<tr>
<td>2009(^a)</td>
<td>34 (1)</td>
<td>14 (2)</td>
<td>3 – 4 (3.3)</td>
</tr>
<tr>
<td>2010</td>
<td>26 (7)</td>
<td>10 (5)</td>
<td>3 – 4 (3.1)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>172 (40)</td>
<td>52 (19)</td>
<td>3 – 5 (3.3)</td>
</tr>
</tbody>
</table>

\(^a\) Nesting success was very low
Figure 2-1. The proportion of nests with extra-pair young in pair ($n = 40$) and cooperative ($n = 19$) groups. Proportions of extra pair young were calculated within each group type. There was no significant difference in the proportion of nests with extra-pair young between pairs and cooperative groups (Fisher’s exact $P = 0.40$).
CHAPTER 3
THE EFFECT OF HABITAT FRAGMENTATION ON POPULATION GENETIC STRUCTURE OF BROWN-HEADED NUTHATCHES (SITTA PUSILLA)

Introduction

Habitat loss and fragmentation is considered to be one of the greatest threats to biodiversity today (Pereira et al. 2012). Increasing human population growth, and the resulting changes to the landscape, has compromised ecosystem integrity and caused habitat loss and isolation. Human land use patterns have resulted in small habitat fragments that limit both the number of species and the sizes of populations that can be supported. Consequently, as habitat fragments become increasingly isolated, populations between habitats become more isolated from each other and are at an increased risk of extinction (Gibbs 2001; Reed 2004; Keyghobadi 2007). One of the greatest causes of concern for habitat loss and fragmentation is the loss of genetic variation, which can have several negative effects on populations. First, this may lead to increased homozygosity, which may lead to a buildup of deleterious mutations. Additionally, since genetic variation provides the raw material for natural selection to act upon, a loss of genetic variation results in a loss of evolutionary potential or the ability to adapt to environmental challenges, and increased risk of population extinction (Reed and Frankham 2003; Frankham 2005).

Landscape genetic studies have examined the role landscape plays on population genetic variation and population structure. Standard isolation-by-distance (IBD) (Wright 1943) models have been shown to constrain migration rates between populations such that we would expect increasing genetic distance with increasing geographic distance (Bohonak 1999; Alcaide et al. 2009). However, the spatial configuration of habitat patches (Jenkins et al. 2010; van Strien et al. 2015) and the
composition of the landscape between patches (the matrix) also affects genetic structure (Ricketts 2001; Zhu et al. 2016). Furthermore, species’ responses to fragmentation are more complex than simply habitat structure, and species’ mobility should also be taken into account (Jenkins et al. 2010; Amos et al. 2012, 2014) when considering habitat connectivity. Birds, because they can fly, are considered highly mobile and thought to be able to disperse across unsuitable habitat easily. However, this may not be the case for sedentary species and/or habitat specialists, and thus species mobility should be considered in the context of habitat fragmentation (Amos et al. 2014) even for birds.

Brown-headed nuthatches are cooperatively breeding birds endemic to open, mature old growth pine forests of the southeastern US, and a small population in the Bahamas. They maintain year-round territories, are considered to be fairly sedentary, and are hypothesized to be weak fliers with limited dispersal abilities (Harrap 2016). Although this has not been examined rigorously, a study in northern Florida found that most observed dispersal in males occurred within 300m of their natal territory (Cox and Slater 2007). Females were not observed and are presumed to disperse farther. They are primary cavity-nesting birds that excavate new cavities each breeding season, and thus must nest in snags that have decayed sufficiently for nest excavation to be possible. While nesting occurs in dead trees, foraging occurs in live pines. Therefore, they do best in forests that are managed with regular burnings (Slater et al. 2013) that reduce the understory but are not sufficiently hot to destroy suitable snag trees. Habitat loss and degradation through fire suppression and land conversion are thought to have contributed to population declines and localized extinctions throughout their range.
(Slater et al. 2013), particularly in Florida. These findings, combined with the fact that they have not been observed re-colonizing suitable habitat, lends support to the hypothesis that they have limited dispersal abilities (Slater et al. 2013). Therefore, I would predict that brown-headed nuthatches would be particularly sensitive to habitat fragmentation, and in areas of fragmented habitat should exhibit reduced gene flow, greater genetic variation among populations, and increased homozygosity.

Southeastern pines, particularly in Florida, were once continuously distributed (Figure 3-1). However, anthropogenic changes to the habitat have resulted in habitat fragmentation, particularly in southern Florida where development has been greatest. This loss of pine habitats corresponds to regions where some brown-headed nuthatch populations have been extirpated (Robertson and Kushlan 1974). In these areas, assuming dispersal among nuthatches is limited, population differentiation among habitat fragments should be increasing, and I would expect to see strong population structure. In contrast, among populations in northern Florida, where the habitat remains more contiguous, I expect to see weak or no population structure since gene flow should be continuing.

However, there are other factors that may contribute to population structuring and lead to different patterns. Cooperative breeders, such as brown-headed nuthatches, generally disperse shorter distances (Walters 2000) than other similar-sized passerines who do not breed cooperatively, which could result in strong population structure even in the contiguous northern habitat (and conversely, could suggest that fragmentation may have a limited impact in the south). Thus, under this scenario, brown-headed nuthatches would naturally exhibit limited gene flow, regardless of
anthropogenic habitat fragmentation. Alternatively, brown-headed nuthatches may have
greater dispersal abilities, and/or greater willingness to disperse through potentially
unsuitable habitat, than has been assumed, leading to a high degree of connectivity in
both southern and northern Florida.

Metcalf et al. (2008) found no population structure among brown-headed
nuthatches on the mainland using mtDNA, although they did find strong genetic
differentiation between the mainland and Bahamas populations. In contrast, Haas et al.
(2010) found evidence of spatial autocorrelation within one population using
microsatellites, suggesting brown-headed nuthatches may naturally exhibit limited gene
flow. Spatial autocorrelation, which can detect sex-biased dispersal (Banks and Peakall
2012), was higher in males (the philopatric sex) than females supporting the
observations by Cox and Slater (2007) that females disperse farther than males. Thus,
females may be maintaining connectivity among habitat patches and could also explain
the differences between the mtDNA data of Metcalf et al. (2008) and the microsatellite
data of Haas et al. (2010).

I used microsatellites to examine the effect of habitat loss and fragmentation on
genetic structure in brown-headed nuthatches in Florida. Specifically, I asked: (1) do I
see greater population differentiation or structure in the more fragmented southern
Florida pine habitat than in more continuously distributed northern Florida pine habitat,
and (2) do I see greater estimates of inbreeding or loss of heterozygosity in more
isolated populations?
Materials and Methods

Sampling

I collected blood samples from brown-headed nuthatches (N=381) from 12 sites throughout Florida (Table 3-1, Figure 3-1). This included 6 populations in the more contiguous pine forested northern FL, and 7 populations in the more fragmented south. Sampling was conducted using procedures described in Cox and Slater (2007). In essence, I lured nuthatches into mist nets using recorded vocalizations of brown-headed nuthatch and/or eastern screech owl (Megascops asio). I attached a federal Fish and Wildlife Service (FWS) band to the leg of each individual and obtained a blood sample (20-40 µl) from the brachial vein before each bird was released. Blood samples were stored in 1 ml of lysis buffer (0.1M Tris-HCl, pH 8.0, 0.1M EDTA, 0.01M NaCl, 1%SDS) at room temperature.

Laboratory Methods

I extracted genomic DNA from blood following the standard protocol outlined for tissues in the Puregene® DNA Purification Kit (Gentra Systems) and quantified the DNA using either a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) or a Synergy HT Multimode Microplate Reader (BioTek).

I genotyped all samples using 9 polymorphic di-nucleotide microsatellite markers developed specifically for brown-headed nuthatch: Spu4-C6, Spu4-E7, SpuA6, SpuE19, SpuL4-3, SpuL4-30, SpuL5-6, SpuL5-22, and SpuL6-16 (Haas et al. 2009) and protocols described in Han et al. (2015). I ran fragments on either an ABI3100 or an ABI3730xl DNA Analyzer (Applied Biosystems) and used either the GeneScan™ 500 ROX size standard (Applied Biosystems) or the GGF500ROX size standard (Georgia Genomics Facility). I used Geneious R6.1 Microsatellite Plugin (Biomatters) to score
allele sizes. As a control, I included the same 8 samples on every plate to ensure consistent sizing of alleles across plates.

**Genetic Variation and Population Differentiation**

I used MICRO-CHECKER (Van Oosterhout et al. 2004) to check for the presence of null alleles and genotyping errors due to stutter or allelic dropout. I used GENPOP 4.2 (Raymond and Rousset 1995; Rousset 2008) to calculate Hardy-Weinberg exact probabilities, Weir and Cockerham's pairwise $F_{ST}$ (Weir and Cockerham 1984) and test for linkage disequilibrium. Significance of multiple comparisons was assessed using the Benjamani-Hochberg correction for false discovery (Benjamini and Hochberg 1995) implemented in R (R Core Team 2008). Confidence intervals around $F_{ST}$ were calculated using the R package DiveRsity (Keenan et al. 2013). DiveRsity was also used to calculate Jost's $D$ and confidence intervals. Observed and expected heterozygosities ($H_o$ and $H_e$) were calculated using GenAlEx (Peakall and Smouse 2012) and differences in heterozygosities among populations were assessed using a one-way ANOVA. Differences between northern and southern populations were assessed using a one-tailed t-test. I used HP-RARE (Kalinowski 2005) to calculate allelic richness ($A_R$) and private allelic richness ($P_A R$) using rarefaction with a sample size of 48 alleles (24 diploid individuals).

Pairwise genetic distances between populations was calculated using GenAlEx. These distances were used to conduct a principal coordinate analysis (PCoA) to visualize genetic clusters and test for isolation-by-distance (IBD) using the Mantel test. An AMOVA was also used to test for genetic differentiation.

I used STRUCTURE (Pritchard et al. 2000), a Bayesian clustering program, to infer the number of distinct genetic groups in our data. I used the admixture model with
α inferred for each parameter and correlated allele frequencies. I ran these two different ways. First, I did not include any prior information on the geographic origin of each individual. Each run consisted of 500,000 burn-in followed by 5,000,000 Markov Chain Monte Carlo (MCMC) iterations. For each value of K ranging from 1-12, I performed 20 runs. Because Bayesian analyses can be sensitive to the priors used, I repeated the STRUCTURE analysis, but included sampling locations as prior information (LOCPRIOR=1, LOCISPOP=1). The LOCPRIOR model is especially useful in situations where signals of population structure may be weak, but will also ignore the location information if the ancestry of an individual conflicts with the sampling location (Hubisz et al. 2009). To determine the most likely value of K in both of the STRUCTURE analyses, I used the Evanno method (Evanno et al. 2005) implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). I used CLUMPP (Jakobsson and Rosenberg 2007) to average the Q scores over the 20 runs and visualized the results using DISTRUCT (Rosenberg 2004).

**Results**

The average number of alleles per locus was 12.162 ± 0.573 for the 12 populations ranged from 10.222 to 14.889 (Table 1). MICRO-CHECKER found evidence for null alleles in 11 of the 12 sites, although no locus consistently had null alleles across all those sites. After the Benjamini-Hochberg correction for false discovery 14 of the 108 locus/population combinations showed significant deviations from Hardy-Weinberg. These were distributed across loci and sites with no single population consistently out of Hardy-Weinberg equilibrium. There was no evidence of linkage disequilibrium between pairs of loci both within and across study areas, indicating loci were independent.
Average $H_0$ among all loci and populations was $0.72 \pm 0.014$ (mean ± SE) and ranged from 0.644 to 0.751 (Table 3-1). When looking among the 12 populations, I found no significant differences in $H_0$ ($F = 0.349$, $p = 0.972$), $H_e$ ($F = 0.228$, $p = 0.995$), $A_R$ ($F = 0.554$, $p = 0.861$), or $PA_R$ ($F = 0.966$, $p = 0.482$) even though the more fragmented southern populations might be expected to have lower genetic variation. However, there was a significant positive correlation between $A_R$ and latitude ($r = 0.209$; $p = 0.015$) and $PA_R$ and latitude ($r = 0.186$; $p = 0.027$), supporting my expectation that genetic variation decreased as you moved southward (Figure 3-2). When I combined the northern populations together, and compared them to the southern populations, I also did not find significant differences in $H_0$ ($t = 0.616$, $df = 106$, $p = 0.539$), $H_e$ ($t = 1.185$, $df = 106$, $p = 0.239$), or $PA_R$ ($t = 1.917$, $df = 106$, p-value = 0.058). I did, however, find significant differences in $A_R$ ($t = 2.357$, $df = 106$, $p = 0.020$) between northern and southern populations suggesting southern populations (10.382 ± 4.805) had significantly lower allelic richness than northern (12.858 ± 6.038), as might be expected if fragmentation was causing a loss of genetic variation. Although the $PA_R$ was not significantly different between the north and the south, some of this may be driven by the fact that $PA_R$ had a right skewed distribution. There were no detectable patterns that could explain the outliers (i.e., they were not limited to a single population), however, it is possible that more data (i.e., loci) would resolve that.

**Population Structure**

Pairwise differentiation between sample sites ($F_{ST}$) was low (Table 3-2), indicating little to no genetic differentiation among populations. However, pairwise $F_{ST}$ values among northern populations were lower than those among southern populations and generally included zero, suggesting there is greater genetic differentiation among
southern populations than among northern populations. Although Jost’s D gives higher values (Table 3-2), the 95% confidence intervals among pairwise estimates among northern populations still included zero, which was not the case among southern populations. Overall, there appears to be less genetic differentiation among northern populations than among southern, as predicted. AMOVA results showed that there was no genetic difference among populations, and suggests that Florida brown-headed nuthatches could be considered a single panmictic population.

STRUCTURE analysis without prior information defined three genetic clusters (Figure 3-3A) that corresponded to a primarily northern group (but includes Starkey and Three Lakes, which I designated as southern a priori), a southern group, and a unique population (St. Sebastian). However, when I consider sampling sites as a priori information, the number of clusters increases to K = 7 (Figure 3-3B) and the southern group breaks into multiple populations while the northern populations remain clustered, indicating greater distinctiveness of the southern populations than of the northern. The PCoA analysis agreed with the results from STRUCTURE, and showed that St. Sebastian was distinct from all other populations (Figure 3-4), as well as that northern populations clustered more closely together than southern populations, indicating southern populations may be more genetically distinct from each other than northern populations.

Discussion

I examined population structure in brown-headed nuthatches in both fragmented (southern) and less fragmented (northern) habitats in Florida. Overall, populations showed high genetic diversity, and there was no evidence for loss of heterozygosity in southern populations (Table 3-1). However, there was a trend towards decreasing
variation in the more southern populations (Figure 3-2), with southern populations exhibiting lower allelic richness than northern populations. Thus, while my results did not provide strong support for my initial prediction, that southern populations would exhibit lower genetic variation and homozygosity, there does appear to be some loss of genetic variation in the south relative to the north.

I found little to no genetic differentiation among my Florida populations based on pairwise estimates of $F_{ST}$, suggesting a single, panmictic population. Although there was no significant differentiation among my sample populations, the lowest pairwise estimates of $F_{ST}$ occurred among comparisons between northern populations and the highest estimates were among the southern populations. Indeed, pairwise estimates of Jost’s D, which is thought to be a better measure of population differentiation (JOST 2008), indicate that at least moderate levels of population differentiation were found in southern Florida.

While my STRUCTURE results supported the idea that there is only limited population structure present among the Florida populations of brown-headed nuthatches, both the STRUCTURE and PCoA analyses highlighted one very distinctive population, St. Sebastian. When I consider the current and native (historical) vegetation in Florida (Figure 3-1), this population appears to be naturally, and likely historically, isolated from our other populations by areas of unsuitable nuthatch habitat – grasslands and marsh. This may be a habitat type that brown-headed nuthatches are reluctant to move through, and could represent a natural landscape barrier to connectivity. Unfortunately, populations south of St. Sebastian where habitat appears to be suitable
were extirpated in the mid-1900s (Slater et al. 2013), so sampling of other populations in that region was not possible.

Although two subspecies of brown-headed nuthatches are currently recognized, S. p. pusilla from the mainland and S. p. insularis from the Bahamas, the mainland subspecies have been divided in the past with S. p. caniceps identified as a southern Florida subspecies that only occurs in peninsular Florida (Bangs 1898). However, this subspecies designation divided populations in the panhandle (and the rest of their range) from populations on peninsular Florida, and so divided my northern populations. Therefore, the subspecific designation did not correspond to the two remaining genetic clusters I identified, since I found no differentiation in the north where the subspecific differentiation had been hypothesized.

By including sampling localities as prior information in my STRUCTURE analyses, I was able to identify the presence of weak population structure (Figure 3-3B) in which all of our southern populations were identified as a unique genetic entities (and the northern ones remained clustered). This is also consistent with the PCoA analyses (Figure 3-4), showing greater differentiation among southern populations than among northern populations. Overall, this supports my hypothesis that there may be greater population genetic structure among the fragmented southern Florida populations, though these results were weaker than I had expected.

The fact that I find limited genetic structure in a species that is non-migratory, considered to be a weak flyer, and a poor disperser, even in areas where habitats have become fragmented, seems to contradict the expectation of greater genetic differentiation that would be predicted from these traits. Reduced functional connectivity
(i.e., dispersal) due to the loss of structural connectivity (i.e., habitat fragmentation) should be a bigger issue for sedentary species, or species with limited dispersal (Amos et al. 2012, 2014). Furthermore, cooperatively breeding birds are thought to be especially sensitive to habitat fragmentation, in part because complex social systems are more likely to be disrupted (Fischer and Lindenmayer 2007), but also because sex-biased differences in dispersal may result in the lack of availability of suitable mates (Dale 2001). In a study on the cooperatively breeding brown treecreepers (Cooper and Walters 2002), fragmentation resulted in low female recruitment, possibly due to low functional connectivity of females between habitat fragments. Collectively, this should have led to very strong, rather than very weak, structuring among my southern populations. However, fragmentation has also been found to lead to increased admixture in some species (Harrisson et al. 2013; Husemann et al. 2015), suggesting dispersal may increase as a result of fragmentation. Additionally, dispersal by females may actually be helping to maintain genetic connectivity across fragments (Amos et al. 2014), and this may explain the patterns I observed in the brown-headed nuthatches.

Species in disturbance-dependent habitats, such as the fire dependent southern pine habitats, may undergo episodes of irruptive dispersal due to fire, which would help maintain gene flow. Indeed, reports of vagrants sighted outside their normal range seems to correspond to periods of high habitat disturbance (Renfrow 2003), such as heavy logging, hurricane damage, drought, and even an outbreak of the Southern Pine Beetle. This pattern of weak genetic structure has also been observed in other bird species that are year-round residents of this habitat such as the Bachman’s sparrow (Cerame et al. 2014) and red-cockaded woodpeckers (Haig et al. 1996), suggesting
irruptive dispersals may be an adaptation to disturbance-dependant habitats and could mitigate the effects of habitat fragmentation. Furthermore, brown-headed nuthatches may be less dependent on pine habitat than previously believed. While habitat degradation may be a contributing factor to observed population declines (Sauer et al. 2011), competition for nest sites with eastern bluebirds (*Sialia sialis*) (Houck and Oliver 1954; McNair 1984) may also be a factor. Indeed, Stanback et al. (2011) found that brown-headed nuthatches will readily nest in nest boxes placed in golf courses when bluebirds can be excluded. Although there is still a dependence on pine habitat for foraging, the fact that brown-headed nuthatches were willing to nest in golf courses suggests that they may be more tolerant to disturbed habitat than previously thought.

Alternatively, fragmentation may be having an effect on genetic structure, but there has been insufficient time for population genetic processes such as genetic drift to have produced sufficient differences among populations that I could detect differences with my sampling (Landguth et al. 2010). Indeed, if I imagine starting with a single, large, panmictic population with high genetic variation, it is likely that it may take a while for differences to become apparent without drastic reduction in population sizes, leading to extreme genetic drift. The weak structure found among the southern populations, but not in the northern, could be an indication that fragmentation is beginning to have some genetic consequences. Examining genetic variation in historic populations could provide insight into whether the weak structure in the south that I observed is a historical signal, or whether my observations represent a change from historical patterns (Callens et al. 2011; Athrey et al. 2012).
Table 3-1. Sample sizes (N), average observed (H$_0$) and expected (H$_e$) heterozygosities, number of alleles (N$_A$), allelic richness (A$_R$) and private allelic richness (PA$_R$) for 12 Florida brown-headed nuthatch populations. Populations are listed from North to South.

<table>
<thead>
<tr>
<th>Population</th>
<th>Code</th>
<th>N</th>
<th>H$_0$</th>
<th>H$_e$</th>
<th>N$_A$</th>
<th>A$_R$</th>
<th>PA$_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blackwater</td>
<td>BW</td>
<td>35</td>
<td>0.73</td>
<td>0.80</td>
<td>14.44</td>
<td>12.81</td>
<td>0.38</td>
</tr>
<tr>
<td>Tall Timbers</td>
<td>TT</td>
<td>32</td>
<td>0.75</td>
<td>0.78</td>
<td>14.89</td>
<td>13.38</td>
<td>0.58</td>
</tr>
<tr>
<td>Apalachicola</td>
<td>AP</td>
<td>33</td>
<td>0.73</td>
<td>0.80</td>
<td>14.89</td>
<td>13.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Osceola</td>
<td>OS</td>
<td>31</td>
<td>0.74</td>
<td>0.80</td>
<td>14.44</td>
<td>13.29</td>
<td>0.52</td>
</tr>
<tr>
<td>Goethe</td>
<td>GO</td>
<td>27</td>
<td>0.74</td>
<td>0.80</td>
<td>12.22</td>
<td>11.86</td>
<td>0.30</td>
</tr>
<tr>
<td>Ocala</td>
<td>OC</td>
<td>29</td>
<td>0.73</td>
<td>0.79</td>
<td>13.33</td>
<td>12.46</td>
<td>0.28</td>
</tr>
<tr>
<td>South</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richloam</td>
<td>RL</td>
<td>29</td>
<td>0.75</td>
<td>0.76</td>
<td>12.22</td>
<td>11.53</td>
<td>0.44</td>
</tr>
<tr>
<td>Starkey</td>
<td>ST</td>
<td>33</td>
<td>0.70</td>
<td>0.73</td>
<td>10.56</td>
<td>9.74</td>
<td>0.10</td>
</tr>
<tr>
<td>Three Lakes</td>
<td>TL</td>
<td>36</td>
<td>0.64</td>
<td>0.73</td>
<td>11.56</td>
<td>10.41</td>
<td>0.01</td>
</tr>
<tr>
<td>St. Sebastian</td>
<td>SS</td>
<td>37</td>
<td>0.72</td>
<td>0.76</td>
<td>10.22</td>
<td>9.20</td>
<td>0.23</td>
</tr>
<tr>
<td>Webb</td>
<td>WE</td>
<td>30</td>
<td>0.74</td>
<td>0.79</td>
<td>12.11</td>
<td>11.40</td>
<td>0.21</td>
</tr>
<tr>
<td>Big Cypress</td>
<td>BC</td>
<td>29</td>
<td>0.75</td>
<td>0.78</td>
<td>10.44</td>
<td>10.02</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Table 3-2. Pairwise estimates of $F_{ST}$ (below diagonal) and Jost’s D (above diagonal) for 12 Florida brown-headed nuthatch populations\(^a\). Populations are arranged from north to south. Values in italics indicate the 95% confidence interval includes 0. Values in bold indicate the 95% confidence interval is >0.10 (moderate differentiation).

<table>
<thead>
<tr>
<th></th>
<th>BW</th>
<th>TT</th>
<th>OS</th>
<th>AP</th>
<th>GO</th>
<th>OC</th>
<th>RL</th>
<th>ST</th>
<th>TL</th>
<th>SS</th>
<th>WE</th>
<th>BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW</td>
<td>-</td>
<td>0.008</td>
<td>0.004</td>
<td>0.005</td>
<td>0.002</td>
<td>0.027</td>
<td>0.049</td>
<td>0.067</td>
<td>0.092</td>
<td>0.156</td>
<td>0.115</td>
<td>0.076</td>
</tr>
<tr>
<td>TT</td>
<td>0.005</td>
<td>-</td>
<td>0.018</td>
<td>0.028</td>
<td>0.016</td>
<td>0.016</td>
<td>0.058</td>
<td>0.080</td>
<td>0.078</td>
<td>0.167</td>
<td>0.090</td>
<td>0.076</td>
</tr>
<tr>
<td>OS</td>
<td>-0.001</td>
<td>0.008</td>
<td>-</td>
<td>0.028</td>
<td>0.018</td>
<td>0.051</td>
<td>0.041</td>
<td>0.107</td>
<td>0.118</td>
<td>0.129</td>
<td>0.086</td>
<td>0.076</td>
</tr>
<tr>
<td>AP</td>
<td>0.002</td>
<td>0.012</td>
<td>0.007</td>
<td>-</td>
<td>0.006</td>
<td>0.016</td>
<td>0.079</td>
<td>0.109</td>
<td>0.127</td>
<td>0.124</td>
<td>0.067</td>
<td>0.075</td>
</tr>
<tr>
<td>GO</td>
<td>0.003</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>-</td>
<td>0.015</td>
<td>0.087</td>
<td>0.073</td>
<td>0.083</td>
<td>0.154</td>
<td>0.073</td>
<td>0.092</td>
</tr>
<tr>
<td>OC</td>
<td>0.006</td>
<td>0.007</td>
<td>0.010</td>
<td>0.002</td>
<td>0.004</td>
<td>-</td>
<td>0.078</td>
<td>0.099</td>
<td>0.074</td>
<td>0.164</td>
<td>0.095</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>RL</td>
<td>0.016</td>
<td>0.016</td>
<td>0.009</td>
<td>0.022</td>
<td>0.025</td>
<td>0.019</td>
<td>-</td>
<td>0.138</td>
<td>0.115</td>
<td>0.172</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>0.025</td>
<td>0.025</td>
<td>0.031</td>
<td>0.028</td>
<td>0.030</td>
<td>0.025</td>
<td>0.033</td>
<td>-</td>
<td>0.116</td>
<td>0.224</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>TL</td>
<td>0.027</td>
<td>0.020</td>
<td>0.032</td>
<td>0.033</td>
<td>0.025</td>
<td>0.021</td>
<td>0.032</td>
<td>0.038</td>
<td>-</td>
<td>0.210</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>0.039</td>
<td>0.049</td>
<td>0.034</td>
<td>0.035</td>
<td>0.046</td>
<td>0.048</td>
<td>0.049</td>
<td>0.069</td>
<td>0.077</td>
<td>-</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>WE</td>
<td>0.023</td>
<td>0.019</td>
<td>0.018</td>
<td>0.019</td>
<td>0.017</td>
<td>0.016</td>
<td>0.021</td>
<td>0.032</td>
<td>0.040</td>
<td>0.064</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>0.018</td>
<td>0.018</td>
<td>0.019</td>
<td>0.018</td>
<td>0.024</td>
<td>0.012</td>
<td>0.030</td>
<td>0.043</td>
<td>0.044</td>
<td>0.058</td>
<td>0.023</td>
</tr>
</tbody>
</table>

\(^a\)BW = Blackwater; TT = Tall Timbers; OS = Osceola; AP = Apalachicola; GO = Goethe; OC = Ocala; ST = Starkey; TL = Three Lakes; RL = Richloam; SS = St. Sebastian; WE = Webb; BC = Big Cypress
Figure 3-1. Sampling localities of brown-headed nuthatches.
Figure 3-2. Average allelic richness for 12 sampled brown-headed nuthatch populations. 
A. Observed heterozygosity ($H_o$); B. Expected heterozygosity ($H_e$); C. Allelic richness ($A_R$); D. Private allelic richness ($P_{AR}$).
Figure 3-3. Structure plots for 12 sampled brown-headed nuthatch populations. A. Analysis without prior information with $K = 3$ as the best model. B. Analysis with sampling sites as prior information (USELOCPRIOR) with $K = 7$ as the best model.
Figure 3-4. PCoA analysis of A. 12 brown-headed nuthatch populations; B. 11 brown-headed nuthatch populations (excluding St. Sebastian)
CHAPTER 4
LIMITED POPULATION STRUCTURE AMONG MOUNTAIN POPULATIONS OF THE PYGMY NUTHATCH (SITTA PYGMAEA)

Introduction

Since genetic variation provides the raw material for natural selection to act upon, understanding the processes that shape current patterns of genetic variation among populations is an important question in evolutionary biology. Current patterns of genetic diversity can be complex, and reflect a mixture of historical, environmental, geographic, and demographic factors. Ultimately, these patterns are determined by the interaction of gene flow with other population genetic factors such as genetic drift, selection and mutation (Hartl and Clark 1997).

Quaternary climate changes have helped to shape current patterns of genetic variation (Avise 1998; Hewitt 2004). During glacial maximums, higher latitudes were covered by ice or permafrost. Those species typically found in high-latitude areas likely dispersed to lower latitudes and elevations, and were able to persist in one or a few glacial refugia (reviewed in Shafer et al. 2010). As the ice sheets retreated, and the climate warmed, those species adapted to cooler temperatures would have retreated back up in latitude and elevation (Fall 1997) resulting in populations distributed in habitat fragments among mountains.

Species in these high elevation habitats are essentially ‘stuck’ on an island of suitable habitat surrounded by an ‘ocean’ of very different vegetation at the lower elevations or valleys (Mayr and Diamond 1976), and are likely isolated from other mountain populations. Habitat fragmentation, due to high elevation mountain habitats, should result in reduced gene flow among mountains, resulting in processes such as
genetic drift and local adaptation driving differentiation among populations. The process is thought to be similar to those faced by species experiencing current anthropogenic habitat fragmentation. Thus, studying high elevation habitats, and the effects of the historical fragmentation on population genetic structure, may provide insights into long-term consequences of ongoing anthropogenic fragmentation.

Strong genetic divergence among mountain populations has been shown in a number of taxa including jumping spiders (Maddison and McMahon 2000), montane grasshopper (Knowles 2001), and large mammals such as the American puma (McRae et al. 2005). Collectively, these studies suggest that many species persist after fragmentation, and though there may be limited dispersal among these populations, loss of genetic variation and/or an increase in homozygosity leading to extinction may not always occur following habitat fragmentation and isolation of populations. However, outside of a few examples of species that persist on high elevation mountain habitats, the long-term effects of habitat fragmentation remain poorly understood (but see Watson 2003).

The pygmy nuthatch (*Sitta pygmaea*) is a cooperatively breeding bird found in close association with long needled pine forests in western North America, which are typically found at higher elevations on mountains (Norris 1958; Kingery and Ghalambor 2001). Because of the distribution of suitable pygmy nuthatch habitat, this species has experienced natural habitat fragmentation over evolutionary timescales. These presumed habitat specialists are sedentary birds that occupy territories year-round with low observed dispersal (< 300m) (Kingery and Ghalambor 2001). There are seven subspecies recognized (Clements et al. 2015) based on morphological differences in
size and plumage, whose ranges correspond to major geographic areas of western North America. This high level of phenotypic variation (e.g., subspecific designation) in pygmy nuthatches, low observed dispersal distances, and naturally fragmented distribution, suggests that there should be strong genetic divergence among populations (i.e., greater genetic variation among populations compared to within populations).

In spite of the expectation of strong genetic divergence, using mitochondrial sequence data, Spellman and Klicka (2006) found little differentiation among pygmy nuthatch populations, with some mitochondrial haplotypes shared among geographically widespread populations (e.g., between populations in British Columbia and Arizona). Furthermore, estimates of genetic variation within populations using the mitochondrial data was higher than expected for a species with limited dispersal and narrow habitat requirements (Zuckerberg et al. 2014). Surprisingly, pygmy nuthatches had more genetic variation within populations than other species with broader habitat requirements. However, Zuckerberg et al. (2014) suggest that the high levels of genetic variation observed within pygmy nuthatch populations may be historical because of the use of mitochondrial markers, which evolve more slowly.

In this study, I used microsatellite markers, which should be able to detect more recent population differentiation than the mitochondrial data used by Spellman and Klicka (2006) because they evolve faster, to analyze the genetic variation of pygmy nuthatches and determine if there is more genetic structure present, than previously identified. The samples used were those collected by Spellman and Klicka (2006), allowing a direct comparison of the genetic structure based on mitochondrial versus microsatellite markers.
Methods

Laboratory methods

I used previously extracted DNA from 120 pygmy nuthatch samples from 11 populations (Figure 4-1) used in Spellman and Klicka (2006). All samples were genotyped using 9 polymorphic di-nucleotide microsatellite markers that were developed for the brown-headed nuthatch: Spu4-C6, Spu4-E7, SpuA6, SpuE19, SpuL4-3, SpuL4-30, SpuL5-6, SpuL5-22, and SpuL6-16 (Haas et al. 2010) and protocols described in Han et al. (2015). All PCRs were performed in 10 µL volumes on a MasterCycler® thermal cycler (Eppendorf). Final PCR products were diluted and multiplexed into the three groups described in Han et al. (2015). Fragments were run on an ABI 3730xl DNA Analyzer (Applied Biosystems) and used the GGF500ROX size standard (Georgia Genomics Facility). I scored allele sizes using Geneious R6.1 (Biomatters). I included the same 8 samples on all plates as a control to ensure consistent sizing across plates.

Genetic variation and population differentiation

I used MICRO-CHECKER (Van Oosterhout et al. 2004) to check for the presence of null alleles and genotyping errors due to stutter or allelic dropout. I used GENEPOP 4.2 (Raymond and Rousset 1995; Rousset 2008) to calculate exact probabilities for Hardy-Weinberg proportions, genic differentiation, and to check for linkage disequilibrium. I also used GENEPOP to calculate Weir and Cockerham’s pairwise $F_{ST}$ (Weir and Cockerham 1984) among sites and calculated the confidence intervals around $F_{ST}$ using the R package DiveRsity (Keenan et al. 2013). Significance of multiple comparisons was assessed using the Benjamani-Hochberg correction for false
discovery (Benjamini and Hochberg 1995) implemented in R (R Core Team 2008). I calculated observed and expected heterozygosities ($H_O$ and $H_E$) using GenAlEx (Peakall and Smouse 2012) and assessed differences in heterozygosities among populations using a one-way ANOVA implemented in R, which was also used for other statistical analyses. I used HP-RARE (Kalinowski 2005) to calculate allelic richness ($A_R$) and private allelic richness ($P_A_R$) using rarefaction with a sample size of 16 alleles (8 diploid individuals). This corresponded to the fewest number of individuals sampled at a locus.

I used a Bayesian clustering approach, implemented in STRUCTURE (Pritchard et al. 2000) to infer the number of distinct genetic groups in my data. I used the admixture model with $\alpha$ inferred for each parameter and correlated allele frequencies, and assumed no a priori information with respect to sampling locality (USELOCPRIOR=0). I performed 20 runs for each value of $K$ ranging from 1 – 11 for pygmy nuthatches. Each run consisted of a burn-in of 500,000 followed by 5,000,000 Markov Chain Monte Carlo (MCMC) iterations. I then repeated the analysis using sampling locations as prior information (USELOCPRIOR=1, LOCISPOP=1). The LOCPRIOR model can be useful in detection population structure when the signal is weak (Hubisz et al. 2009). I used the Evanno method (Evanno et al. 2005), implemented in STRUCTURE HARVESTER (Earl and von Holdt 2012), to determine the most likely value of $K$. I averaged the $Q$ scores over the 20 runs using CLUMPP (Jakobsson and Rosenberg 2007) and visualized the results using DISTRUCT (Rosenberg 2004). A Mantel’s test for isolation-by-distance was tested using the
isolation by distance web service (Jensen et al. 2005) and Rousset’s genetic distance (Rousset 1997).

**Results**

The average number of alleles per locus was 7.081 and ranged from 2.273 to 11.273 (Table 4-1). Average $H_O$ among all loci and populations was 0.667 and ranged from 0.611 to 0.765 (Table 4-2). There were no significant differences among populations in various measures of genetic variation: $H_O$ ($F = 0.42, P = 0.934$), $H_E$ ($F = 0.348, P = 0.965$), $A_R$ ($F = 0.932, P = 0.508$), or $P_A_R$ ($F = 1.519, P = 0.146$).

Nevertheless, Nuevo Léon had the lowest observed values for $H_O$, $H_E$, $A_R$, and number of alleles ($N_A$), suggesting that this population had lower genetic variation compared to other populations. Although MICRO-CHECKER detected the possibility of null alleles in three sites, I found no consistent evidence of null alleles across loci or sites. The loci were independent as there was no evidence of linkage disequilibrium between pairs of loci both within and across sites.

Although I expected to see greater genetic variation among populations, that is not what I observed. Pairwise $F_{ST}$ estimates between populations (Table 4-3) generally indicated little to no genetic differentiation among populations, even those from different ends of the range (e.g., samples from British compared to samples from southern California). The exception to this was Nuevo Léon in northern Mexico, which showed moderate to high differentiation with all other localities (range 0.147 to 0.198).

I found $K=3$ genetic clusters (Figure 4-2) in the STRUCTURE analysis without priors. This corresponded to a western group, an eastern group, and a unique population (Nuevo Léon). Interestingly, when I used sampling localities as prior
information, I actually lost genetic clusters (Figure 4-3) and Nuevo Léon is the only population that separates from the other populations. There was no evidence of isolation by distance found ($r = 0.4398$, $p = 0.083$; Figure 4-4).

**Discussion**

In this study, I examined population differentiation in pygmy nuthatches occupying high elevation mountain habitats of the western United States. Although I predicted significant population differentiation due to the fragmented nature of this habitat, this was not the case with pygmy nuthatches. This is in sharp contrast to results shown in other taxa, including other bird species, that occupy habitats in these regions (e.g., McCormack et al. 2008; Knowles 2001; DeChaine and Martin 2005), all of which exhibit strong population differentiation using only mitochondrial markers. Thus, even for highly vagile taxa like birds, species in high elevation habitats appear to exhibit limited or no dispersal among populations.

Instead, I found three genetic clusters using microsatellites (Figure 4-2) that corresponded to a cluster including the Madrean Sky Islands in the southwestern United States and the Transvolcanic Belt in Mexico; a second cluster including southern British Columbia and the mountains along the California coast; and finally, a unique population, Nuevo Léon, that occurs in the Sierra Madre Oriental mountain range in Mexico. While I identified greater structure than had been previously identified using mitochondrial sequence data (Spellman and Klicka 2006), this was still substantially less variation than I had expected given the sedentary nature of pygmy nuthatches, their narrow habitat requirements, and their isolation due naturally fragmented high elevation habitats.
One possible explanation for the low variation among populations I found is that regular gene flow is occurring among sample locations. Although these birds are presumably sedentary (Kingery and Ghalambor 2001), occupying territories year-round, it is possible that individuals may be willing to disperse over long distances and through lowland habitat, at least when seeking a breeding site. This seems unlikely given broad scale movements have not been observed (Ghalambor 2003), although it has not been rigorously studied. Observations of pygmy nuthatches in lower elevations are uncommon (Kingery and Ghalambor 2001), suggesting that these birds do not regularly disperse through lowland habitat. However, without careful study, dispersal among high elevation populations cannot be ruled out.

Another possibility is that pygmy nuthatch populations have not reached genetic equilibrium. Slater et al. (2013) suggested that pygmy nuthatches were located in a single glacial refugium, whose range expanded as the last ice sheets retreated ~12,000 years ago. If there was high genetic diversity in the refuge population, it is possible that there has not been sufficient time to accumulate differences and observe strong population structure even in the absence of gene flow among sampling locations (Keyghobadi et al. 2005; Zuckerberg et al. 2014). Microsatellites, because they evolve faster, should provide greater resolution than mitochondrial markers and indeed, that is what I observed (Frankham et al. 2002). However, even with the increased rate of evolution, the time may have been insufficient to allow strong genetic structure to be evident.

Studies in the white-breasted nuthatch (S. carolinensis), a species with a broader distribution and habitat requirements (Grubb and Parosudov 2008) than pygmy
nuthatches, found genetic structure using 4000 loci (Manthey and Moyle 2015) in populations within the Madrean Sky Islands. This is in contrast with analysis of mitochondrial data in white-breasted nuthatches (Spellman and Klicka 2007), and my data in pygmy nuthatches, which found only a single genetic cluster in this region. Therefore, it is possible that many more loci will be necessary to detect more genetic structure. This would suggest that even if more structure exists among pygmy nuthatch populations, it is relatively weak since only limited (this study) or no (Spellman and Klicka 2006) structure has been found with smaller datasets, and if more structure is to be detected, it will only be with substantially larger data set sizes.

The levels of genetic variation I found within each study locality in the United States was comparable to the diversity found by (Zuckerberg et al. 2014) using mitochondrial data in these populations. Nuevo Léon had lower genetic variation than my other sampled locations, suggesting that this population has been isolated for longer than my other sample sites. However, heterozygosity in Nuevo Léon is still higher than would be expected for isolated populations in fragmented habitat. Further study of pygmy nuthatch populations in Mexico is warranted since very little is known about their status there (Kingery and Ghalambor 2001). If, Nuevo Léon is indeed a unique population, then it should be managed to preserve the level of variation found in this population.

Although I expected to find greater genetic structure in pygmy nuthatches and lower heterozygosity within populations, that was not what I found. Instead, there was very little structure across the range of this species. A similar pattern is seen in populations of brown-headed nuthatches (Chapter 3), which experienced habitat
fragmentation more recently; there was weak population structure and little to no differentiation among populations in southern Florida. This could suggest that the consequences of fragmentation may be less severe for species with high genetic variation. Alternatively, the lack of population differentiation found among most of my sites suggests the possibility that some level of gene flow still occurs. This, combined with my brown-headed nuthatch results, may suggest that nuthatches overall disperse longer distances than has previously been assumed.
Table 4-1. Diversity statistics for 9 microsatellite loci. \( N_A \): number of alleles; \( H_E \): expected heterozygosity; \( H_O \): observed heterozygosity

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>( N_A )</th>
<th>( H_E )</th>
<th>( H_O )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4C6</td>
<td>2.273</td>
<td>0.328</td>
<td>0.356</td>
</tr>
<tr>
<td>L4E7</td>
<td>9.000</td>
<td>0.832</td>
<td>0.802</td>
</tr>
<tr>
<td>A6</td>
<td>3.818</td>
<td>0.406</td>
<td>0.347</td>
</tr>
<tr>
<td>E19</td>
<td>5.909</td>
<td>0.584</td>
<td>0.591</td>
</tr>
<tr>
<td>L43</td>
<td>9.091</td>
<td>0.809</td>
<td>0.771</td>
</tr>
<tr>
<td>L56</td>
<td>10.727</td>
<td>0.855</td>
<td>0.897</td>
</tr>
<tr>
<td>L430</td>
<td>11.273</td>
<td>0.853</td>
<td>0.803</td>
</tr>
<tr>
<td>L522</td>
<td>4.455</td>
<td>0.654</td>
<td>0.705</td>
</tr>
<tr>
<td>L616</td>
<td>7.182</td>
<td>0.796</td>
<td>0.808</td>
</tr>
</tbody>
</table>

Table 4-2. Sample sizes (N), average observed (\( H_O \)) and expected heterozygosities (\( H_E \)), number of alleles (\( N_A \)) average allelic richness (\( A_R \)) and private allelic richness (\( P_A_R \)) for 11 pygmy nuthatch (\( Sitta pygmaea \)) populations. Samples are arranged from North to South.

<table>
<thead>
<tr>
<th>Population</th>
<th>( N )</th>
<th>( H_O )</th>
<th>( H_E )</th>
<th>( N_A )</th>
<th>( A_R )</th>
<th>( P_A_R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>British Columbia</td>
<td>11</td>
<td>0.667</td>
<td>0.695</td>
<td>7.444</td>
<td>6.59</td>
<td>0.38</td>
</tr>
<tr>
<td>Siskiyou</td>
<td>9</td>
<td>0.765</td>
<td>0.719</td>
<td>7.111</td>
<td>6.69</td>
<td>0.40</td>
</tr>
<tr>
<td>Mono</td>
<td>11</td>
<td>0.636</td>
<td>0.661</td>
<td>7.444</td>
<td>6.45</td>
<td>0.41</td>
</tr>
<tr>
<td>Pacific Coast</td>
<td>11</td>
<td>0.616</td>
<td>0.680</td>
<td>6.667</td>
<td>6.03</td>
<td>0.13</td>
</tr>
<tr>
<td>Transverse Ranges</td>
<td>15</td>
<td>0.689</td>
<td>0.691</td>
<td>8.111</td>
<td>6.35</td>
<td>0.16</td>
</tr>
<tr>
<td>San Diego</td>
<td>11</td>
<td>0.687</td>
<td>0.690</td>
<td>6.556</td>
<td>5.73</td>
<td>0.40</td>
</tr>
<tr>
<td>Coconino</td>
<td>13</td>
<td>0.709</td>
<td>0.710</td>
<td>9.222</td>
<td>7.02</td>
<td>0.65</td>
</tr>
<tr>
<td>Chiricahua</td>
<td>12</td>
<td>0.759</td>
<td>0.727</td>
<td>8.111</td>
<td>6.71</td>
<td>1.04</td>
</tr>
<tr>
<td>Sacramento</td>
<td>11</td>
<td>0.667</td>
<td>0.698</td>
<td>7.889</td>
<td>6.73</td>
<td>0.47</td>
</tr>
<tr>
<td>Nuevo Léon</td>
<td>8</td>
<td>0.611</td>
<td>0.576</td>
<td>3.778</td>
<td>3.78</td>
<td>0.37</td>
</tr>
<tr>
<td>Morelos</td>
<td>8</td>
<td>0.625</td>
<td>0.630</td>
<td>5.556</td>
<td>5.56</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Table 4-3. Pairwise estimates of FST (below diagonal) and Jost's D (above diagonal) for 11 pygmy nuthatch (*Sitta pygmaea*) populations\(^a\). Populations are arranged from north to south. Values in bold indicate the 95% confidence interval > 0.10.

<table>
<thead>
<tr>
<th></th>
<th>BR</th>
<th>SI</th>
<th>MN</th>
<th>PC</th>
<th>TR</th>
<th>SD</th>
<th>CO</th>
<th>CH</th>
<th>SA</th>
<th>NL</th>
<th>MR</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR</td>
<td>–</td>
<td>0.052</td>
<td>0</td>
<td>0.056</td>
<td>0.026</td>
<td><strong>0.221</strong></td>
<td>0.033</td>
<td>0.076</td>
<td><strong>0.186</strong></td>
<td>0.366</td>
<td>0.161</td>
</tr>
<tr>
<td>SI</td>
<td>0.020</td>
<td>–</td>
<td>0.014</td>
<td>0.096</td>
<td>0.010</td>
<td><strong>0.198</strong></td>
<td>0.077</td>
<td>0.069</td>
<td>0.095</td>
<td><strong>0.248</strong></td>
<td>0.088</td>
</tr>
<tr>
<td>MN</td>
<td>0.001</td>
<td>0.016</td>
<td>–</td>
<td><strong>0.111</strong></td>
<td>0.003</td>
<td>0.116</td>
<td>0.059</td>
<td>0.045</td>
<td>0.098</td>
<td><strong>0.242</strong></td>
<td>0.126</td>
</tr>
<tr>
<td>PC</td>
<td>0.032</td>
<td>0.036</td>
<td>0.039</td>
<td>–</td>
<td>0.052</td>
<td><strong>0.181</strong></td>
<td>0.059</td>
<td>0.137</td>
<td>0.102</td>
<td><strong>0.434</strong></td>
<td>0.139</td>
</tr>
<tr>
<td>TR</td>
<td>0.017</td>
<td>0.011</td>
<td>0.000</td>
<td>0.030</td>
<td>–</td>
<td>0.062</td>
<td>0.069</td>
<td>0.064</td>
<td>0.095</td>
<td><strong>0.327</strong></td>
<td>0.060</td>
</tr>
<tr>
<td>SD</td>
<td>0.074</td>
<td>0.065</td>
<td>0.053</td>
<td>0.071</td>
<td>0.041</td>
<td>–</td>
<td><strong>0.185</strong></td>
<td>0.149</td>
<td>0.179</td>
<td><strong>0.413</strong></td>
<td>0.158</td>
</tr>
<tr>
<td>CO</td>
<td>0.015</td>
<td>0.033</td>
<td>0.028</td>
<td>0.027</td>
<td>0.031</td>
<td>0.060</td>
<td>–</td>
<td>0.093</td>
<td>0.069</td>
<td><strong>0.382</strong></td>
<td>0.073</td>
</tr>
<tr>
<td>CH</td>
<td>0.029</td>
<td>0.030</td>
<td>0.021</td>
<td>0.037</td>
<td>0.025</td>
<td>0.061</td>
<td>0.028</td>
<td>–</td>
<td>0.033</td>
<td><strong>0.432</strong></td>
<td>0.111</td>
</tr>
<tr>
<td>SA</td>
<td>0.046</td>
<td>0.037</td>
<td>0.031</td>
<td>0.031</td>
<td>0.028</td>
<td>0.074</td>
<td>0.028</td>
<td>0.015</td>
<td>–</td>
<td><strong>0.282</strong></td>
<td>0.075</td>
</tr>
<tr>
<td>NL</td>
<td><strong>0.165</strong></td>
<td><strong>0.153</strong></td>
<td><strong>0.168</strong></td>
<td><strong>0.198</strong></td>
<td><strong>0.171</strong></td>
<td><strong>0.186</strong></td>
<td><strong>0.163</strong></td>
<td><strong>0.168</strong></td>
<td><strong>0.147</strong></td>
<td>–</td>
<td>0.388</td>
</tr>
<tr>
<td>MR</td>
<td>0.050</td>
<td>0.053</td>
<td>0.047</td>
<td>0.049</td>
<td>0.032</td>
<td>0.066</td>
<td>0.038</td>
<td>0.045</td>
<td>0.029</td>
<td><strong>0.191</strong></td>
<td>–</td>
</tr>
</tbody>
</table>

\(^a\) BR = British Columbia; SI = Siskiyou; MN = Mono; PC = Pacific Coast; TR = Transverse Ranges; SD = San Diego; CO = Coconino; CH = Chiricahua; SA = Sacramento; NL = Nuevo Léon; MR = Morelos
Figure 4-1. Sampling localities of 11 pygmy nuthatch populations. Shaded areas represent major mountain ranges
Figure 4-2. STRUCTURE results of 11 pygmy nuthatch populations with no priors. A. Delta K showing the most likely value of K = 3 from StructureHarvester. B. STRUCTURE bar plot for K = 3.
Figure 4-3. STRUCTURE results for 11 sampled pygmy nuthatch populations analyzed using prior information (USELOCPRIOR). A. DeltaK plot from StructureHarvester showing K = 2 is the most likely value of K. B. STRUCTURE bar plot for K = 2.
Figure 4-4. Mantel's test for correlation between Rousset's genetic distance and geographic distance ($r = 0.4398$, $p = 0.083$) from 1000 randomizations. The line shown is the RMA regression line.
CHAPTER 5
THE EFFECTS OF LIMITED SAMPLE SIZES ON THE PROGRAM STRUCTURE

Introduction

Understanding how landscape changes affect genetic variation is a key question for population genetics and conservation genetics. One of the first questions asked when starting a new study is, “how many samples per population do I need?” It is commonly believed that large sample sizes are needed with highly polymorphic loci such as microsatellites (Ruzzante 1998) to get precise estimates of genetic distance such as $F_{ST}$. Kalinowski (2005), using simulations, showed this to be the case to a certain point whereupon sampling more individuals has little effect, and that this is determined by the degree of genetic distance or population differentiation. More individuals should be sampled when population differentiation is smaller (~100 individuals when $F_{ST} = 0.01$) than is necessary when the amount of population differentiation is larger (~20 individuals when $F_{ST} = 0.05$). However, precision around estimates of $F_{ST}$ (or other distance measures) can also be improved by increasing the number of loci sampled (Kalinowski 2005) requiring fewer sampled individuals.

Estimates of genetic variation (e.g., allelic richness, heterozygosity) are other genetic measures used to determine distinctiveness of populations. However, sample sizes can bias these estimates (Petit et al. 1998). As a way around this problem, rarefaction to the smallest sample size is typically used (Kalinowski 2004). This should give a more accurate estimate of the total number of alleles present in a population. However, more recent empirical studies focused on microsatellite data (Pruett and Winker 2008; Hale et al. 2012) have shifted the focus towards estimating informative allele frequencies (i.e., allele frequencies representative of the population) rather than...
trying to get estimates of total number of alleles. This is driven, in part, by the fact that very rare alleles are less likely to be due to shared ancestry or recent gene flow and therefore, are uninformative for most population genetic studies (Hartl and Clark 1997). Instead, 25 – 30 individuals per population are enough to get accurate estimates of genetic variation (Pruett and Winker 2008; Hale et al. 2012), but that standard errors for these estimates (e.g., heterozygosity, allelic richness) should also be reported.

Conservation managers rely on accurate estimates of genetic variation to make policy recommendations, yet sampling is often constrained in the study of rare or endangered species. Large numbers of samples may be unavailable simply given the fact that population sizes are so small. Although sampling 25-30 individuals is an improvement on the 100 individuals (Ruzzante 1998; Kalinowski 2005) previously thought to be necessary, this may still be too high for some studies. Even for species with larger population sizes, obtaining samples can still be challenging if, for example, the organisms inhabit remote and inaccessible locations. Moreover, the number of available microsatellite loci are generally low compared to those available in model systems, so it is not always possible to compensate for small numbers of samples by greatly increasing numbers of loci. So, what happens when sampling is limited? Pruett and Winker (2008) found that even for smaller samples (5 – 10 individuals), accurate estimates can be found by including the sampling error, as the error range (which is larger in smaller samples) will include the true value. However, Hale et al. (2012) recommends avoiding sample sizes lower than 20 individuals because the error range is larger, unless more loci are added.
Another commonly used method of identifying genetically distinctive populations is to use individual Bayesian clustering methods such as STRUCTURE (Pritchard et al. 2000) that group individual genotypes together into unique genetic clusters. STRUCTURE has been found to perform well even in cases of weak population structure (Latch et al. 2006; Hubisz et al. 2009), but it is not known how the software performs with limited sampling. In this study, I assessed the impact of limited sampling on conclusions of population genetic structure.

**Methods**

To explore the effect of sample size on my ability to identify population structure, I used data previously obtained from 381 brown-headed nuthatches (Chapter 3) to generate 100 new datasets by randomly subsampling 10 individuals from each of the 12 populations. I then used STRUCTURE (Pritchard et al. 2000) to infer the number of distinct genetic groups assuming no *a priori* information, as well as using the LOCPRIOR model with sampling locality as *a priori* information. For these analyses I used a burn-in of 100,000 and 1,000,000 MCMC iterations. For each value of K ranging from 1 – 12, I performed 10 replicates per dataset. I determined the most likely value of K using the Evanno et al. (2005) method, implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). Q scores over the 10 replicates were averaged in CLUMPP (Jakobsson and Rosenberg 2007) and visualized using DISTRUCT (Rosenberg 2004).

**Results**

The STRUCTURE analyses on my jackknifed brown-headed nuthatch data varied based on evaluation of the most likely K using the Evanno method, which is the most commonly used approach for identifying the number of genetic clusters. The
average likely value of $K$ estimated from STRUCTURE HARVESTER was $2.65 \pm 0.92$ without priors and $3.13 \pm 1.68$ with prior information (Figure 5-1). However, if I examine the STRUCTURE bar plots from each of the jackknifed datasets, the most likely value is $K = 1$ in analyses without prior information, since there appears to be no distinct clusters apparent in the bar plots. Because the Evanno method is based on the rate of change of the log probability of the data between two successive $K$ values, $K = 1$ cannot be evaluated. The bar plots from the analysis using prior information is much more variable but usually picks up the distinctiveness of the St. Sebastian population. This is in contrast to the results found when using the full set of brown-headed nuthatch data (where I sampled 25 - 35 individuals per population), in which $K = 3$ without priors and $K=7$ with priors (Chapter 3).

**Discussion**

Although one might be able to obtain reasonable estimates of unbiased heterozygosity with 10 individuals per sample, sampling 25 – 30 individuals improves these estimates (Pruett and Winker 2008; Hale et al. 2012). However, the consequences of sample sizes on estimates of genetic structure have not been previously assessed. I found that sampling only a small number of individuals in a population (10, rather than the recommended 25-30) results in underestimating the level of genetic structure present in the population. Given that the St. Sebastian population is quite distinctive (Chapter 3), the failure to identify this under most conditions emphasized the low power such limited samples have. The use of sampling locality as prior information improved estimates only marginally but still failed to equal the number of populations ($K = 3$) identified without using prior information when all samples were
used. This small number of samples also failed to identify instances where weak population structure was present.

Obtaining sufficient sample sizes may not always be possible, particularly when studying rare or endangered species, or species located in remote areas. In such cases, population assignment methods, such as STRUCTURE, would fail to detect unique genetic clusters. Most studies suggest that when sampling is limited, increasing the number of loci could improve estimates of genetic variation and genetic distance (Ruzzante 1998; Kalinowski 2005; Pruett and Winker 2008; Hale et al. 2012). It remains to be seen, however, if the use of new sequencing technologies, that can generate thousands of loci, could resolve the sampling issue for genetic clustering methods.

How does this relate to the results of Chapter 4, where we observed lower structure in pygmy nuthatches than we expected? Although the results of the jackknifing suggest that some of my conclusions on the pygmy nuthatches may be an artifact of the limited sampling, the use of sampling location as a prior should have identified strong structure (since St. Sebastian was generally identified as distinct under those conditions). Thus, given what has been found in other studies of high elevation mountain species, and the putative sedentary lifestyle of pygmy nuthatches, I would still have predicted higher levels of genetic structure (separating into many different populations, rather than just 2-3) in pygmy nuthatches than was found in my study (Chapter 4).
Figure 5-1. Proportion of jackknifed datasets for each value of K analyzed without prior information (green) and with priors (blue)
CHAPTER 6
CONCLUSIONS

Understanding the processes that shape patterns of population genetic variation is important in evolutionary biology and conservation biology. Current patterns of genetic variation reflect a mixture of historical, environmental, and behavioral factors, and understanding that variation can help elucidate the processes that may have led to it. In this study I characterized the mating system in a population of brown-headed nuthatches at Tall Timbers Research Station in Tallahassee, FL. Then I examined the effect of habitat fragmentation in brown-headed nuthatches (affected by anthropogenic fragmentation) and their sister species, the pygmy nuthatch (affected by natural fragmentation). Finally, the role of sample sizes on estimates of genetic structure was examined using subsampling of the brown-headed nuthatch data.

My results show that brown-headed nuthatches have a complex mating system. Auxiliary adults were primarily male offspring related to one or both parents at the nest, though some were unrelated to either parent. I also found a high level of extra-group paternity in both pair and cooperative groups, although in all cases the extra-pair offspring occurred in nests where the breeding male also sired offspring. Auxiliaries generally did not contribute genetically to the nest they attended (though it is possible at least some of these males contributed genetically to other nests). I also found at least one case of incest (and possibly two), suggesting that mechanisms to avoid inbreeding are imperfect. This becomes a greater conservation concern in fragmented habitats where the potential to mate with a relative increases. Although cooperative breeders should have a lower effective population sizes since some adults forgo independent reproduction, the observed extra-pair paternity in brown-headed nuthatches should
increase effective population sizes, particularly if auxiliaries sired some of the extra-pair offspring. Overall, genetic variation in brown-headed nuthatches is likely to be higher than would be expected for a cooperatively breeding bird.

I found little genetic differentiation among the brown-headed nuthatch populations. However, as expected, I found greater genetic structure in southern Florida populations, which have been more affected by fragmentation, than northern Florida, where there is largely contiguous suitable habitat. The high degree of admixture seen in the STRUCTURE bar plots suggests that gene flow may still be occurring among fragmented populations in southern Florida. Although habitat fragmentation is expected to decrease dispersal distances (Walters et al. 2004), this may not be the case for species in disturbance-dependent habitats, such as the fire-dependent pine forests inhabited by both brown-headed and pygmy nuthatches. Instead, species in these habitats may be capable of large dispersal distances (Cerame et al. 2014; Renfrow 2003) to allow dispersal from burning or recently burned habitats that may not be suitable. This may also explain the low genetic structure found in the pygmy nuthatches.

In addition, both species have high levels of genetic variation, so an alternative explanation is that populations have not reached equilibrium after fragmentation (Landguth et al. 2010). While this may be likely for brown-headed nuthatches (where fragmentation is recent), it would seem less likely for pygmy nuthatches given the timescale of fragmentation for this species. Studies of other taxa in this region exhibit strong population differentiation, including birds with broader habitat preferences than pygmy nuthatches (e.g., DeChaine and Martin 2005; Knowles 2001; McCormack et al. 2008). However, Zuckerberg et al. (2014) suggest that effective population sizes for this
species following population expansion after glaciation was large, which would slow down the process of reaching genetic equilibrium. With the current data, it is difficult to evaluate whether fragmentation has had an effect in brown-headed nuthatches, or if the weak structure that was observed is due to historical factors, which would require greater historical structuring in south Florida than in the north. The use of historical samples from museums could help address this.

Alternatively, extra-pair paternity, as was observed in brown-headed nuthatches, may be contributing to the high levels of genetic variation observed. Other studies have reported a positive correlation between genetic diversity and extra-pair paternity (Petrie and Kempenaers 1998; Griffith et al. 1999; Gohil et al. 2013, but see Lie et al. 2015). Thus, extra-pair paternity may be helping to maintain high levels of genetic variation within populations, and helping to counter the effects of fragmentation and small population sizes. Although the genetic mating system of pygmy nuthatches remains to be studied, it is also a cooperatively breeding species and its close relationship with the brown-headed nuthatch may indicate it has a similar genetic mating system.

I also found that limited sampling of individuals per population resulted in underestimating the level of population structure present. The use of sampling locality a priori on STRUCTURE analyses only marginally improved the analysis, but still failed to identify the number of genetic clusters identified using the complete brown-headed nuthatch data set. This is a concern for the study of rare or endangered species where samples are naturally limited, as well as species in remote or inaccessible locations, where sampling is often constrained. One possible solution is to increase the number of loci sampled (Pruett and Winker 2008; Hale 2012). It is now possible to get thousands
of loci with new sequencing technologies which should help with this problem. I am currently using ddRADSeq to examine this issue.
LIST OF REFERENCES


Robertson, W.B., Kushlan, J.A. 1974. The southern Florida avifauna.


BIOGRAPHICAL SKETCH

Kin-Lan Han graduated from Cornell University with a Bachelor of Science in Animal Sciences. She earned a Master of Science degree in Behavior, Ecology, Evolution and Systematics at the University of Maryland – College Park before joining the Doctor of Philosophy program in zoology at the University of Florida. She received her Ph.D. from the University of Florida in 2016.