THE POPULATION HISTORY OF THE DOWNY WOODPECKER (*PICOIDES PUBESCENS*) IN NORTH AMERICA: INSIGHTS FROM GENETICS, ECOLOGICAL NICHE MODELING AND BIOACOUSTICS

PAULO C. PULGARÍN-RESTREPO
BSc. Biology/Zoology, University of Antioquia, 2006

A Thesis

Submitted to the School of Graduate Studies

of the University of Lethbridge

in Partial Fulfilment of the

Requirements for the Degree

[MASTER OF SCIENCE]

Department of Biological Sciences

University of Lethbridge

LETHBRIDGE, ALBERTA, CANADA

ABSTRACT

The last Quaternary ice age strongly influenced the distribution of most plants and animals. Here I used genetics, ecological niche modeling and bioacoustics to understand the possible historical patterns behind the current distribution of the Downy Woodpecker (*Picoides pubescens*) in North America. Analyses of mtDNA sequences and seven microsatellites loci suggest low genetic differentiation among populations (a maximum of two genetic groups), however population structure is subtle. Ecological niche modeling suggests several refugia SE of US, and some restricted areas east and west of the Rocky Mountains with ecological suitable conditions for the species at 18-21 kya. The analysis of the pik call suggested no geographic variation in the frequency and temporal variables studied. It is likely that the Downy Woodpecker expanded and colonized northern North America quickly after the LGM from a southern refugium.

ACKNOWLEDGEMENTS

I am indebted to my advisor Dr. Theresa M. Burg for the support during these two years. I also want to thank my committee Dr. Andrew Iwaniuk, Dr. Andy Hurly and Dr. Ken Otter for their feedback and discussion during this project. Dr. Drew Rendall and Dr. Edward Miller also contributed with ideas and encouragement. Brendan Graham, John Hindley, Kim Dohms and Linda Lait for all the help in the laboratory and field, and for the good times over these two years. I owe a lot to Pinar Kocak and Graham family for all the love, support and extraordinary moments during the completion of this program. My family and friends in Colombia also supported me extensively during these years.

The following institutions and individuals provided valuable tissue samples, information and recordings: American Museum of Natural History, Burke Museum, Canadian Museum of Natural History, Field Museum of Natural History, Los Angeles County Museum, Louisiana State University Museum of Zoology, Marjorie Barrick Museum (UNLV), New Brunswick Museum, North Carolina Museum of Natural History, Royal Ontario Museum, Royal Alberta Museum, University of Michigan Museum of Zoology, University of Northern British Columbia. Macaulay Library of Natural Sounds, British Library Sound Archive, Borror Laboratory of Bioacoustics and Xenocanto. Catherine Thexton, Kevin Colvert, Nathan Pieplow, Alen T. Chantier, Martyn Steward and Don Boucher kindly provided recordings from their personal archives.

I am indebted to all the people that helped during the two field seasons, especially Joanne Vinnedge and her family, Natalie Freeman, Robert and Patricia Perry, Felipe Estela, Adriana Sanchez, Ben Cunningham, Adrian Gail, Sunny Castillo and Susana Salazar. NSERC (TMB) and Alberta Innovates Technology Futures (TMB) funded this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A - adenine
AK - Alaska
AOU - American Ornithologists' Union
ATP - adenosine triphosphatase
AUC - area under curve
AR - allelic richness
BC - British Columbia
BCR - British Columbia Revelstoke
bp - base pairs
C - cytosine
CA - California
CAB - central Alberta
CBC - central British Columbia
CO - Colorado
dB - decibels
DC - Washington DC
dNTP - deoxyribonucleotide triphosphate
DelF - bandwidth
DelT - call duration
DFA - discriminant function analysis
DNA - deoxyribonucleic acid
EDTA - ethylenediaminetetraacetic acid
ENM - ecological niche modeling
FL - Florida
F_{ST} - Wright's fixation index
G - guanine
GBIF - global biodiversity information facility
GIS - geographic information system
H - heavy strand of mitochondrial DNA
Hap - haplotype
HCl - hydrochloric acid
H_d - haplotype diversity
He - expected heterozygosity
HiF - maximum frequency
Ho - observed heterozygosity
Hz - hertz
IBD - isolation by distance
ID - Idaho
ILWI - Illinois/Wisconsin
IN - Indiana
K - cluster
km - kilometres
KY - Kentucky
kya - thousand of years ago

L - light strand of mitochondrial DNA
LA - Louisiana
LGM - last glacial maximum
LoF - minimum terminal frequency
MB - Manitoba
MD - Maryland
MANOVA - multivariate analysis of variance
MI - Michigan
MLNS - Macaulay Library of Natural Sounds
MO - Missouri
MT - Montana
mtDNA - mitochondrial DNA
NaCl - sodium chloride
NC - North Carolina
NJ - New Jersey
NL - Newfoundland
NSNB - Nova Scotia/New Brunswick
OH - Ohio
ON - Ontario
OR - Oregon
PA - Pennsylvania
PCR - polymerase chain reaction
Q3 - 3rd quartile of the frequency
Q - assignment probability
r - raggedness index
s - seconds
SAP - shrimp alkaline phosphatase
SC - South Carolina
SDS - sodium dodecyl sulfate
SSD - sum of squared deviations
T - thymine
taq - DNA taq polymerase
TE - Tris-HCl EDTA buffer
TN - Tennessee
Tris-HCl - tris(hydroxymethyl)aminomethane
U - units
UT - Utah
US - United States of America
VI - Vancouver Island
WA - Washington
WV- Western Virginia
π - nucleotide diversity
°C - degrees Celsius

CHAPTER ONE

GENERAL INTRODUCTION

1.1 BACKGROUND

Natural populations of animals, plants and other organisms are historically and geographically intertwined. The history of the biota is strongly influenced by the history of the landscape. For example, the emergence of mountain ranges, river basins, deserts, and global-scale climatic phenomena such as the Quaternary glaciations have influenced the demographic dynamics of many taxa, and promoted speciation (Hewitt 1996, Avise 2000). Studying the patterns, mechanisms and processes behind these past events is vital to understanding how biological diversity is generated and maintained (Hey & Machado 2003, Cracraft & Donoghue 2004).

During the last part of the Pleistocene, 18-21 thousand years ago (kya), the retreat of the ice sheets was a key event promoting the expansion and contraction of the geographical ranges of many species of animals and plants. Some species became extinct, some dispersed to new areas, and others persisted in ice-free refugia (Pielou 1991). At the end of the last ice age, the climate and landscape were dramatically different from today; the earth was colder and drier, and ice sheets covered one third of the land (Hewitt 2000). The impact of the ice sheets and the subsequent deglaciation was more extreme in boreal and temperate areas of the world, and less extreme in tropical areas (Pielou 1991, Hewitt 2000).

Europe and North America have been the focus of most studies on post-glacial expansion, and the effects of the last glacial maximum (LGM) on organisms from a geological and biological point of view is better understood. In Europe, the Iberian Peninsula, the Balkans and some places in the Mediterranean were common refuges for alder (*Alnus glutinosa*), bears (*Ursos arctos*), grasshoppers (*Chorthippus parallelus*), hedgehogs (*Erinaceous europeus*), Moor Frogs (*Rana arvalis*) and Tawny Owls (*Strix aluco*), from which these species colonized Northern Europe and Asia following glacial retreat (Hewitt 1999, Brito 2007, Knopp & Merilä 2008).

In North America, the Cordilleran and Laurentide ice sheets were present as far as south as 40 ^{o}N with the Pacific Northwest, Northeast coast, Southwest US and Beringia serving as the major refuges for a diverse array of organisms. Some examples of species that remained in some refugial areas include birds: the Chestnut-backed Chickadee (*Poecile rufescens*), Steller's (*Cyanocitta stelleri*) and Mexican Jays (*Aphelocoma ultramarina*) (Burg et al. 2006, Burg et al. 2005, McCormack et al. 2008); mammals: Northern Flying Squirrel (*Glaucomys sabrinus*), Red-backed Voles (*Clethrionomys gapperi*), American Marten (*Martes americana*); and amphibians and reptiles: Spotted Salamander (*Ambystoma maculatum*) and Western Diamondback Rattlesnake (*Crotalux atrox*) (Waltari et al. 2007). In other words, "every little thing to be found in the ice-covered area of North America must, obviously, be descended from ancestors that, at glacial maximum, were living in an ice-free area" (Pielou 1991).

Despite the knowledge gained from multiple studies over the last two decades, understanding how different lineages and species responded in a spatio-temporal context to the retreat of the ice sheets following the LGM is still in its infancy (Milá et al. 2007). Where were refugia located? How long did it take for organisms to colonize ice-free areas after the glaciers retreated? Even simple questions like these are not easy to answer for many taxa due to the analytical and methodological problems of inferring past events from current information, and the absence of fossil data (Jaramillo-Correa et al. 2009). Nevertheless, studying the response of natural populations to large-scale climatic events is critical not only to understanding current patterns of diversity, but also in predicting the demographic and population trends of ongoing and future climatic changes (Thomas et al. 2004, Wormworth & Mallon 2006).

1.2 METHODOLOGICAL APPROACH

One way to study historical patterns is by using DNA. Before the rise of the molecular markers as a favorite tool for systematics and population genetic studies, most research about past events (such as post-Pleistocene dispersal) relied on phenotypic data from morphology, physiology or behavior (Avise 2004). The introduction of molecular markers in comparative studies permitted direct comparisons of genotypic information derived from nucleic acids and proteins, and increased the amount of data available to reconstruct historical patterns, since a typical genome may contain more than a billion nucleotides (Avise 2004). Molecular markers are a powerful tool because their elemental

components (e.g. nucleotides) are universal for all the species, becoming a yardstick to specific or comparative studies (Avise 2004).

Moreover, it is impressive what genetic data can tell us about historical processes, such as demographic changes, extinction or speciation. In the case of historical demographic changes, recent population expansions are exemplified by a starburst pattern where a single, central haplotype is found in many individuals and most other haplotypes differ from the common haplotype by one or two mutations (Fig 1.1a). Molecular markers can also be used to trace ancestral lineages or population subdivision. In this case, populations from the same glacial refugia are characterized by a collection of similar haplotypes strongly differentiated from haplotypes found in other isolated areas (Fig 1.1b). Accordingly, a population's or species' history is imprinted in their DNA (Rogers 1995).

Two of the most powerful molecular markers used to solve population and demographic questions are mitochondrial DNA (mtDNA) sequences and microsatellite loci. MtDNA analyses were introduced to population genetics in the late 1970s, and completely altered the field of systematics and population genetics, including creating a new field called phylogeography (Avise 2009). Mitochondrial DNA sequences evolve rapidly due to inefficient DNA repair, are uniparentally inherited and exhibit no recombination. These characteristics allow us to reconstruct the history of matrilineal lineages based on haplotypes (Avise 2004). A more powerful analogy comes from the fact that all individuals present in the current population shared a common ancestor, or

coalesced, at various points in time. So individuals from different populations are connected to each other in a genealogical hierarchical system (Avise 2004). Therefore, for the purposes of studying historical process such as demographic expansions, contractions or population differentiation, mtDNA is one of the most widely used and reliable tools to understand past biological processes.

The other type of molecular marker used in this study is microsatellite loci, which are short tandem iterations of a simple sequence motif. The core motifs can be a one to six nucleotide repeat (e.g. dinucleotide AC, tetranucleotide TAGA). Microsatellites are found throughout eukaryote genomes and are bi-parentally inherited (Ellegren 2004). Microsatellites represent one of the most variable types of DNA sequences in the genome. They are hypervariable and the polymorphism is derived mostly from variation in length (e.g. allele size) rather than the information contained in the sequence. In the early 1990s, microsatellites became very popular among population geneticists using them to answer all sorts of biological questions from forensics to gene flow among populations (Tautz et al. 1986, Schlötterer 2004). Despite some pitfalls related to complex mutational processes and PCR artifacts (stutter bands), microsatellites are a powerful tool used to study population structure, recent gene flow and parentage (Jarne & Lagoda 1996).

Another way to study historical patterns is through the use of geographical and climatic information. Specifically by using paleo-reconstructions of the climate and landscape, it is possible to infer habitats and areas where species may have persisted

based on niche conservatism (Wiens & Graham 2005). Niche conservatism suggests that species tend to have similar physiological and climatic requirements over time and, therefore current climatic and habitat characteristics can be used to predict historical distributions. For example, if a species was better adapted to a high temperature regime 100,000 years ago, there is a high probability that the same species is still found in high temperature habitats today (Ackerly 2003). Nonetheless, niche conservatism and ENM should be taken with caution in specific situations, since species niche and ecological characteristics can change quickly over time (Wiens & Graham 2005). The introduction of ecological niche modeling (ENM) methods has provided analytical tools to study how climate changes have influence species' distributions (Waltari et al. 2007). In a few words, ENM uses climate layers from the present and from the past, locality points (lat and long coordinates) and an algorithm (in this case maximum entropy), in order to infer a suitable historical habitat.

A third way to examine gene flow and population connectivity is through acoustic data. Bioacoustics, specifically the study of temporal and frequency characteristics derived from analyzing spectrograms (Podos & Warren 2007), have provided a complementary tool to studying population variation. The use of acoustic characteristics to study topics such as population structure or speciation is based on the fact that some acoustic traits are used in key events such as courtship. Therefore any possible modification of a standard signal can promote reproductive isolation if the signal is not recognized (Baptista & Kroodsma 2001). Modifications of a standard acoustic signal can be driven by historical isolation (such as a Pleistocene refugium), geographic distance or

acoustic habitat specialization (Koetz et al. 2007). Each of these three attributes applies to the Downy Woodpecker, a species that has been influenced by the Pleistocene glaciation; is found from Florida to Alaska, and has a distribution that encompasses several type of habitats (Jackson & Ouellet 2002).

Only a handful of studies have used an integrative approach involving genetic, acoustic data, ecological niche modeling (ENM) to examine patterns of post-glacial expansion and colonization in North American birds (Knowles et al. 2007, McCormack et al. 2008). The use of multiple methods is favored over a single method for several reasons. First, when studying historical biogeography, it is not possible to directly recreate past events. Therefore, combining different independent methods (e.g. ENM, bioacoustics) to generate explanations is more powerful (Hugall et al. 2002). For example, before the 1990s any hypotheses of the possible refugial areas for species at the LGM were based on palynological or fossil data (e.g. Wells 1983). When data were missing, interpretations were limited (Waltari et al. 2007). Two decades later, the combination of the Geographic Information Systems (GIS), locality points and climate information can be used to predict the historical distribution of a species or even its future range (Thomas et al. 2004).

Second, the use of information from independent sources opens a torrent of data that can be integrated with information collected in the present day. The possibility of obtaining locality records from museum specimens, DNA, or vocal recordings (as is the case for this project) are perfect examples. This procedure can be time and cost-effective,

and offers the possibility of answering more complex biological questions. Nonetheless, a disadvantage of combining information from different areas such as vocalizations, genetic data and ecological niche modeling lies in the fact that conclusions are usually correlative explanations lacking a statistical or modeling baseline.

1.3 SPECIES AND AREA OF STUDY

Continental North America has been a climatically and biologically dynamic place during the last two million years, especially after the last part of the Quaternary ice ages, 21 kya, when the polar ice sheets were last at their maximum extent (Hewitt 2000). As previously stated, two immense glaciers covered most of Canada, and parts of the northern US at the LGM, while some areas may have remained ice-free such as the Queen Charlotte Islands (also called Haida Gwaii) and parts of Newfoundland (Pielou 1991). Most plants and animals lived in reduced areas south of the ice sheets, with less suitable conditions; quite different from present-day North America, where ~ 20,000 vascular plants, more than 800 species of birds, and more than 640 species of mammals dominate a diverse landscape from Florida to Alaska (Gaston 2000). The changing conditions during the last thousand years make this continent a natural laboratory to study patterns of colonization, geographic variation of morphological traits, adaptive radiation and extinction (Klicka & Zink 1997, 1999, Weir & Schluter 2004).

This thesis focuses on a member of the North American avifauna, the Downy Woodpecker (*Picoides pubescens*). As with other members of the North American bird

fauna such as Juncos, Warblers and Sparrows (Klicka & Zink 1997, 1999, Milá et al. 2006, 2007), there is preliminary evidence that populations and distribution of this species was strongly influenced during the late Pleistocene (Ball & Avise 1992, Zink 1997). Ball & Avise (1992) used a limited sample size (n = 51) from a smaller geographic area, and a less powerful molecular marker, and found no population genetic structure in the Downy Woodpecker. However no other study has attempted to test a post-glacial expansion hypothesis using two quick-evolving molecular markers and larger geographic sampling (either in numbers and area) for the Downy Woodpecker, as in the present study.

The Downy Woodpecker inhabits a variety of vegetation types including coniferous, deciduous and mixed forest and tall shrubbery (Jackson & Ouellet 2002). Downy Woodpeckers can be found in the wilderness, in city parks, wooded areas, orchards and farm groves. The species is ubiquitously distributed in North America from Alaska to Florida (Fig 1.2). It is a year round resident, non-migratory species, breeding from March to July. The Downy Woodpecker is an obligate cavity nesting species that feeds mostly on insects, many of which are prejudicial to humans (Godfrey 1986). This species has an important role in forest bird communities because it makes cavities that are used by other bird species that cannot excavate their own cavities. The Downy Woodpecker is common in forested areas east of the Rocky Mountains, peaking in numbers in the Mid-Atlantic States and Midwest based on Christmas Bird Counts and Breeding Bird Survey data (Jackson & Ouellet 2002). Some local movements between

and within seasons have been reported, but it is not clear how often and how important these movements are in homogenizing populations (Browning 1995).

The Downy Woodpecker is represented by at least seven subspecies (Fig 1.2), which differ mostly in characteristics such as body mass and plumage coloration. Larger birds inhabit northern latitudes and higher elevations, and smaller birds live further south and at lower elevations (Ouellet 1977, Jackson & Ouellet 2002). Birds in the east are whiter than the ones in the west (e.g. more white spots in the wings). The eastern US and Canada contain the subspecies *medianus* and *pubescens*, characterized by their small size and grayish underparts. In the west, there are more races, some of which intergrade (Jackson & Ouellet 2002). The subspecies *glacialis* and *nelsoni* occur from Alaska, west and east of the Rockies down to British Columbia, *nelsoni* extends further into Alberta and possibly as far east as W Ontario. Finally, *P. p. leucurus* is found from the Kenai Peninsula in Alaska to California and the Great Basin; *P. p. gairdnerii* occurs from SW British Columbia, to western Washington down to NW California; and *P. p. turati* from north central Washington to central Oregon and humid coastal California. The five western subspecies are difficult to differentiate even when using several of the morphological traits and plumage (Browning 1997).

1.4 WHY IS THE DOWNY WOODPECKER A GOOD MODEL SPECIES?

The Downy Woodpecker is a potential model species to study historical biogeography in North America because the current range encompasses an enormous

area with complex climate and active geological history. Moreover, the present day distribution includes areas that were glaciated and ice-free starting 21 kya. As such Downy Woodpeckers are ideal to study important questions on the patterns and processes leading to the current distribution. One of the main aspects that makes this species a good model is its non-migratory behavior. All the evidence suggests that individuals are sedentary throughout the year. Banding data and casual observations indicate the species perform short distant movements between seasons (especially females) and during the winter in search of better foraging habitats (Jackson & Ouellet 2002). Browning (1995) for instance reported recoveries from 32 km up to 1080 km. One would expect that the absence of long distance migration and high site fidelity would result in genetically and vocally hierarchically structured groups across the range. Under these assumptions, a null hypothesis of genetic and vocal differentiation among populations can be tested.

Many long distance migrants with a continental distribution have been shown to be highly philopatric; however, many such as the American Redstart (*Setophaga ruticila*) exhibit reduced population genetic structure (Colbeck et al. 2008). Distinguishing whether or not the lack of population genetic structure is associated with the migratory behavior (e.g. potential for dispersal while moving between wintering and breeding grounds) or the reflection of a demographic expansion after the LGM from a single glacial refugium, adds another layer of complexity in the case of highly migratory species (Colbeck et al. 2008). Other highly migratory species have shown population genetic structure between eastern and western North America, which suggests reduced gene flow following prolonged isolation in two refugia (Zink 1996). As the Downy Woodpecker is

not a migratory species, this minimizes the confounding effects of long distance migration on gene flow reducing the number of phenomena and possible scenarios that could explain the current genetic structure. Finally, studying the patterns of genetic and acoustic variation will help to assess the subspecific status of many populations. Currently, seven subspecies are recognized by the American Ornithologists' Union (Fig 1.2); however, there is no evidence that plumage and morphological variation match genetic or vocal traits in the Downy Woodpecker (Ouellet 1977, Jackson & Ouellet 2002).

1.5 THESIS OVERVIEW

This thesis uses two rapidly-evolving molecular markers, one from the mitochondrial genome and the other from the nuclear genome, to examine range wide patterns of genetic variation. The first marker contains a group of mitochondrial genes ATPase 6 and 8 and a small transfer RNA subunit. Mitochondrial DNA sequences have been extensively used during the last 20 years to study genealogical relationships of matrilineal lineages in relation to geography, and powered the field of phylogeography (Avise 2009). MtDNA is haploid and inherited by the maternal line, hence has a reduced effective population size and does not recombine (Avise 2000). The second type of marker is microsatellites, which are commonly used in the study of population structure and gene flow because of their high levels of polymorphism and high mutation rate (10^{-3} to 10^{-6} per generation) (Schlötterer 2004). Microsatellites are codominant, Mendelian markers with bi-parental inheritance (Jarne & Lagoda 1996). Unlike mtDNA,

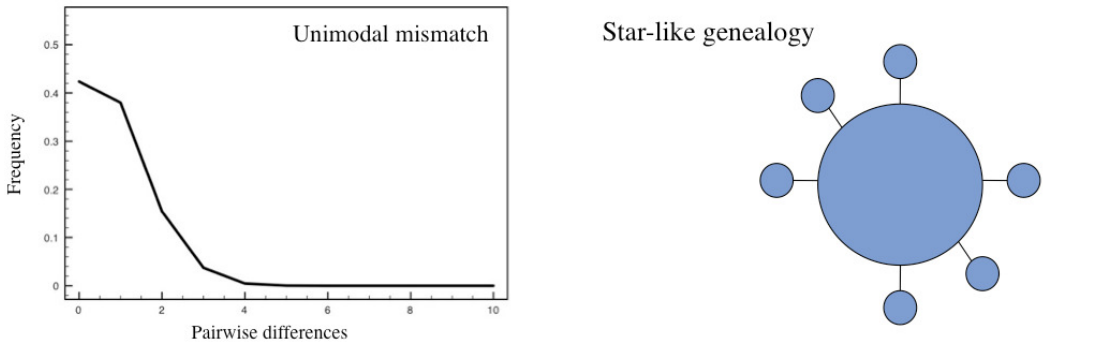
microsatellites do recombine and show recent patterns of gene flow and population connectivity (Nims et al. 2008). Thus, the combination of these two markers offers high resolution and complement each other.

The second component of this thesis uses ecological niche modeling (ENM) to reconstruct the distribution of the Downy Woodpecker in possible refugia at the LGM (18-21 kya). ENM uses point locality data and current and past climatic information to predict suitable areas of habitat based on presence records (Phillips et al. 2006). ENM has become a very powerful tool to predict past, present and future distribution of species, and has been successfully combined with phylogeographic approaches (Hugall et al. 2002).

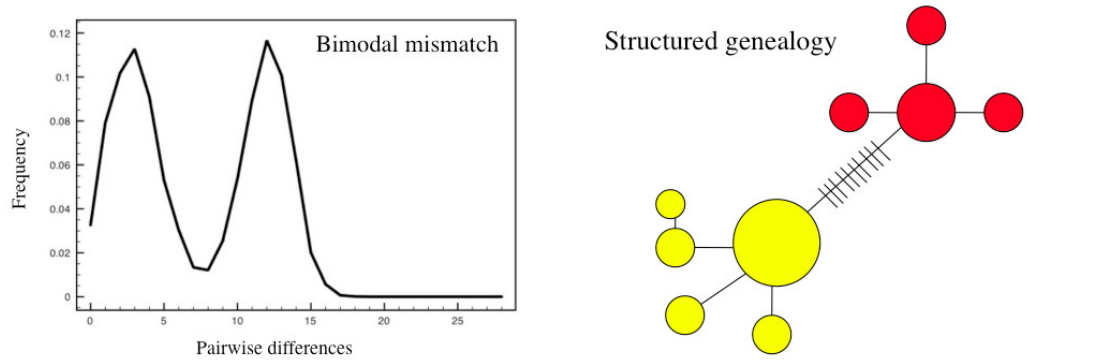
The third approach examines vocalizations. Specifically, a type of vocalization called the pik call is characterized in a geographic context. Vocal analysis is a powerful means of detecting patterns of differentiation caused by the absence of gene flow due to geographic barriers, distance or incipient speciation (Isler et al. 1998, Irwin et al. 2000). The absence of gene flow will produce vocally distinct, diagnosable populations based on frequency and time characteristics. There are two disadvantages of using vocalizations to study population differentiation in a widespread species. First ecological adaptation to habitat characteristics can cause vocal traits to diverge producing a pattern similar to the one in genetically isolated populations (Baptista & Kroodsma 2001) (not possible to test using only recording archives). Second, in many cases a large number of vocalizations

(from different ages and sexes) need to be analyzed to detect differences (Isler et al. 1998).

Using these three approaches (molecular, ENM and acoustics), I construct a comparative framework to study three simple questions that are developed in detail in chapters two and three. First, how this species colonized new areas and expanded its range as the conditions became favorable after the glaciers retreated? Second, where did Downy Woodpeckers live during the LGM? And third, does the acoustic variation in the pik call predict the genetic variation in a geographic context?



a - Hypothetical scenario 1 - genetic signatures



b - Hypothetical scenario 2 - genetic & signatures

Figure 1.1a-b. Predicted mismatch distributions and genealogical (haplotype) networks under two different demographic scenarios. Scenario one denotes a species that has undergone recent population expansion from a single refugium after the LGM. Scenario two represents a species that has remained stable in two isolated refugia (depicted by two different colors) during and after the LGM. Each circle represents a genetic variant (haplotype – size proportional to the number of individuals) and each line a mutational step. Number of mutational steps between red and yellow circles is indicated by the number of cross hatches.

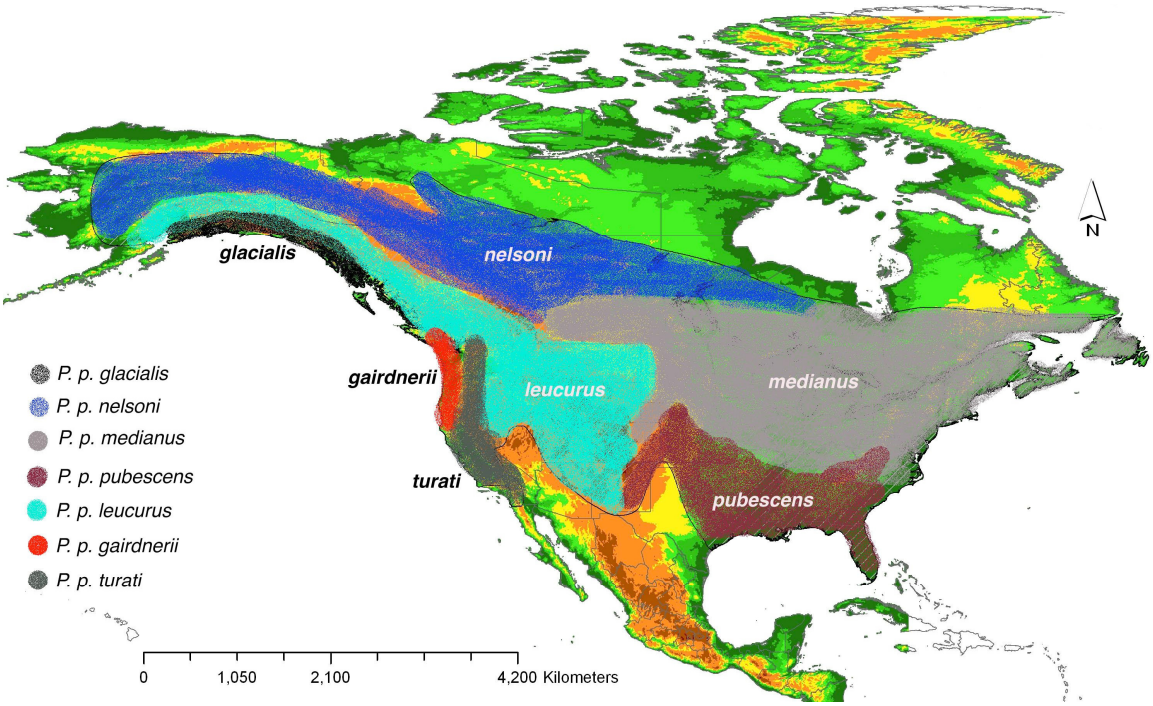


Figure 1.2. General distribution of the seven subspecies of Downy Woodpecker (*Picoides pubescens*) recognized by the American Ornithologists' Union (AOU) (following Jackson & Ouellet 2002). In general individuals are darker in the west compared to east, and smaller in the south compared to north. Several subspecies intergrade at some extent. Background represents elevation.

CHAPTER TWO

RECENT EVOLUTIONARY HISTORY OF THE DOWNY WOODPECKER INFERRED FROM PHYLOGEOGRAPHY, POPULATION GENETICS AND ECOLOGICAL NICHE MODELING

ABSTRACT

The glacial cycles of the Pleistocene have been recognized as an important, large-scale historical process that strongly influenced the demographic and genetic patterns of current Nearctic bird species. I tested hypotheses related to postglacial expansion in the Downy Woodpecker, a common member of the forest bird communities in North America with a continental distribution. DNA sequences from the mitochondrial tRNA-Lys and ATPase 6 and 8 genes, and microsatellite data from seven loci were combined with a paleoecological model of the species' distribution at 21,000 years ago (21 kya). Analyses of 312 individuals from 23 geographic areas suggested little differentiation, shallow genealogical relationships, and limited population structure across the species' range. Contrasting patterns of mtDNA and microsatellites highlight differences in the resolution of the markers. Microsatellites, which are better at detecting recent differences, revealed two geographic groups where populations along the eastern edge of the Rocky Mountains (Montana, Utah, Colorado, and Southern Alberta) were genetically isolated from the rest of the populations. Mitochondrial DNA, an important marker to detect historical patterns, recovered only one group. However, populations in Revelstoke, BC and Idaho contained high haplotype diversity, and in general were characterized by the absence of the most common haplotype. The distribution model from the ENM analysis

revealed the SE US, and east of Rocky Mountains as the most probable areas where the Downy Woodpecker survived 21 kya. Therefore, genetic and ENM suggest a sudden expansion from one or two possible areas south of the ice sheets. This study highlights the importance of combining multilocus data with ecological niche modeling to test biogeographical hypotheses.

Keywords: Colonization, ecological niche modeling, genetic variation, mitochondrial DNA, microsatellites, North America, Picidae, and Pleistocene.

2.1 INTRODUCTION

Studies of the phylogeography and population genetics of Nearctic birds have uncovered an array of different trends (Avise 2009). By examining genetic divergence and population genetic structure, one can infer past patterns. From a historical point of view, the combination of low intraspecific genetic diversity and absence of population genetic structure is one of the patterns found in passerines and other groups (Ball & Avise 1992, Milá et al. 2000, Zink 1997). Some avian groups lack population genetic differentiation despite high morphological and ecological differentiation as shown in *Junco hyemalis* and *J. phaeonotus* (Milá et al. 2007).

A second pattern is high genetic differentiation and geographic structure (Drovetski et al. 2004, Scribner et al. 2003). In North American birds, there is a general east-west split present in both genetic data and in plumage, and this reflects independent evolutionary histories and restricted gene flow (Zink 1997). This second pattern of genetic variation applies to species restricted to specific ecosystems and habitats such as the Chestnut-backed Chickadee (*Poecile rufescens*) in the Mesic temperate forest of Canada and US (Burg et al. 2006), or the Mexican Jay (*Aphelocoma ultramarina*) in the sky islands of SW US and Mexico (McCormack et al. 2008).

Most of the studies mentioned previously used molecular markers with different modes of inheritance, namely, microsatellites and mitochondrial DNA (mtDNA) (Avise 2009, Shafer et al. 2010, Zink 1997). The use of mtDNA sequences has prevailed over

other markers because its matrilineal-haplotypic nature and absence of recombination. Those features are ideal for building genealogies, and making inferences about population history. MtDNA also allows the estimation of divergence times among populations based on a molecular clock approach (Brito & Edwards 2009, Larsson et al. 2009).

The processes behind the genetic patterns can be explained by different modes of inheritance (matrilineal or biparental) of different molecular markers. Genetic patterns are also explained by other factors such as coalescent times (time at which two gene variants or populations split from a common ancestor), mutation rate (how fast the marker evolved), and gene sorting which are all stochastic processes contributing to intrinsic differences (Edwards & Beerli 2000). Other factors such as selective sweeps (positive selection), migration (gene flow), isolation by distance (correlation between genetic and geographic distance), demographic expansion and population bottlenecks, among others, could explain genetic structure and differentiation at intraspecific levels (Zink 1997, Avise 2009).

Large-scale climate changes, such as those occurring during the last Pleistocene glacial cycles, are some of the most influential phenomena responsible for shaping the genetic structure of current populations of North American birds (Klicka & Zink 1997, Hewitt 2000). Specifically, the consequences of the ice sheets retreating after the last glacial maximum (LGM, 18-21 kya), followed by the colonization of the newly ice-free areas produced different genetic patterns (Fig 1.1) as individuals expanded from single or

multiple refugia (Pielou 1990, Burg et al. 2005, Milá et al 2007, Shaffer et al. 2010). One extreme is species with genetically structured populations displaying a gene genealogy (haplotype network) with highly differentiated groups and collections of unique mitochondrial haplotypes or alleles restricted to specific areas, in addition to high nucleotide and haplotype diversities (Burg et al. 2005). This pattern is typical of colonization out of multiple refugia, by a large number of individuals and low levels of gene flow among populations. At the other extreme are species with reduced genetic structure and weakly differentiated populations. The expectations for species with low genetic structure are shallow gene genealogies with one common haplotype and numerous less common haplotypes that differ by a few mutational steps from the main one (Slatkin & Hudson 1991). Populations also exhibit low levels of nucleotide diversity and no or very few private alleles. In most cases, these characteristics are typical of species that have undergone recent, rapid population expansions from a reduced number of individuals, usually from a single refugium (Fig 1.1) (Rogers 1995, Milá et al. 2000, McCormack et al. 2008). Patterns seen can be species-specific as a consequence of different demographic or evolutionary histories (Rogers & Harpending 1992).

In this chapter, I test the following predictions related to postglacial expansion in the Downy Woodpecker.

1. Downy Woodpecker populations expanded and colonized parts of North America after the LGM from a single (or limited) southern refugium. I predicted that: A) populations should display genetic and genealogical signatures typical of a species that has undergone recent expansion such as a star-like genealogy (Fig

1.1a) and unimodal mismatch distributions and B) populations should be characterized by limited genetic structure and extensive gene flow.

2. Downy Woodpecker populations expanded and colonized previously glaciated North America from multiple refugia, either north (Beringia) or south (contiguous US) of the ice sheets. I predict that: A) populations should be characterized by a structured genealogy, with at least two genetically differentiated groups, and a bi-modal mismatch distribution (Fig 1.1b) and B) some populations should contain unique genetic variants (haplotypes or alleles).

To test these predictions, I used a series of population genetic statistics and statistical phylogeographic approaches derived from mitochondrial DNA sequences and microsatellite data. By using both markers I will have two independent perspectives on the patterns of genetic variation among populations, and higher resolution to untangle recent demographic events. Finally, I compare the results from the genetic data with the construction of a distribution model for the Downy Woodpecker at the LGM (21 kya) based on ecological niche modeling.

2.2 METHODS

2.2.1 Sampling - Three hundred and twelve individuals were sampled through: fieldwork, toe-pads from museum bird skins, and tissue loans from different institutions. Most samples were collected during late spring-summer and include individuals from across the majority of the species' range (Fig 2.1) and all seven subspecies recognized by the American Ornithologists' Union (AOU) (Fig 1.2) (Jackson & Ouellet 2002).

2.2.2 DNA extraction - Total genomic DNA was extracted from muscle, blood, feather shaft and toe-pads from bird skins using a modified Chelex extraction (Walsh et al. 1991, Burg et al. 2003). Ten microlitres of blood-ethanol mix, or the equivalent in muscle, feather shaft or toe-pad, was placed in a 1.5 mL microcentrifuge tube and incubated at 50 °C for 30 min to allow the ethanol to evaporate. A 300 µL solution of DNA extraction buffer (0.1 M Tris-HCl pH 8.0, 0.05 M EDTA, 0.2 NaCl and 1% SDS) containing 5% w/v Chelex, 500 micrograms of proteinase K and 250 micrograms RNase was added, and the sample was incubated at 50 °C overnight on a rotating wheel. Samples were vortexed at 10,000 rpm for 60 s, and 200-250 µl supernatant of each sample was transferred to a new 1.5 mL microcentrifuge tube containing 300 µl 5% w/v Chelex in low TE buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA).

2.2.3 Mitochondrial DNA amplification and sequencing - The tRNA-Lys, ATPase 6 and 8 genes were sequenced for 251 individuals. DNA amplification was carried out in 25 µl reactions under the following conditions: ~100 ng DNA, 250 µM dNTP, 2.5 mM $MgCl_2$,

5 pmol of each primer (L8929 - COII 5'- GGMCARTGCTCAGAAATCTGYGG -3' and H9855 - ATP6 5'-ACGTAGGCTTGGATTAKGCTACWGC -3' Sorenson et al. 1999), and 2 units (U) of Taq polymerase. Reactions consisted of: one cycle of 120 s at 94 °C, 45 s at 58 °C, 60 s at 72 °C; 37 cycles of 30 s at 94 °C, 45 s at 58 °C, 60 s at 72 °C; and one cycle of 300 s at 72 °C and 20 s at 4 °C. PCR amplicons were visualized on 0.8% agarose gels (with ethidium bromide) run in 1X TBE buffer. Five microlitres of PCR products were treated with 2 µl of the following mixture: 0.1 units Exonuclease, 0.1 units of SAP, and incubated for 15 minutes at 37 °C followed by 15 minutes at 80 °C to inactivate the enzymes. Sequencing consisted of 10 µl reaction volume: 0.25 µl Big Dye (ver. 3.1), 3.5 µl 2.5X Big Dye buffer, 2.5 pmol of primer (L8929 COII), and approximately 100 ng of purified PCR product. The program consisted of one cycle of 120 s at 96 °C; 25 cycles of 30 s at 96 °C, 15 s at 50 °C and 240 s at 60 °C. PCR products were cleaned using a $NaC_2H_3O_2$ precipitation. Samples were left to incubate at 4 °C for 20 minutes and centrifuged at 13,000 rpm for 20 minutes, followed by two sequential rinses to remove residual salts: first with 100 µl cold 70% ethanol, spun at 13,000 rpm for 5 minutes, and second 35 µl cold 70% ethanol centrifuged at 13,000 rpm for 5 minutes. Ethanol was removed after each step. Samples were left to evaporate in the dark at room temperature for 60 min, re-suspended in 10 µl of hi-di formamide, and sequenced using an ABI3130 sequencer.

2.2.4 Microsatellite amplification and scoring - Two hundred and fifty seven individuals were genotyped at seven microsatellite loci using previously published primers developed for different species of woodpeckers: DIU5 (Ellegren et al. 1999), DMA2, DMD9,

DMD111, DMC115, DMC118 (Vila et al. 2008), and Ptri3 (Välimäki et al. 2008). All 'forward' primers were modified with the addition of a M13 sequence on the 5' end to allow for the incorporation of a fluorescently labeled M13 primer during PCR. The forward primer of DMC115 (5'-TGTCAGAGATGGTTCATGGGTGCACT-3') from Vila et al. (2008) and Ptri3 (5'-GCAAAAGCCAGTTCCTGTGCATGG-3') from Välimäki et al. (2008) were modified to improve PCR amplifications. All loci were amplified using a two-step procedure that consisted of: one cycle of 120 s at 94 °C, 45 s at a specific annealing temperature, and 60 s at 72 °C; seven cycles of 60 s at 94 °C, 30 s at a specific annealing temperature, 45 s at 72 °C. Annealing temperatures were 45 and 48 °C for DMA2, DMD9, DMD118; 48 and 50 °C for DMC111; 50 and 52 °C for: DIU5, DMC115 and Ptri3. DNA amplification was carried out in 10 µl reactions under the following conditions: ~200 ng DNA, 100 µM dNTP, 2.5 mM (for DMC111, DMD118, DMA2, DIU5, Ptri3) and 1.5 mM (for DMD9, DMC115) $MgCl_2$, 5 pmol of each primer (and M13 primer) and 1 U of Taq polymerase. Amplification products were loaded on a Li-COR 4300 analyzer, and visualized using fluorescent tags attached to M13 primers. DNA fragments were sized using Li-COR size standards and internal controls (three samples run on every load). Both Paulo C. Pulgarín and Theresa Burg checked microsatellite scores for every locus independently. We also developed a check plate containing multiple samples from each gel to confirm the original scores.

2.2.5 Mitochondrial DNA genetic analyses - Sequences were aligned manually using MEGA version 4 (Tamura et al. 2007). Two measurements of DNA polymorphism, nucleotide (π) and haplotype (H_d) diversity indices (Nei 1987), were calculated for each

population using DnaSP version 4.50.1 (Rozas et al. 2003). Nucleotide diversity is the average number of nucleotide differences per site between any two DNA sequences selected at random from the same population. Haplotype diversity combines the number and frequency of different genetic variants in the population and ranges from 0 (all haplotypes are identical) to 1 (all haplotypes are different) (Nei 1987). To understand how the genetic variation was partitioned among and within populations, Wright's fixation index F_{ST} was calculated in Arlequin version 3.1 (Excoffier et al. 2005). F_{ST} measures the proportion of genetic variance due to allele frequency differences among populations (predefined sampling sites) relative to the total population (all samples) (Holsinger & Weir 2009).

To assess the influences of genetic drift and gene flow, I tested for a positive association between genetic ($F_{ST}/(1-F_{ST})$) and geographic distances with those expected under a stepping-stone model of population structure (Hutchison & Templeton 1999) using isolation by distance (IBD) analysis. ($F_{ST}/(1-F_{ST})$), Rousset's distance, was originally proposed as an alternative, and more robust, method to perform direct comparisons between genetic and geographic (linear or two dimensional) distances Rousset (1997). Before Rousset's distance, the relationship between genetic and geographic distance was usually assessed by using a single F-statistic for all the populations.

If the IBD correlation is significant and positive, one would expect more migrants to be exchanged between proximate populations, rather than between distant areas. A

matrix of distances among sampling locations was calculated using the Geographic Distance Matrix Generator (GDMG) version 1.2.3 available at http://biodiversityinformatics.amnh.org/open_source/gdmg/index.php. This tool is widely used in IBD analyses and uses a spherical function to calculate distance based on latitude and longitude coordinates. A unified latitude and longitude point for every population was calculated by converting the available coordinates to a convex polygon, and the polygon was used to calculate a centroid. Polygons and centroids were calculated using DIVA-GIS version 7.2.3 (http://www.diva-gis.org). Significance of the IBD analysis was assessed with a Mantel test and 500,000 permutations in GenAlEx version 6 (Peakall & Smouse 2006).

Phylogeography and demographic history of the Downy Woodpecker were studied using two complementary approaches. First, genealogical relationships among haplotypes were visualized using the software TCS version 1.21 (Clement et al. 2000). This program produces a parsimony-based minimum-spanning network that displays the number of possible connections (base pair differences) between haplotypes with a 95% of confidence per link. The network is analogous to a phylogeny but when studying populations of the same species a network is better because it helps to visualize complex patterns that are not recovered by phylogenetic methods.

Fu's F_s and Ramos-Onsins' R_2 tests of neutrality were calculated to detect signatures of a recent population expansion (Fu 1997, Ramos-Onsins & Rozas 2002). F_s and R_2 are very powerful at detecting population expansions (increases in effective

population size) by estimating departures from neutrality. Large, significant negative F_s values reject the null hypothesis of population stasis, and low (< 0.05), significant R_2 values indicate an excess of rare alleles and young mutations. Both were implemented in DnaSP version 4.50.1 and significance was evaluated with 10,000 coalescent simulations. Similarly, Tajima's D was used to test against selective neutrality and population equilibrium (Tajima 1989). Negative and significant values of Tajima's D suggest either population expansion or purifying selection. This last statistic was calculated in Arlequin version 3.1 (Excoffier et al. 2005). Additionally, I performed the Fu and Li's D* and F* tests in order to see if background selection will better explain the patterns of genetic diversity in this species. If Fu and Li's D* and F* are significant and Fu's F_s and Tajima's D are not, then background selection is more likely creating the observed pattern (Fu 1997).

Finally, I used mismatch distributions to help visualize signatures of demographic expansion for all the samples combined, and to test against the null hypothesis of population growth (Rogers 1995, Rogers & Harpending 1992). The simulated distribution for the nucleotide pairwise differences (under population growth model) was compared with the observed distribution using the sum of square deviations (SSD). Likewise, the raggedness statistic (r) was calculated to test if the frequency of pairwise nucleotide differences followed a smooth curve expected from a growing population, comparing the expected and observed (r). Low values of r suggest a good fit of the model. Both SSD and r were calculated in Arlequin version 3.1 (Excoffier et al. 2005).

2.2.6 Microsatellite genetic analyses - MICRO-CHECKER (Oosterhout et al. 2004) was used to detect errors in scoring and the presence of null alleles. Allele frequencies, allelic richness, observed (*Ho*) and expected heterozygosity (*He*) were calculated for every population and each locus using the package GenAlEx version 6 (Peakal & Smouse 2006). Linkage disequilibrium and Hardy-Weinberg equilibrium were assessed (using Fisher's exact probability test) in the web version of GENEPOP version 4.0.10 (Raymond & Rousset 1995). The Markov chain parameter was set to 10,000 dememorizations, 1000 batches, and 10,000 iterations per batch for both analyses.

Genetic differentiation among sampled populations was evaluated using F_{ST} (Weir & Cockerham 1984) implemented in GENETIX version 4.05 (Belkhir et al. 1996). A total of 10,000 permutations was used to calculate the significance of the F_{ST}. For microsatellite analysis of population structure, F_{ST} is preferred over R_{ST} when the number of loci does not exceed 20 (Gaggiotti et al. 1999) as is the case with this study.

Another way to study population structure is using a Bayesian clustering method in which individuals are assigned to populations using a probabilistic approach. STRUCTURE version 2.3.3 (Pritchard et al. 2000) was used to infer the number of clusters (K). STRUCTURE assigns individuals to clusters independently of where they were sampled. K was set to values ranging from 1 to 16 during the initial exploratory runs (500,000 generations with a burn in of 100,000), and was then reduced to K=1-5 for the final analyses because of the reduced number of clusters detected in the initial runs. The final analysis was run for 1,000,000 generations with a burn in of 500,000 assuming

population admixture (same alpha for the degree of admixture), correlated allele frequencies, and including *a priori* population information. Ten replicates were performed at each value of K for the initial and final runs. Average ln Pr (X|K) was plotted against K to visualize the number of possible clusters as suggested by the authors. Although STRUCTURE represents a novel approach, the biological meaning of the number of groups (K) can be problematic to infer when there are few available loci or the populations are not strongly differentiated. Likewise, because it is difficult to decide which K captures the major structure in the data due to similar ln Pr (X|K) values (the yardstick to infer K), two complementary methods were employed. One is the Bayes factor, where averaged posterior probabilities for each K are compared in order to select the more meaningful K (Pritchard et al. 2000). The second is the ΔK method, suggested by Evanno et al. (2005), implemented using the on-line tool STRUCTURE Harvester (http://taylor0.biology.ucla.edu/struct_harvest/). ΔK method is a procedure that eliminates the possibility of overestimating the potential number of populations or groups due to the fluctuating nature of probability ln Pr (X|K) as K increases.

An isolation by distance (IBD) analysis was carried out for microsatellites in order to test the relationship between genetic and geographic distance. Refer to mitochondrial DNA analyses for a complete description and explanation of this method and program used.

2.2.7 Ecological niche modeling (ENM) - To gain a better understanding of the possible areas where the Downy Woodpecker survived during the LGM, a distribution model was

constructed using the maximum entropy method implemented in Maxent version 3.3.3e (Phillips et al. 2006). Maxent estimates maximum entropy distributions using climatic variables to predict non-negative probabilities on a target distribution using presence records (Stockman & Bond 2007). The analysis is based on 19 bioclimatic variables (Appendix 2) from the WorldClim dataset, publically available and extensively used in ENM (see Hijmans et al. 2005, Waltari et al. 2007). For all analyses, I used the default convergence threshold and maximum number of iterations (500), using 25% of the localities for model training. Maxent produces a continuous probability value (0 to 1 using the logistic default output) as an indicator of the presence/suitability for the species based in the principle of maximum entropy (Phillips et al. 2006). Distribution modeling for the LGM was done using paleoclimatic data drawn from the Community Climate System Model (Otto-Bliesner et al. 2006).

A set of 27,633 localities based on collection sites for museum specimens and breeding bird data were used to generate current and historical distribution maps. All the localities (latitude and longitude data) were downloaded from the Global Biodiversity Information Facility (GBIF) data portal (http://data.gbif.org), and included the localities used in the genetic analyses. Duplicate records were removed as part of the default setting in Maxent to avoid pseudo-replication. To test if the modeled distribution of the Downy Woodpecker during the LGM corresponds to a realistic model, I used two basic techniques implemented in Maxent: The first is the omission plot. If the data used for training deviates from the predicted omission line then the model is statistically weak. The second, if the area under the curve (AUC), which tests if the model predicted with

the training data (25% of the total locality records) is similar to the testing data, and if both are not better than a random model. Values range from 0 to 1, and values under 0.5 are not better than a random model (Phillips et al. 2006).

2.3 RESULTS

2.3.1 Genetic diversity - A total of 850 base pairs of DNA sequence was obtained for 251 individuals. No insertions or deletions were found. Global nucleotide ($\pi = 0.00097$) and haplotype ($H_d = 0.5839$) diversities were relatively wide ranging from $H_d = 0.182$ to 1.0 (Table 2.1). Among the 52 haplotypes none was differentiated by more than six base pairs, and all the populations shared a general widespread haplotype except Revelstoke, BC. A regression between H_d or π vs latitude or longitude found no significant correlation or trend for any of the comparisons (H_d vs. latitude, $r^2 = 0.0287$, $P = 0.474$; H_d vs. longitude, $r^2 = 0.0016$, $P = 0.865$; π vs. latitude, $r^2 = 0.0031$, $P = 0.804$; π vs. longitude, $r^2 = 0.0207$, $P = 0.539$). In general, Alaska, Washington and Southern Alberta were the least diverse regions in terms of H_d and π, and Missouri, Idaho, Revelstoke (BC), Illinois-Wisconsin and Louisiana exhibited the most diversity for the two genetic indices above.

Micro-checker analysis showed no evidence of scoring errors due to allele dropout or stutter for any of the loci. The only evidence for null alleles arising from an excess of homozygotes was in DIU5 as suggested by the Monte Carlo simulation on randomized alleles within populations (see micro-checker manual for specific details). Both Illinois-Wisconsin (DMA9, $P = 0.0217$), and Vancouver Island (DMC118, $P = 0.0360$) deviated from Hardy-Weinberg equilibrium, but neither was significant after correction for multiple tests. There was no evidence of linkage disequilibrium among loci as indicated by Fisher's exact test (P values ranged from 0.121 to 1.000).

The seven-microsatellite loci had moderate levels of genetic variation among loci and across populations. The number of alleles per locus (and average number) ranged from 4 (2.7 ± 0.159) in DMA9 to 17 (9.4 ± 0.608) in DMC118 (Table 2.2 only contains the average number of alleles). On average, Alaska had the lowest number of alleles (3.3 ± 0.71), while Nova Scotia – New Brunswick had the highest (6.1 ± 1.08) (Table 2.1). Genetic diversity, measured indirectly using H_e, in microsatellites was relatively uniform across the range with no association with latitude (r^2 = 0.1325, P = 0.1703) or longitude (r^2 = 0.0422, P = 0.4557).

2.3.2 Population structure - Mitochondrial-based pairwise F_{ST} comparisons across 20 sampling locations revealed minor differences among most pairs of populations (Table 2.3). Most of the significant differences were among Revelstoke, Idaho, Montana and Michigan (independently) relative to the rest of the sampling locations. Revelstoke was the only population where Hap1 (see phylogeography and demographic history and Table 2.5) was absent. Hap1 was also rare in ID and found in a single bird (n=6). Plotting the genetic similarities (F_{ST}/(1-F_{ST})) as a function of the geographic distance among populations revealed no significant relationship (Mantel test, r^2 = 0.0091, P = 0.220) (Fig 2.2).

Microsatellite-based pairwise F_{ST} comparisons among 15 sampling locations ranged from -0.011 to 0.086. In general, Alaska, Utah, Colorado and Montana were significantly different with the rest of the populations (Table 2.4). These results suggest

subtle, but significant, population structure across the species' range. The isolation by distance analysis found a strong trend, but no significant correlation between genetic and geographic distance ($r^2 = 0.1103$, $P = 0.070$) (Fig 2.3).

The Bayesian clustering analysis estimated at K = 2 as the most probable number of Downy Woodpecker subgroups; the same hierarchical number of groups was recovered by the ΔK method (Evanno et al. 2005) (Fig 2.4), and was supported by the Bayes factor where K = 2 (0.999) was the most probable number of groupings (K = 3 - Bayes factor 8.350×10^{-6} and K = 4 - Bayes factor 2.108×10^{-19}). Based on assignment probabilities (coefficient of ancestry – $Q > 0.6$) for each individual, one of the groups includes Utah, Montana, Colorado and Southern Alberta (group 1), and the second group includes most of the other populations (group 2). The exceptions to these two exclusive groups were Alaska, Michigan and Central Alberta in which most individuals are equally likely to be assigned to group 1 or 2 (Fig 2.5a and 2.5b). Independent analyses of groups 1 and 2 (with inclusion of Alaska, Michigan and central Alberta) suggest no substructure within the groups.

2.3.3 Phylogeographic and demographic history - Genealogical relationships among 52 mitochondrial haplotypes created a starburst pattern, a common, central haplotype different from most other less frequent haplotypes by one or two nucleotide differences; which suggests a demographic expansion. Furthermore the absence of any geographic break in the distribution of the haplotypes in the TCS network suggests limited phylogeographic structure of breeding populations of the Downy Woodpecker in North

America (Fig 2.6). A single widespread haplotype (Hap1) was present in all the populations except Revelstoke, BC. Hap1 was the most frequent haplotype present in 162 individuals (62%), followed by Hap8 (9 ind. – 3.5%), Hap4 (6 ind. – 2.3%), Hap42 (5 ind. – 2%) and Hap21 (4 ind. – 1.59%). The rest of the 47 haplotypes were present in one, two or three individuals (Table 2.5, Fig 2.6).

The hypothesis of constant population size was rejected by Fu's F_s (-8.3204, $P < 0.0001$) and Ramos-Onsins' R_s (0.7930, $P < 0.0001$). Likewise, Tajima's D rejected a scenario of selective neutrality and population equilibrium (-2.65, $P < 0.0001$). Fu and Li's D* (-6.54, $P < 0.02$) and F* (-5.77, $P < 0.02$) were both negative and significant, therefore a scenario of background selection cannot be totally rejected. Nonetheless, a unimodal mismatch distribution signals that Downy Woodpecker populations have undergone periods of population growth (goodness-of-fit test found no significant differences between the observed and expected frequencies – SSD = 0.00034, $P = 0.75$) (Fig 2.7). This was also confirmed by the raggedness index, which failed to reject the null hypothesis of population expansion (r = 0.068, $P = 0.36$).

2.3.4 Ecological niche modeling - The resulting models for the current and LGM 21 kya Downy Woodpecker distribution in North America were constructed based on ~ 16,271 presence records used for testing and 5,423 for training. The predicted models for the current distribution and at 21 kya were statistically robust using guidelines suggested by Phillips et al. (2006). The omission plot (not shown) displayed a good match between the predicted omission by the program algorithm and the training data (25% of locality

records). Both the training data (5,423 records) and the test data (16,271 records) predicted similar models, as the area under the curve test - AUC = 0.942 (values can range from 0 to 1, and values less than 0.5 are not better than a random model).

The predicted model for the current distribution suggests a strong correspondence with other sources of information related to the current species' distribution (e.g. Jackson and Ouellet 2002) (Fig 2.8). The predicted distribution during the LGM (18-21 kya) suggests two main areas (Fig 2.9). The largest area corresponds to the southeastern US from North Carolina west to Texas, with a proportion of the suitable habitat outside the current coastal line (sea level dropped up to 120 m 21 kya – Ray & Adams 2001). The second main area corresponds to the Rockies with higher distribution probabilities in the east: Arizona, Utah, and Colorado, and west: California and Oregon (Fig 2.9). At the LGM these areas were covered with open boreal woodlands, semi-arid temperate woodlands or scrub, forest steppe, subalpine parkland, temperate steppe parkland and taiga (Ray & Adams 2001). Based on current habitat preferences (Jackson & Ouellet 2002), it seems that all of the above could have been potential places for the Downy Woodpecker to persist.

2.4 DISCUSSION

2.4.1 Congruencies and contrasts between mtDNA and microsatellites - This study represents the first comprehensive analysis on patterns of genetic variation in the Downy Woodpecker in the context of postglacial expansion and colonization in North America. The results derived from mtDNA and microsatellite data suggest two perspectives on the genetic structure of the Downy Woodpecker. Both markers showed limited genetic structure across sampled locations, with most geographic groups characterized by a wide range of haplotype and nucleotide diversities, and similar expected heterozygosities. Microsatellites loci faster than mtDNA, and provide a more current perspective on the genetic variation) supported at least two genetically differentiated groups; whereas mtDNA (which is good at recovering more historical patterns of genetic variation) recovered only one genetic group.

In general, the data were not able to support prediction number two, which suggested a structured haplotype network, unique genetic variants and a bi-modal mismatch distribution. Instead, the absence of unique haplotypes or private alleles restricted to any specific geographic region and a unimodal mismatch distribution suggest a common history as predicted for scenario one. Microsatellite and mtDNA data suggest Alaska has the lowest genetic diversity (Table 2.1) compared with the other sampled populations, which also support the idea of a recent founder event from populations with low genetic diversity. In comparison, Washington has low nucleotide ($\pi = 0.00047$) and haplotype ($H_d = 0.294$) diversities in mtDNA that contrast with one of the highest mean

allelic richness (3.87) and *He* (0.612) for nuclear markers. Southern Alberta (π = 0.00036, H_d = 0,295, *He* = 0.576) shows a similar pattern to Washington.

Patterns in genetic diversity indices for mtDNA and microsatellites have shown to be contrasting when they are studied in the same species. This may reflect the mutational process intrinsic to them, but also the effective population size associated with nuclear and mitochondrial genomes (Burg & Croxall 2001, Brito 2007). It also suggests that there might be current-day processes such as recent limited gene flow between contemporary populations, a pattern that could be detected by the microsatellites but not the mtDNA (Beadell et al. 2010).

As microsatellites are biparental markers, the current perspective is reflecting the mixed population history of males and females, including processes such as change in population sizes for the studied areas, and any historical fluctuation in the number of males and females. The mtDNA only reflect the historical patterns of female lineages and populations, therefore a direct comparison needs to be taken with caution. Controlling for such types of effects was not in the scope of this study.

Mutational variance in the seven microsatellite loci studied could also explain the differences between the two types of markers. In the microsatellite analysis, basically seven different regions with possibly different mutation rates and genomic history were combined together to infer population structure. It is possible that some locus has evolved more rapidly than others, and that the genotypes based on seven markers are not

reflecting the overall genetic structure of the species. In humans, a recent study using 678 autosomal loci have shown that population history inferred from a wider representation of the genome can result in a different perspective on the number of historical populations in the New World native people (Want et al. 2007).

2.4.2 Population structure - F_{ST} based differences suggest mostly congruent patterns between mtDNA and microsatellite data, which is expected under equilibrium due to the different effective population sizes plus other factors, such as higher mutation rates in microsatellites than mtDNA (Ellegren 2004, Brito 2007). For instance, Southern Alberta, Montana, Utah and Colorado were well differentiated in the microsatellite analysis (Fig 2.5a and Fig 2.5b), but not significantly different from the other populations using mtDNA (Fig 2.6, Table 2.3). Nevertheless, the IBD analyses for mtDNA and microsatellite data suggest gene flow is occurring at all geographic distances, and most variation (> 90% in AMOVAs, not shown) is contained within populations. Therefore, over a large geographic scale, the effects of genetic drift may be stronger than gene flow in Alaska and the populations east of the Rocky Mountains, a pattern seen at a different geographic scale in the Greater Prairie Chicken (*Tympanuchus cupido*) (Johnson et al. 2003).

As gene flow is not apparently impeded by geographic distance in the Downy Woodpecker, one would expect individuals to be moving around at a high frequency. In an independent study, Browning (1995) found that over 113 band recoveries of Downy Woodpeckers, 94% were hatch year and after hatch year females moving during different

seasons (on average 218 km). Three of the recoveries were 583 km (winter to summer), 591 km (fall to winter) and 1080 km (summer to winter) away from banding locations. This scenario of females moving seasonally, even at low frequencies, can result in the homogenization of different populations (Kimura & Weiss 1964, Hewitt 1996).

If females' movements (or dispersal) are higher in some geographic areas, this would have had a strong influence in the genetic architecture of different populations, for both microsatellites and mtDNA. Although not fully understood, it seems that Downy Woodpeckers disperse (mostly females) during the breeding season searching for food and empty territories (Greenwood & Harvey 1982). If there is differential dispersal between areas due to bird density or resources, it is not known, a matter that needs to be studied to understand how much gene flow is occurring in current populations (Jackson & Ouellet 2002).

2.4.3 Demographic history and ecological niche modeling - The results of this study support scenario number one, a postglacial expansion over the present range by the Downy Woodpecker during the last 21 kya from a single refugium. Several independent lines of evidence agree with this. First, a star-like haplotype network characterized by a predominant haplotype and accompanied by a collection of low frequency peripheral haplotypes differing by one or two mutational steps from the main haplotype (Ball et al. 1988, Ball & Avise 1992, Slatkin & Hudson 1991, McCormack 2008). Second, the absence of high genetic divergences among geographic groups even with reduced structure as in this case (Stenzler et al. 2009, Milá et al. 2000). Third, a unimodal

mismatch distribution not different from a model of sudden expansion, which suggests an increase in effective population size (Rogers & Harpending 1992). More support comes from the neutrality test, where all, Tajima's D, Fu's F_s and R_s rejected the null hypothesis of population stability. However, background selection could be responsible of the genetic patterns visualized in the mtDNA as Fu and Li's F* and D* were also significant.

The distribution model for the Downy Woodpecker at 21 kya suggests two general refugial areas, one in the southeastern US and a second near the Rocky Mountains in the west from Arizona to Oregon, separated by a gap comprised of New Mexico, Nevada and Colorado (Fig 2.9). Assuming that this distribution model represents a realistic scenario, it strongly favors prediction number one, and the expectations of one or limited number of southern refugia.

The predicted suitable habitat 21 kya does not support any additional isolated areas north of the ice sheets. Therefore ENM does not support isolation in a northern refugium, such as Beringia, for Downy Woodpeckers. From a genetic point of view, both mtDNA and microsatellite data show Alaska has the lowest genetic diversity of all the populations (and lacks unique genetic variants), suggesting that it was recently colonized from other populations. The genetic affinities of non-sampled populations in the Midwest such as Oklahoma, Kansas, Nebraska, South Dakota, North Dakota, Manitoba, and Saskatchewan are unknown, however, based on the current genetic patterns it is unlikely that the inclusion of samples from these areas will change the general patterns depicted here.

The findings support an earlier study that suggested limited phylogeographic structure in the Downy Woodpecker as assessed by restriction analyses of mtDNA of a limited number of samples (n = 51) (Ball & Avise 1992). However, their restriction analyses (or single locus mtDNA phylogeography) did not detect the limited, but significant, structure found east of the Rocky Mountains in Montana, Colorado, Utah and Southern Alberta. It is possible that the Rocky Mountains have acted as a dispersal barrier for those populations as they are isolated by the rough topography of the landscape, but also because they are connected to the other western populations by the boreal forest or other major forested areas. From the mtDNA point of view Ball & Avise (1992) neither detected the subtle genetic uniqueness of the Revelstoke and Idaho (characterized by lack of the common haplotype and high H_d and π), which might reflect historical isolation within or nearby the Clearwater refugium (see Shaffer et al. 2010), however, obscured by recent gene flow or positive selection.

The use of microsatellites and sequence data in this study offered a much more complete and better explanation when compared with Ball & Avise (1992) findings. Multilocus (mtDNA and microsatellites combined in this case) phylogeography has become a popular approach because it can help to infer if the genetic variation corresponds to current phenomena (such as habitat fragmentation) or past events (such as population expansion), as recently shown in the Black-backed Woodpecker (*Picoides arcticus*) (Pierson et al. 2010).

In summary, I was able to support scenario number one which favors a pattern of recent postglacial expansion in the Downy Woodpecker. All the predictions based on the signatures of the genetic data (low genetic diversity, unimodal mismatch distribution, star-like haplotype network and significance of all the neutrality tests) were met and suggest that the species has undergone population growth and demographic expansion during recent times as conditions became favorable. The distribution model for the LGM indicates that this species probably persisted in two reduced areas south of the ice sheets. In general those areas (east and west) reflect the two general groups identified by the microsatellite data, however the east-west groups are not supported by the distribution of the mtDNA haplotypes. Notwithstanding, the distribution of haplotypes 4 and 8 shared by individuals in populations from Central British Columbia to Colorado, and the absence of haplotype 1 in Revelstoke, BC, could suggest a more stable and complex history in this part of their range.

Table 2.1. Genetic diversity statistics for mtDNA and microsatellite across populations and loci. Number of segregating sites (SS), number of haplotypes (# Hap), nucleotide (π) and haplotype (H_d) diversities; average number of alleles (# alleles), allelic richness (AR), and observed (*Ho*) and expected (*He*) heterozygosities.

		Mitochondrial DNA					Microsatellites				
	Population	**n**	**SS**	**# Hap**	H_d	π	**n**	**# Alleles**	**AR**	***Ho***	***He***
AK	**Alaska**	11	1	2	0.182	0.00021	11	3.29	2.84	0.458	0.491
CBC	**Central BC**	7	3	4	0.714	0.00101	*	*	*	*	*
BCR	**BC Revelstoke**	9	6	5	0.861	0.00275	*	*	*	*	*
VI	**Vancouver Island**	*	*	*	*	*	11	4.71	*	0.685	0.620
ID	**Idaho**	6	7	6	1.000	0.00310	*	*	*	*	*
WA	**Washington**	13	2	3	0.295	0.00036	14	5.14	3.87	0.699	0.612
OR	**Oregon**	8	3	4	0.643	0.00089	*	*	*	*	*
CAB	**Central Alberta**	19	6	5	0.462	0.00097	19	6.00	3.75	0.644	0.587
SAB	**Southern Alberta**	13	2	3	0.295	0.00036	13	5.00	3.61	0.638	0.576
MT	**Montana**	19	3	4	0.719	0.00108	19	5.57	3.89	0.653	0.612
CO	**Colorado**	19	5	6	0.538	0.00073	23	5.29	3.65	0.645	0.584
UT	**Utah**	13	2	3	0.410	0.00051	16	4.57	3.21	0.626	0.539
ON	**Ontario**	*	*	*	*	*	15	3.57	*	0.529	0.507
NSNB	**Nova Scotia/New Brunswick**	15	5	5	0.476	0.00079	32	6.14	3.88	0.687	0.626
NL	**Newfoundland**	18	3	4	0.314	0.00039	16	4.71	3.41	0.613	0.568
MI	**Michigan**	19	8	8	0.614	0.00099	19	5.71	3.66	0.616	0.583
ILWI	**Illinois/Wisconsin**	18	8	9	0.758	0.00117	20	6.00	3.8	0.625	0.596
MO	**Missouri**	14	6	7	0.846	0.00182	15	5.71	3.81	0.671	0.580
NC	**North Carolina**	10	2	3	0.378	0.00047	14	4.71	3.69	0.647	0.598
WV	**Western Virginia**	5	1	2	0.400	0.00047	*	*	*	*	*
LA	**Louisiana**	4	2	3	0.833	0.00118	*	*	*	*	*
FL	**Florida**	9	2	3	0.417	0.00052	*	*	*	*	*

*For microsatellites only populations with > 10 individuals were used. For mitochondrial DNA, sequences from Ontario and from Vancouver Island (except one) were not available for population comparisons. One individual from California and Vancouver Island were included in the haplotype network analyses but excluded here.

Table 2.2. Average and SE sample size for (N), number of alleles (Na), number of effective alleles (Ne), information index (I), observed heterozygosity (*Ho*), expected and unbiased expected heterozygosity (*He*) calculated for each locus.

		Locus						
		DMD118	**DMC111**	**DMA2**	**DlU5**	**DMA9**	**Ptri3**	**DMC115**
N	**Mean**	16.200	15.867	16.533	16.733	16.867	13.867	9.800
	SE	1.324	1.230	1.376	1.255	1.348	1.142	1.254
Na	**Mean**	9.400	6.133	2.733	3.467	2.667	7.333	3.800
	SE	0.608	0.336	0.248	0.307	0.159	0.361	0.368
Ne	**Mean**	5.684	3.799	1.438	1.570	1.988	4.889	3.079
	SE	0.379	0.264	0.046	0.111	0.041	0.283	0.268
I	**Mean**	1.916	1.471	0.524	0.615	0.752	1.724	1.169
	SE	0.071	0.079	0.041	0.081	0.024	0.053	0.101
Ho	**Mean**	0.868	0.715	0.324	0.307	0.522	0.871	0.795
	SE	0.028	0.041	0.025	0.051	0.028	0.018	0.067
He	**Mean**	0.811	0.708	0.295	0.319	0.494	0.785	0.639
	SE	0.015	0.033	0.021	0.046	0.011	0.014	0.049

Table 2.3. MtDNA pairwise F_{ST} values are above the diagonal; P values below the diagonal. Values in italics are significant prior to corrections for multiple tests and values in bold are significant P values after the Benjamini-Hochberg correction for multiple tests. Abbreviations for localities are in Table 2.1 and Figure 2.1.

	AK	**CBC**	**BCR**	**ID**	**WA**	**OR**	**CAB**	**SAB**	**MT**	**CO**	**UT**	**NSNB**	**NL**	**MI**	**ILWI**	**MO**	**NC**	**WV**	**LA**	**FL**
AK	*	0.041	0.239	0.121	-0.004	0.024	-0.002	-0.004	0.079	-	0.033	-0.013	-0.010	-0.039	-0.012	0.066	0.009	0.050	0.159	0.009
CBC	0.130	*	0.156	0.043	0.037	0.001	0.016	0.037	0.008	0.003	0.046	0.010	0.048	0.001	0.001	0.039	0.011	-0.022	0.009	0.011
BCR	**<0.001**	*0.040*	*	0.081	0.229	0.128	0.238	0.229	0.199	0.242	0.251	0.233	0.299	0.224	0.214	0.183	0.201	0.128	0.105	0.201
ID	*0.027*	0.177	0.148	*	0.156	0.061	0.155	0.156	0.108	0.166	0.166	0.147	0.214	0.125	0.132	0.109	0.112	0.031	0.007	0.112
WA	0.871	0.208	**<0.001**	*0.010*	*	-0.017	0.009	-0.040	0.054	-	0.033	-0.004	-0.001	-0.011	-0.001	0.076	0.007	0.018	0.125	0.007
OR	0.180	0.600	0.101	0.079	0.678	*	0.011	-0.017	0.028	-	0.003	0.004	0.032	-0.004	-0.002	0.044	0.004	-0.024	0.020	0.004
CAB	0.501	0.278	**<0.001**	*0.008*	0.434	0.285	*	0.009	0.086	0.024	0.029	0.005	0.018	0.014	0.014	0.084	0.000	-0.028	0.035	0.000
SAB	0.838	0.219	**<0.001**	*0.013*	0.999	0.658	0.429	*	0.020	-	0.033	-0.004	-0.001	-0.011	-0.001	0.076	0.007	0.018	0.125	0.007
MT	0.060	0.379	**0.002**	*0.031*	0.111	0.276	**0.006**	0.246	*	0.072	0.096	0.078	0.099	0.074	0.075	0.118	0.068	0.036	0.079	0.068
CO	0.600	0.461	**<0.001**	**0.003**	0.712	0.908	0.258	0.741	*0.033*	*	0.000	-0.001	0.011	0.008	0.016	0.089	0.001	-0.020	0.063	0.001
UT	0.424	0.178	**<0.001**	**0.003**	0.489	0.438	0.225	0.483	*0.029*	0.465	*	0.020	0.036	-0.013	0.017	0.078	0.029	0.024	0.106	0.029
NSNB	0.953	0.471	**<0.001**	**0.001**	0.889	0.579	0.391	0.900	*0.015*	0.541	0.265	*	0.004	-0.011	0.006	0.051	-0.009	-0.032	0.043	-0.009
NL	0.864	0.187	**<0.001**	**0.001**	0.736	0.267	0.427	0.733	*0.009*	0.451	0.154	0.271	*	-0.001	0.001	0.097	0.009	0.014	0.136	0.009
MI	0.999	0.626	**<0.001**	*0.010*	0.964	0.706	0.502	0.959	*0.010*	0.466	0.746	0.888	0.964	*	0.007	0.074	-0.017	-0.047	-0.012	-0.017
ILWI	0.654	0.524	**<0.001**	**0.001**	0.478	0.590	0.275	0.505	**0.004**	0.074	0.184	0.431	0.734	0.147	*	0.039	-0.011	-0.045	0.008	-0.011
MO	0.067	0.193	**<0.001**	*0.022*	*0.032*	0.116	**<0.001**	*0.021*	**<0.001**	**0.001**	*0.025*	0.087	**<0.001**	**<0.001**	0.082	*	0.049	0.001	0.015	0.049
NC	0.409	0.360	**<0.001**	**0.005**	0.547	0.449	0.416	0.564	0.104	0.537	0.303	0.802	0.556	0.912	0.784	0.112	*	-0.005	0.066	0.000
WV	0.563	0.939	0.072	0.320	0.646	0.900	0.692	0.646	0.299	0.739	0.357	0.813	0.661	0.895	0.914	0.394	0.778	*	0.025	-0.005
LA	0.174	0.676	0.188	0.388	0.237	0.552	0.287	0.211	0.180	0.271	0.156	0.471	0.237	0.502	0.578	0.363	0.384	0.442	*	0.066
FL	0.411	0.359	**<0.001**	**0.003**	0.568	0.445	0.443	0.539	0.121	0.516	0.284	0.781	0.566	0.915	0.798	0.116	0.999	0.764	0.353	*

Table 2.4. Microsatellites pairwise F_{ST} values are above the diagonal; and *P* values below the diagonal. Numbers in bold are significant after the Benjamini-Hochberg correction for multiple tests and numbers in italics are significant before corrections for multiple tests. Abbreviations for localities are in Table 2.1 and Figure 2.1.

	AK	**VI**	**WA**	**CAB**	**SAB**	**MT**	**UT**	**CO**	**ILWI**	**ON**	**MI**	**NSNB**	**NL**	**MO**	**NC**
AK	*	0.080	0.062	0.052	0.047	0.047	0.086	0.071	0.039	0.030	0.036	0.055	0.060	0.051	0.024
VI	*0.016*	*	0.002	0.034	0.028	0.037	0.074	0.051	0.019	0.018	0.046	0.031	0.028	0.008	0.021
WA	**<0.001**	0.408	*	0.017	0.030	0.012	0.072	0.042	0.010	0.024	0.007	0.007	0.019	0.013	-0.007
CAB	**0.003**	*0.016*	*0.041*	*	0.002	0.004	0.030	0.004	-0.001	0.023	0.003	0.015	0.008	-0.004	0.006
SAB	*0.021*	0.080	*0.025*	0.354	*	0.020	0.043	0.015	0.016	0.038	0.019	0.020	0.014	0.000	0.022
MT	**0.004**	*0.018*	0.117	0.263	0.062	*	0.023	-0.007	0.006	0.016	0.002	0.030	0.038	0.014	0.000
UT	**<0.001**	**0.001**	**<0.001**	**0.005**	**0.007**	*0.042*	*	0.004	0.036	0.059	0.019	0.073	0.081	0.064	0.053
CO	**<0.001**	**0.001**	**0.002**	0.252	0.096	0.810	0.263	*	0.017	0.028	0.016	0.049	0.050	0.028	0.029
ILWI	*0.024*	0.141	0.173	0.521	0.121	0.240	**0.004**	0.040	*	-0.000	0.002	0.014	0.011	0.005	-0.002
ON	0.245	0.155	0.102	*0.044*	*0.050*	0.152	*0.013*	*0.012*	0.459	*	0.015	0.015	0.021	0.011	0.008
MI	*0.024*	**0.008**	0.232	0.290	0.057	0.359	*0.050*	*0.044*	0.372	0.167	*	0.009	0.030	0.022	0.004
NSNB	**0.003**	*0.050*	0.234	*0.032*	0.054	**0.005**	**<0.001**	**<0.001**	0.070	0.132	0.107	*	0.004	0.007	0.002
NL	**0.007**	0.069	0.061	0.182	0.115	**0.003**	**<0.001**	**0.001**	0.151	0.111	0.015	0.302	*	-0.011	0.013
MO	*0.014*	0.269	0.108	0.615	0.466	0.096	**<0.001**	*0.013*	0.272	0.252	0.037	0.228	0.849	*	0.001
NC	0.082	0.107	0.729	0.229	0.054	0.445	**0.002**	*0.015*	0.535	0.266	0.307	0.384	0.154	0.397	

Table 2.5. Distribution of the mtDNA genetic variants (haplotypes) in twenty-two sampling areas across the Downy Woodpecker range.

Hap	AK	CBC	BCR	VI	ID	WA	OR	CA	CAB	SAB	MT	CO	UT	NSNB	NL	MI	ILWI	MO	NC	WV	LA	FL
1	10	4		1	1	11	5	1	14	11	9	13	9	11	15	12	9	5	8	4	2	7
2	1		2																			
3		1										1										
4		1			1						4											
5		1																				
6			3																			
7			1																			
8			2			1	1			1	3	1										
9			1																			
10					1																	
11					1																	
12					1																	
13					1																	
14						1																
15							1					2	1									
16							1															
17									1													
18									2													
19									1								1					
20									1					1								
21										1	3											
22												1										
23												1										
24													2			1						
25													1									
26														1				1				
27														1								
28														1		1						
29															1							
30															1		1					
31															1							
32																1						
33																1					1	
34																1						
35																1						
36																1						
37																	1					
38																	1					
39																	1					
40																	1					
41																	1					
42																	2	3				
43																		1				
44																		1				
45																		2				
46																		1				
47																			1			
48																			1			
49																				1		
50																					1	
51																						1
52																						1
Total	**11**	**7**	**9**	**1**	**6**	**13**	**8**	**1**	**19**	**13**	**19**	**19**	**13**	**15**	**18**	**19**	**18**	**14**	**10**	**5**	**4**	**9**

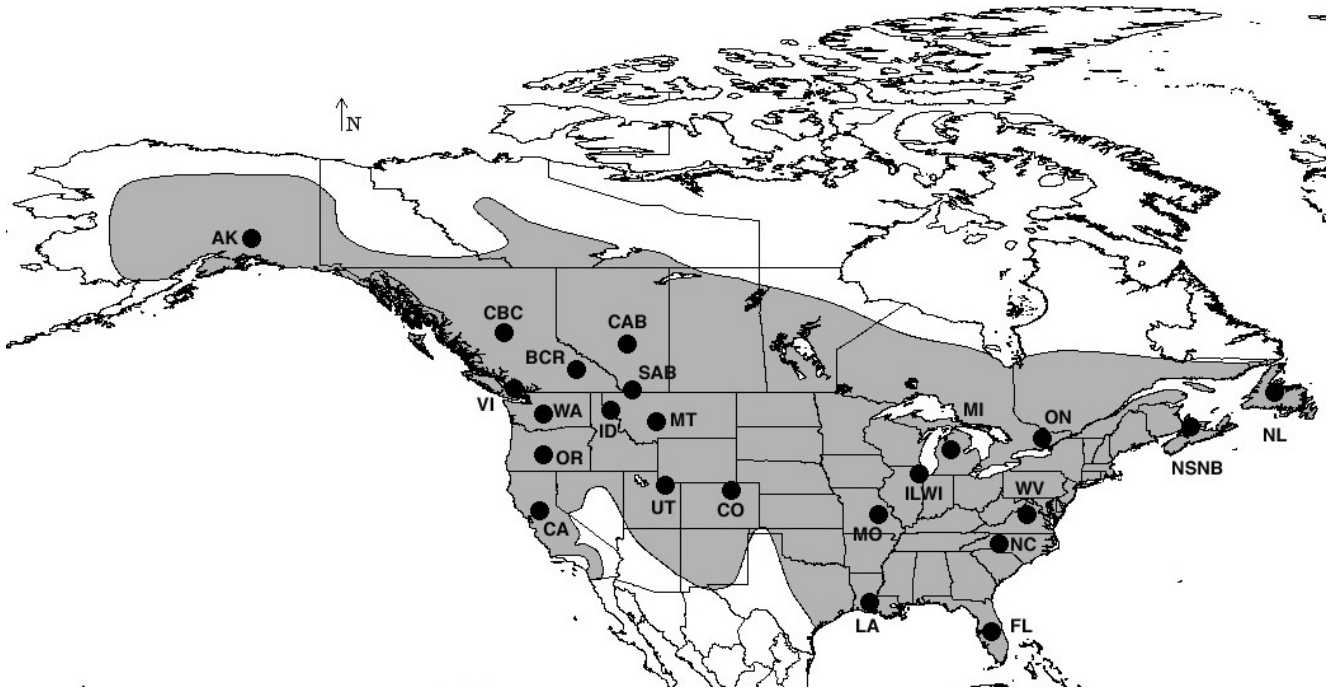


Figure 2.1. Distribution of the 23 sampling locations included this study. Abbreviations for localities are in Table 2.1. Grey area indicates the general distribution of the species (modified from Ridgely et al. 2007).

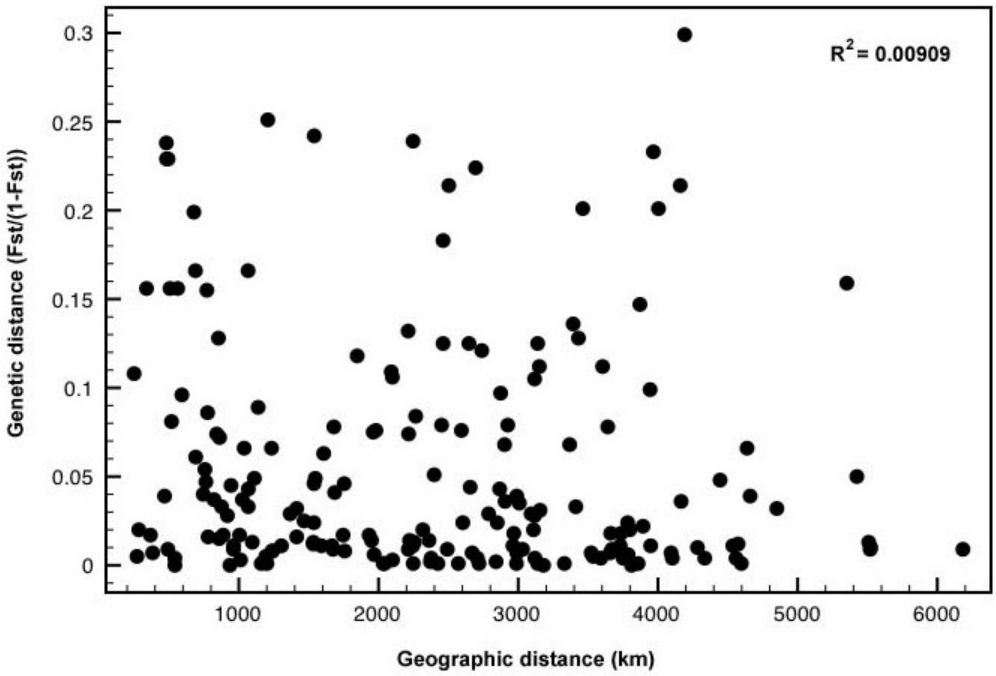


Figure 2.2. Pairwise comparisons of geographic and genetic distance among 20 geographic groups for which mtDNA data were available ($r^2 = 0.0091$, $P = 0.220$).

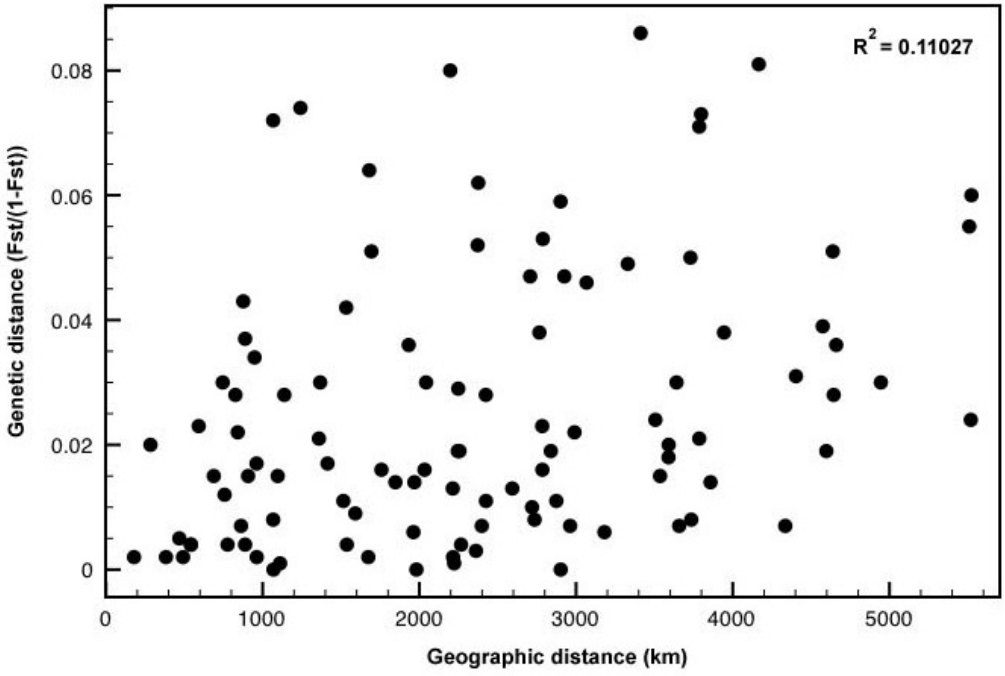


Figure 2.3. Isolation by distance analysis for 15 populations (microsatellites) suggested no relationship between genetic and geographic distance ($r^2 = 0.1103$, $P = 0.070$).

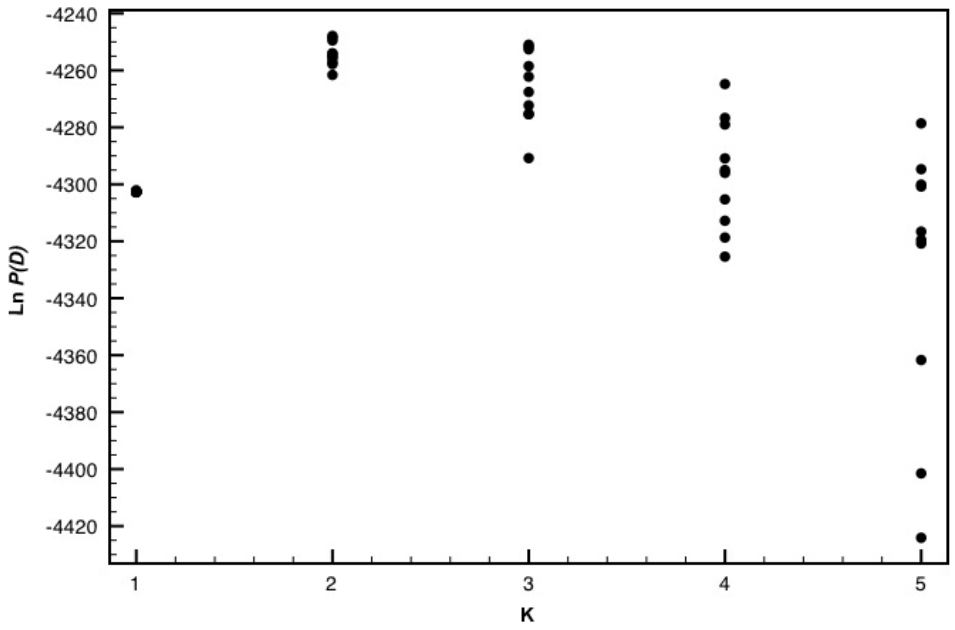
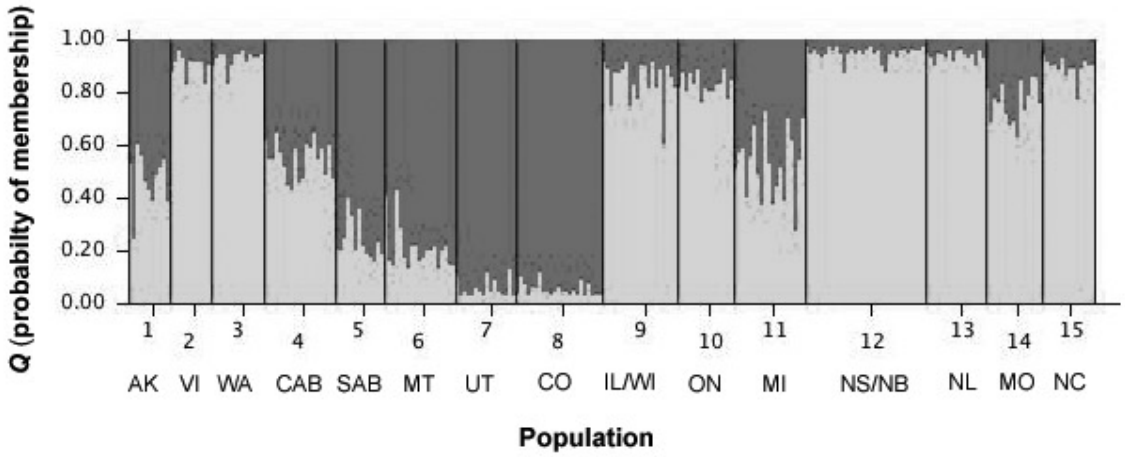
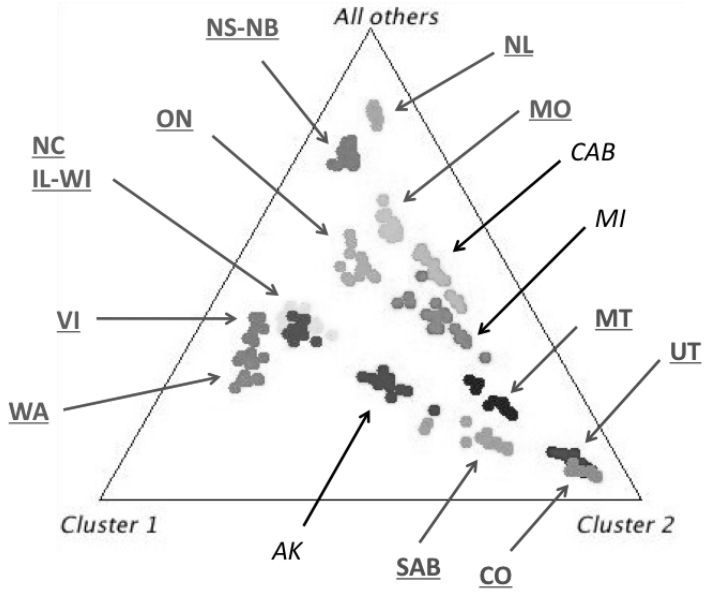


Figure 2.4. Posterior probability values from ten runs per given K (possible number of groups) in the program STRUCTURE. K=2 is suggested by the lower mean value and using the Evanno et al. (2005) method. The higher the ln *P(D)*, the higher the probability of K being the most likely hierarchical group.



a.



b.

Figure 2.5 a. Barplot of assignment probabilities for 257 individuals from 15 populations. Geographic groups were arranged to display the probability (*Q*) gradient for the sampled geographic areas. Each vertical bar (within the black lines) represents an individual from one population. *Q* values > 0.6 are regarded as confident for population assignment of an individual. b. STRUCTURE triangle plot displaying clustering patterns among 15 populations. Each point represents an individual. Names in bold and underline are populations for which individuals had assignment probabilities (*Q*) > 0.6; in italics are populations where individuals could not be assigned with confidence to groups in bold underlined (left or right) based on *Q*.

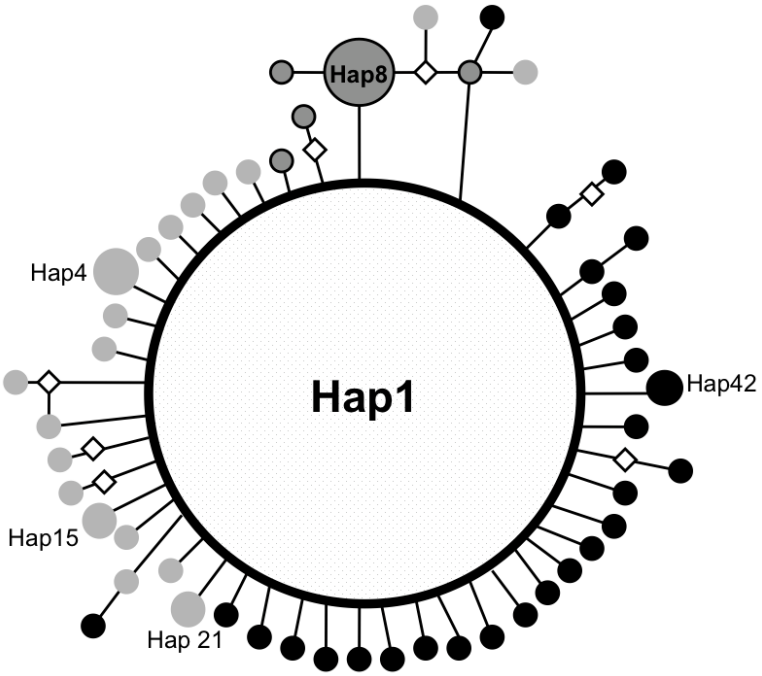


Figure 2.6. Minimum-spanning network of mtDNA haplotypes. Only haplotypes with four or more individuals are labeled (see Table 2.5). Hap1 was present in all the sampled populations with the exception of SEBC. Unsampled/inferred haplotypes are indicated with a diamond, haplotypes found exclusively in the east (east of LA, MO, and ILWI) are in black and those in the west in light grey (west of CO and MT) with exception of two black haplotypes. Haplotypes in dark gray and black outline represent SE BC population.

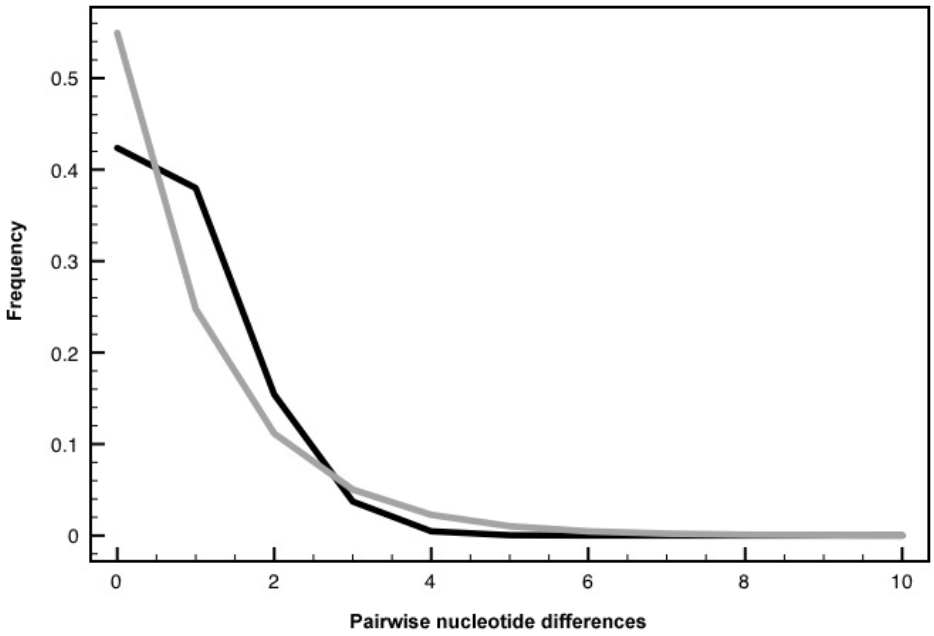


Figure 2.7. Mismatch distributions of pairwise nucleotide differences between mtDNA haplotypes revealed a unimodal pattern and no significant differences between the observed and expected frequencies. Sum of squared deviation = 0.000344, $P = 0.75$. Black line is the observed frequency and gray line is the expected frequency.

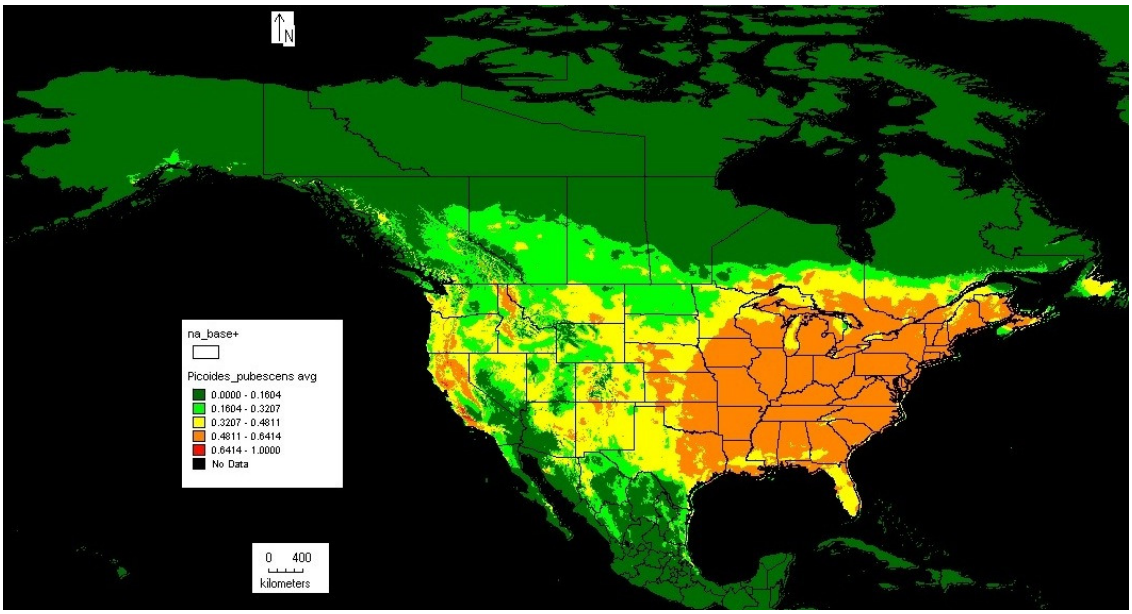


Figure 2.8. Current suitable habitat for the Downy Woodpecker as modeled with 19 bioclimatic variables. Light green represents areas with low presence probability and orange to yellow with higher. Probabilities range from 0 to 1. Dark green is the background.

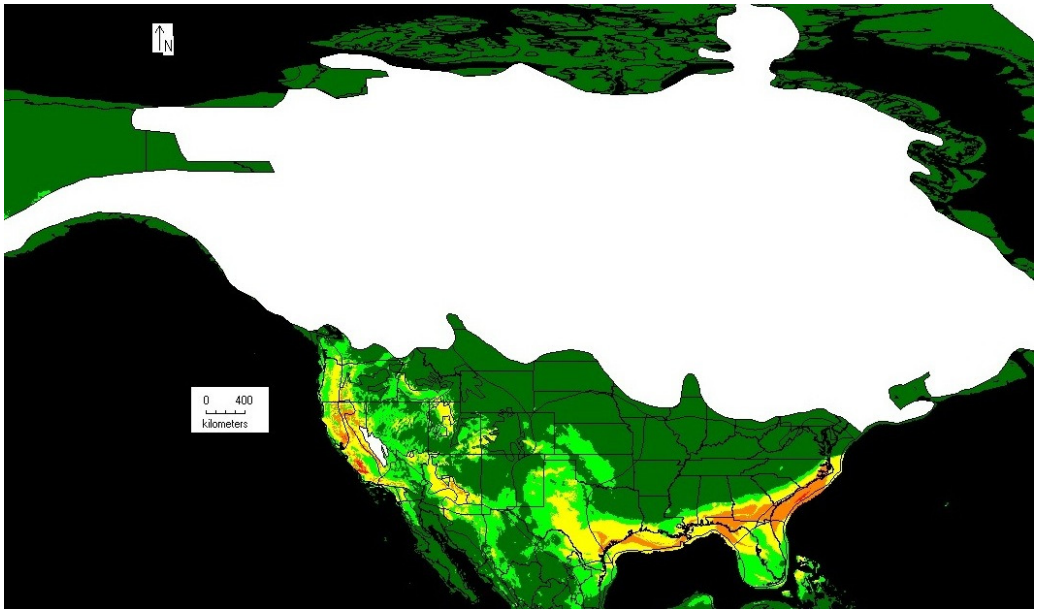


Figure 2.9. Modeled suitable habitat during the LGM (21 kya) for the Downy Woodpecker. Light green represents areas with low presence probability and orange to yellow with higher presence probability. White area at the top indicates the distribution of the ice sheets from Ray & Adams (2001).

CHAPTER THREE

GEOGRAPHIC VARIATION IN THE DOWNY WOODPECKER "*PIK*" CALL: A DESCRIPTIVE AND QUANTITATIVE ANALYSIS

ABSTRACT

The Downy Woodpecker (*Picoides pubescens*) is a common and widespread bird in North America. Despite many studies on ecology and breeding biology very little is known about geographic variation of acoustic signals. I carried out a quantitative and descriptive analysis of Downy Woodpecker pik calls (n = 317) from 42 individuals and 22 localities across the species' range to test whether or not this acoustic signal exhibits geographic variation among geographic areas and subspecies. An analysis of individual variation revealed that it is possible to identify and discriminate individual birds by the combination of frequency and temporal variables. I found no evidence of geographic variation in any of the six spectro-temporal variables, or differences between localities or among any of the four studied subspecies. There was no evidence for a relationship between latitude and any of the frequency or temporal variables. These results suggest that the pik calls exhibit more macro-geographic than individual variation.

Keywords: Acoustic communication, frequency, latitude, North America, Picidae, *Picoides pubescens*.

3.1 INTRODUCTION

The study of bird vocalizations in a geographic context has traditionally been one of the most active topics in avian bioacoustics, with most of the interest directed towards songbirds (Podos & Warren 2007). Other avian taxonomic groups for which vocal learning has not been demonstrated have received relatively less attention (Miller 1986, 1996, Bretagnolle & Genovis 1997). Species with extensive geographic ranges are of special interest because of geographic variation in selective pressures due to variation in environment, competition, or dispersal patterns (Irwin 2000, Koetz et al. 2007).

Woodpeckers are a widespread taxonomic group for which there is no evidence of vocal learning. Despite many studies on the ecology, systematics and evolution of woodpeckers (Short 1982, Koenig & Mumme 1987, Benz et al. 2006), very few analyses of geographic variation in vocalizations have been made (Winkler & Short 1978, Dodenhoff et al. 2001, Tremain et al. 2008). The limited published work suggests geographic variation in calls for the Great Spotted Woodpecker (*Dendrocopus major*) and the Hairy Woodpecker (*Picoides villosus*) (Winkler & Short 1978), and more recently geographic variation was detected in the drumming (non-vocal form of communication) of the Black-backed Woodpecker (*Picoides arcticus*) (Stark 2002).

Here, I studied acoustic variation in the Downy Woodpecker (*Picoides pubescens*), a common species throughout continental North America, found in a variety of habitats from 0 to 1500 meters above sea level, and a year-round resident from Alaska

to Florida (Fig 3.1). Taxonomists and systematists have recognized up to seven subspecies based on plumage and morphology (Fig 1.2), which suggests limited gene flow among populations, but whether this actually reflects genetic or vocal differences is unknown (Ouellet 1977, Jackson & Ouellet 2002).

The Downy Woodpecker sound repertoire consists of at least three main vocal signals plus non-vocal drumming and tapping (for more details see Jackson & Ouellet 2002). I focused on the pik call, a vocalization used in a wide array of situations, and regarded as a single-element signal (Winkler & Short 1978) (Fig 3.2). Pik calls might reflect motivational/excitement states, are used in territorial defense, in signaling individual position (sender's location), and also contains individual identity information (Kilham 1962, Short 1982, Dodenhoff 2002). The pik call is used year round, however it is used more frequently during the breeding season, and it is involved in mating behavior, therefore might be subjected to strong evolutionary pressures since it could reflect individual fitness (Dodenhoff 2002).

In this chapter I aim to qualitatively and quantitatively characterize the Downy Woodpecker pik call, specifically to understand whether it exhibits geographic variation at a continental scale. Based on the ecology, natural history, vocal behavior, distribution, and non-migratory behavior of the Downy Woodpecker, it is possible to test general predictions about the geographic variation of this call.

First, according to the non-migratory behavior of the Downy Woodpecker (Browning 1995), putative non-vocal learning, and presence of seven different subspecies, I predicted that the pik call should differ among distant geographic areas, and among subspecies, and also along latitudinal gradients. In theory, the geographic isolation of populations should result in populations with distinctive acoustic characteristics due to the lack of gene flow (Isler et al. 1998, Irwin et al. 2000). Accordingly, if the current subspecies are the result of recently independent evolutionary histories (possibly not interbreeding), one can predict that every subspecies could be characterized (partially or totally) by a distinctive pik call. Finally, it has been suggested a negative relationship between body size and peak frequency (pitch), whereby larger birds produce lower frequency vocalizations and vice versa (Patel et al. 2010). In general, Downy Woodpeckers are smaller in the south (with some exception in the southeast part of the range) and bigger in the as predicted by the Bergman's rule of body size variation across latitude in vertebrates (a positive relationship between body size and latitude, the northern the bigger) (see James 1970). Therefore one can predict that in the Downy Woodpecker the peak frequency (pitch) should decrease from south to north.

Second, it is possible to test some predictions based on the ecology, vocal behavior and habitat preferences. If the pik call is a multifunctional vocal trait used in many types of behavioral context (including individual identity) (see chapter one) one can predict that the Downy Woodpecker pik call will varies mostly among individuals irregardless the subspecies or geographic group. Finally, the variation in the pik call could be predicted based on the type of habitat this woodpecker uses, as supported by the

acoustic adaptation hypothesis (call will be different depending on transmission properties of the habitat) (Bretagnolle & Genovis 1997, Dingle et al. 2008, Koetz et al. 2007). It is known that the Downy Woodpecker uses different types of or forests and natural areas (e.g. coastal California vs. Alaska) (Jackson & Ouellet 2002), then, it is possible to expect pik calls to be distinctive among different type of habitats, irregardless of the geographic area or subspecific differences.

3.2 METHODS

3.2.1 Call selection, measurements and analysis - Recordings were obtained from sound archives, commercial audio CDs and private archives (Appendix 2). Spectrograms were produced only from recordings with a clear fundamental frequency and first harmonic. A total of 317 calls from 42 individuals and 22 localities recorded between 1956 and 2010 were used (Fig 3.1, Table 3.1). The identity of each bird was assessed using consecutive calls from files that contained recordings from a single bird. Information from the recorder (usually in the file) was also used to confirm if pik calls from one or more birds were in each recording. Additionally, when two individuals are recorded, there is a differential pattern in the amplitude values that can be easily recognized. Calls with similar amplitude values (intensity of the call) were classified as belonging to a single bird in recording archives that had more than one individual vocalizing.

Spectrograms from up to seven individuals per state/province were measured. Most localities were represented by a single individual (Fig 3.1, Table 3.1). I measured 2-12 calls per individual (average = 8.00 ± 2.54 SD). It was not possible to control for the effects of sex, recording equipment, habitat characteristics or age because of the limited number of available recordings.

In cases where more than 12 calls per individual were available, the clearest calls of highest amplitude were used for the analyses. Calls were analyzed using Raven 1.3 (Charif et al. 2008). Raven workspaces (a type of file that saves the on-screen

measurements done by the cursor, and captures the data in a text file) were created for every individual, and variable values were automatically stored on a selection table. All the variables were chosen from Raven's measurement panel. Spectrogram parameter settings were: window Hann: 150 samples, 3 dB filter bandwidth: 460 Hz, FFT: 512, overlap 80%. All recordings had a sampling rate of 44.1 kHz.

Time intervals between calls were ignored because the number of calls and tempo of calling varies with behavioral context. The following variables were measured on each call: HiF = maximum frequency; LoF = minimum terminal frequency; LoFSu = minimum frequency at the initial part of the note; DelT = call duration; DelF = bandwidth and Q3 = 3rd quartile of the frequency (Fig 3.2). All variables were measured on the first harmonic (fundamental frequency). Only single calls separated by > 0.20 seconds were used because "double calls" have been reported in this species (Winkler & Short 1978). I saw no more than five "double calls" during the call selection.

3.2.2 Geographic, subspecific and individual variation analyses - In order to test for individual variation and identity, I used seven birds from Ontario, and five birds from Ohio, and ran the analysis for each group independently (Ontario and Ohio have the largest number of calls from single individuals and from the area in general). Individual variation and identity was also tested using a MANOVA (individual as fixed effect). Discriminant function analysis (DFA) was used to see if individuals were correctly classified using individual calls. All the analyses were carried out in SPSS v.17 (SPSS

Inc., Chicago, IL) and PAST v. 2.02 (Hammer et al. 2001) using a level of significance of $a < 0.05$.

For every individual, mean values were estimated for each of the six variables (Table 3.1). In order to test for general geographic variation patterns, individuals were grouped by sampling areas based on latitude and longitude data (Table 3.1). Subspecific differences were tested by grouping individuals into four different subspecies *sensu* AOU (Jackson & Ouellet 2002): *P. p. pubescens*, *P. p. medianus*, *P. p. leucurus and P. p. gairdnerii* (Fig 1.2, Table 3.2). A multivariate analysis of variance (MANOVA) was used to test if the distribution of the samples based on geographic or subspecific groupings explain a significant amount of the variance when all the six variables are combined (rather than testing one variable among groups using several individual analysis of variance). I tested for correlations between latitude and maximum and minimum terminal frequency, and latitude and call duration using a least squared regression.

Additionally, I used a DFA to indentify differences in call characteristics among subspecies. DFA uses a linear combination of variables (six in this case) to characterize differences among groups of samples. Variables are translated into one or more discriminant functions that are also employed to classify individuals (or calls) to predefined correct or incorrect groups, in this case, subspecies.

3.3 RESULTS

Downy Woodpecker pik calls were characterized as "n" shape, with marked ascending and descending frequency modulation with some degree of asymmetry, usually skewed at the left side of the note. The fundamental frequency ranges from 1354.8-3750.6 Hz in a brief period of time (0.0334 ± 0.0058 s), and was similar among subspecies and most sampling sites (Table 3.1 and Table 3.2). In general, most notes displayed two to three harmonics with similar amplitude (dB) values; still the fundamental frequency had higher amplitude values on average (not shown). Call morphology and shape among locations were qualitatively and quantitatively different (Fig 3.3); variation (mostly in the frequency variables) was considerable among individuals (Fig 3.4a).

The discriminant function analysis of individual variation using samples from Ontario (seven individuals - 58 calls) and Ohio (five individuals - 40 calls) independently suggest that individuals can be discriminated and classified correctly using a combination of information from different acoustic variables. In Ontario 93.1% of the calls were correctly classified in seven individuals (Fig 3.4a) and in Ohio 92.5% were correctly classified in five individuals (plot not shown). Although I could not control for the effects of sex, age, habitat or recording equipment; it seems that pik call characteristics are relatively similar in birds (different individuals) recorded over a ~50 year interval. Individuals from different years were identified by the discriminant function analysis, but calls did not cluster based on when recordings were made.

None of the four subspecies was characterized by any distinctive acoustic trait; however, subspecies *leucurus* and *gairdnerii* in the west had relatively lower maximum (HiF) and minimum (LoF) frequencies (3600.9 - 3623.7 Hz; 1189.0 - 1254.7 Hz, respectively) compared to the subspecies at the east (5782.5 - 5791.3 Hz; 1311.3 - 1394.2 Hz, respectively) (Table 3.2). As the eastern subspecies *pubescens* and *medianus* accounted for 73% of the sampled individuals, and the western subspecies for the rest, sampling bias might explain the differences. The MANOVA of the six spectro-temporal variables failed to detect any significant differences among the four subspecies (Wilk's $\lambda = 0.767$, $F = 0.635$, $P = 0.839$). Additionally, the discriminant function analysis with 317 calls and 42 individuals, assigned 26.2% of the calls and individuals to the correct subspecies. Individuals were closely grouped in the canonical functions, with no clear distinction among them (Fig 3.4b).

Finally, I found no significant correlations between latitude and maximum frequency (HiF - $r^2 = 0.085$, $P = 0.893$), minimum terminal frequency (LoF - $r^2 = 0.000004$, $P = 0.979$), or call duration (DelT - $r^2 = 0.059$, $P = 0.276$) (Fig 3.5a-c), which suggests that overall variation in the Downy Woodpecker pik call is not structured or changes gradually with latitude or longitude.

3.4 DISCUSSION

To my knowledge, this is the first quantitative and descriptive study of the Downy Woodpecker pik call in a geographic context. The pik call is a single element vocal trait peaking at ~ 3,750 ± 58 Hz (fundamental frequency) and averaging 0.034 ± 0.010 s in duration, corroborating an earlier study with very limited sampling across the species' range (Short 1982). In general, these results show no evidence of geographic variation or differentiation in the pik call for several common frequency and temporal variables; however, individual variation was evident, and predominant, when the variables were combined in multidimensional space (Fig 3.4).

These findings are in agreement with the predictions based on ecology and vocal behavior in which individual variation should be predominant over geographic variation because of the strong necessity of individual identity or specific habitat adaptations. The present results are in agreement with those of Dodenhoff (2002) where the pik call frequency and time variables varied a lot among individuals, and allowed individual recognition over years. In the present study I was not able to test differences in the same individual across years, however, one would expect that if individual recognition is key in territory defense or signaling fitness, variation between years should not prevail or hinder individual variation (Sung & Miller 2007).

The other challenge to understanding whether or not the pik call varies geographically is the multifunctional nature of the pik call (used in different behavioral

contexts from courtship to defense) (Winkler & Short 1978, Dodenhoff 2002). No studies have addressed if the Downy Woodpecker changes temporal or frequency characteristics of the pik call in order to transmit a specific message, for instance, using only high frequency notes to advertise an intruder in the territory or maybe using longer calls for mating. It is known that the frequency of call delivery can change (number of calls per minute), but almost nothing is known about call structure (Winkler & Short 1978). In other words, context dependent differences could confound geographic variation.

Latitude and geographic distance were not associated with any of the frequency and time variables investigated in this study. In practice, vocalizations in species with extensive geographic ranges have been used to demonstrate that song/call variation patterns occur in concert with genetic and plumage traits. For example in vocal learners such as passerines and hummingbirds (Irwin 2000, Irwin et al. 2008, Cadena & Cuervo 2010, González et al. 2011), nonetheless, vocalizations have been poor predictors of subspecific differences in non-vocal learners such as in the Bobwhite (*Colinus cristatus*) (Goldstein 1978).

While I had a larger number of samples from the eastern US, there were fewer from Canada and the western US. I did not find evidence of subspecific or clinal variation (south to north) in any of the acoustic parameters studied (Fig 3.5). Even with the sampling limitations, and not being able to control for sex, age, year, habitat or recording equipment, these results suggest that the pik call can not predict four of the subspecies currently recognized by the AOU, which is also in agreement with the genetic data in

chapter two. None of the four subspecies sampled here can be recognized or diagnosed based on the combination of the six call variables studied here. These conclusion need to be taken with caution for subspecies *gairdnerii* and *leucurus* where sampling was very limited (Table 3.2).

In general, my results are in conflict with general theoretical predictions based on body size and latitude (Browning 1997, Jackson & Ouellet 2002), subspecies or geographic distance. The present analyses suggest that the pik call is not a useful trait to study geographic variation because of the extensive individual variation, but it can be a powerful tool to study population dynamics and behavior at a local scale. It is recommended to complement the present study with more direct sampling, specifically in the west and north, in order to support this observation. Further studies on geographic variation should also focus on the rattle call or drumming, acoustic signals that more context-specific, and for which geographic variation has been suggested at least in other species of temperate woodpeckers (Winkler & Short 1978, Stark 2002).

Overall, this study highlighted a lack of representation of Downy Woodpecker pik calls in sound archives, with a high proportion of the available recordings represented by drums and rattle calls. There was also an overrepresentation of calls in the eastern United States and Canada, compared with the rest of the range, which probably reflects both a higher abundance of this species the east (Jackson & Ouellet 2002) or the historical accumulation of recordings in this part of the continent. More recordings are needed to

understand patterns of geographic variation in this and other poorly understood species with continental-range distribution.

Table 3.1. Summary of descriptive statistics for five frequency (Hz) and one temporal (s) measurements for 317 calls from 22 locations across USA and Canada (Fig 3.2). For abbreviations see Fig 3.2.

Locality	no. ind.	no. calls	LoF (Hz)	LoFSu (Hz)	HiF (Hz)	DelF (Hz)	DelT (s)	Q3 (Hz)
Alaska	1	12	1335.3	1875.9	3572.6	2237.2	0.02792	3158.1
Alberta	1	8	1317.6	1677.4	3458.1	2140.5	0.02425	3079.2
Arkansas	1	9	1536.2	1741.1	4056.9	2520.7	0.03178	3604.1
Colorado	1	9	1435.2	1916.5	4128.5	2693.3	0.03878	3703.7
Florida	1	9	1225.0	1631.9	4114.8	2889.8	0.04000	3694.1
Indiana	1	8	1726.4	2177.1	4365.1	2638.7	0.02913	3962.0
Kentucky	1	6	1299.3	1501.0	3477.6	2178.3	0.02617	3172.5
Manitoba	1	9	1063.0	1362.6	3841.2	2778.3	0.04178	3387.9
Maryland	2	13	1269.8	1620.1	3919.6	2649.7	0.03656	3488.3
Michigan	3	23	1227.5	1743.8	3816.8	2589.3	0.03128	3376.7
New Jersey	1	2	1333.2	1536.9	3851.6	2518.4	0.03300	3402.2
North Carolina	1	11	1373.5	1736.6	3725.5	2352.0	0.03191	3312.2
Ohio	5	40	1197.9	1663.6	3839.6	2641.7	0.03591	3403.3
Ontario	7	58	1371.1	1735.7	3693.5	2322.3	0.03084	3303.7
Oregon	1	5	1921.3	2153.3	3389.4	1468.0	0.03300	3152.4
Pennsylvania	1	6	1480.6	1696.4	3684.4	2203.7	0.04900	3287.3
South Carolina	1	7	1047.0	1430.7	3490.2	2443.1	0.03357	3026.9
Tennessee	3	19	1525.7	1840.0	3717.0	2191.2	0.02940	3359.6
Utah	4	26	1173.6	1547.9	3551.5	2377.9	0.03819	3159.2
Washington	3	18	1189.0	1599.1	3600.8	2411.8	0.03261	3221.1
Washington DC	1	11	1440.7	1670.3	3290.6	1849.8	0.03491	2897.1
West Virginia	1	8	1316.5	1619.8	3926.2	2609.6	0.03938	3499.1
Total	**42**	**317**						
Mean			**1354.8**	**1703.6**	**3750.6**	**2395.8**	**0.0341**	**3347.8**
SD			**202.75**	**200.30**	**268.17**	**320.85**	**0.01**	**247.11**

Table 3.2. Summary of descriptive statistics for five frequency (Hz) and one temporal (s) measurements for four Downy Woodpecker subspecies. For abbreviations see Fig 3.2 and subspecies Fig 1.2. SD: standard deviation.

Subspecies	N	LoF	LoFSu	HiF	DelF	DelT	Q3F
pubescens	7	1394.2	1723.0	3791.3	2397.1	0.0322	3388.0
	SD	219.3	177.8	230.4	261.3	0.00410	225.1
medianus	24	1311.3	1687.7	3782.5	2471.2	0.0340	3369.1
	SD	177.5	197.7	254.2	269.1	0.0060	230.1
leucurus	7	1254.7	1666.0	3623.7	2369.0	0.0348	3225.4
	SD	119.1	180.3	278.9	227.2	0.0070	259.3
gairdnerii	4	1189.0	1599.1	3600.9	2411.9	0.0326	3221.2
	SD	81.2	222.0	190.5	160.7	0.0041	187.8

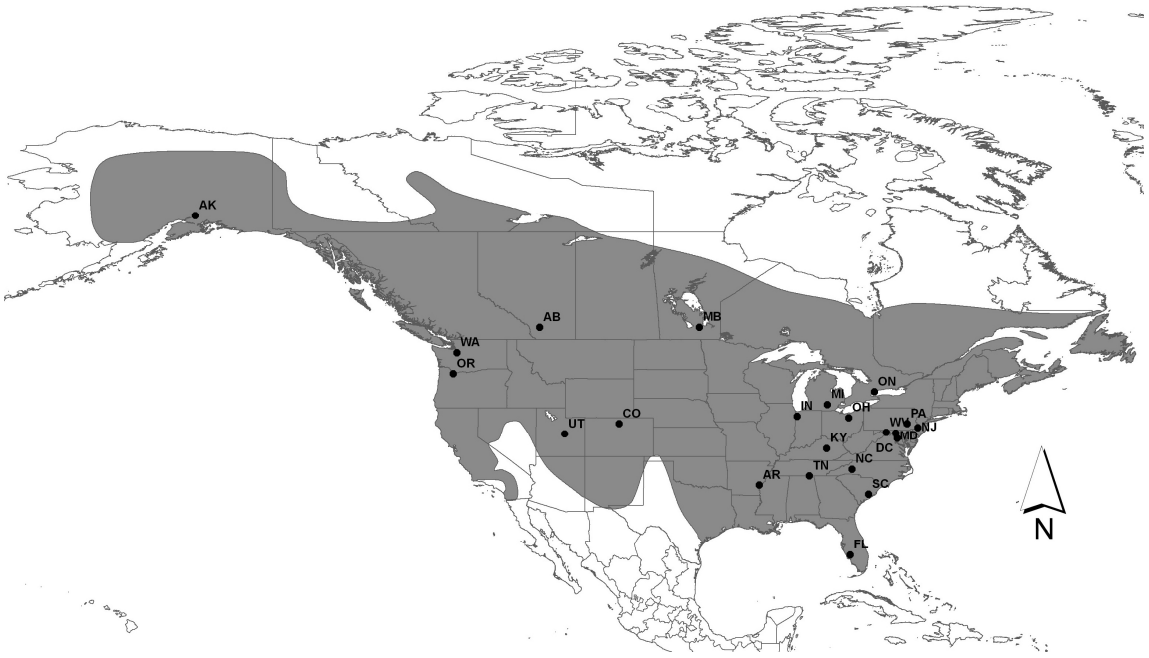


Figure 3.1. Map summarizing geographic samples of Downy Woodpecker recordings using in the study. Background gray area is the distribution of the species. Distributional layer was obtained from Ridgely et al. (2007).

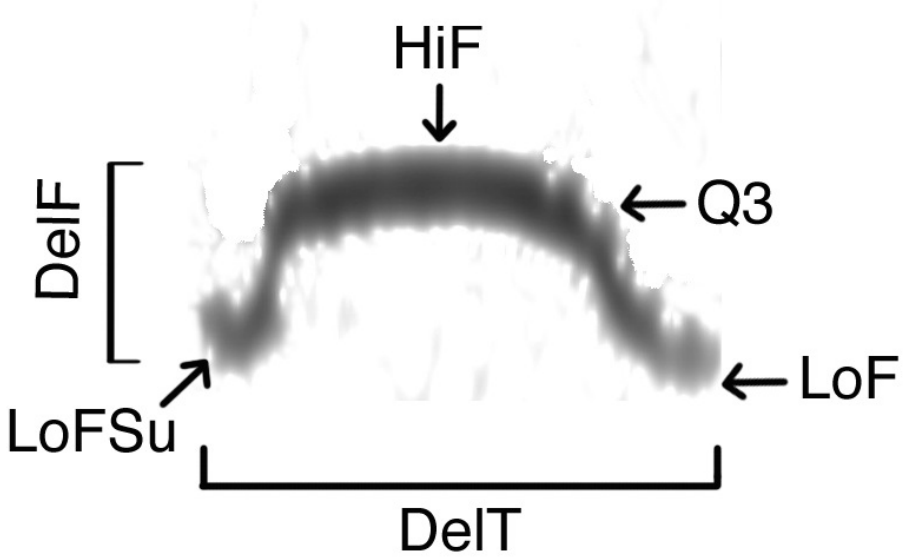


Figure 3.2. Typical pik call of Downy Woodpecker. Only the first harmonic is displayed. Frequency (Hz) and time (s) measurements were taken using an on-screen cursor in Raven 1.3. HiF = maximum frequency; LoF = minimum terminal frequency; LoFSu = minimum frequency at the initial part of the note; DelT = call duration; DelF = bandwidth and Q3 = 3rd quartile of the frequency.

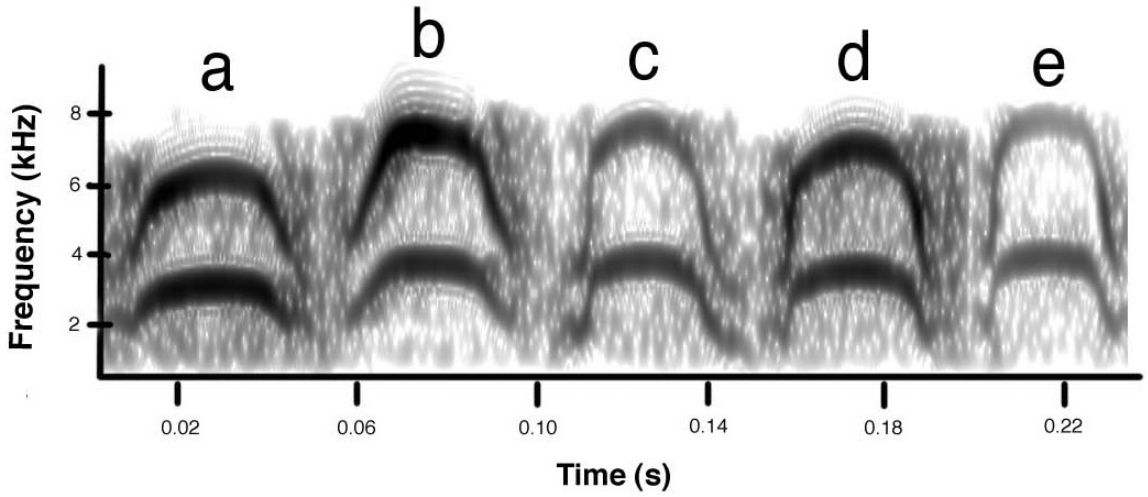
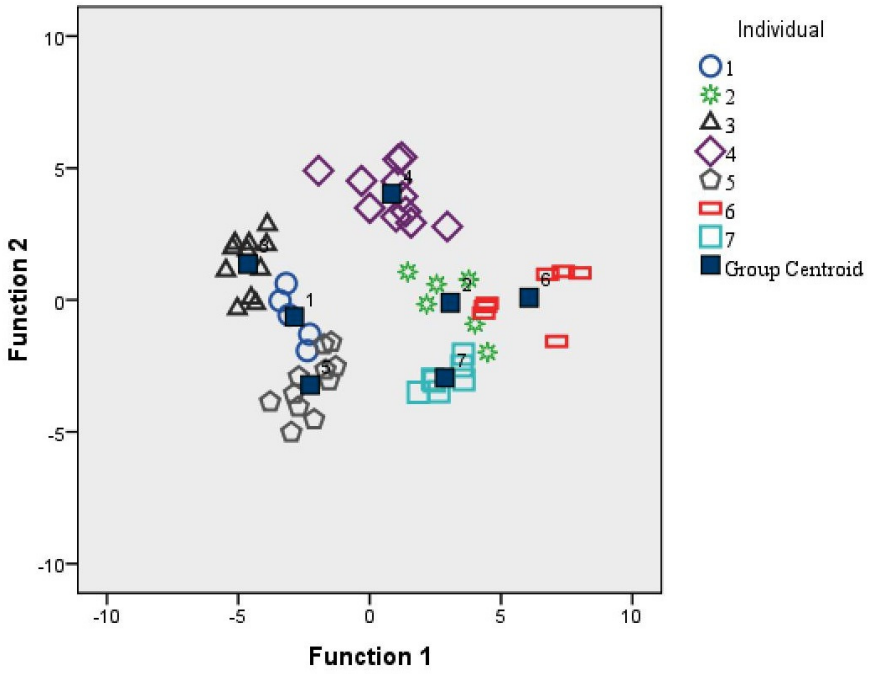
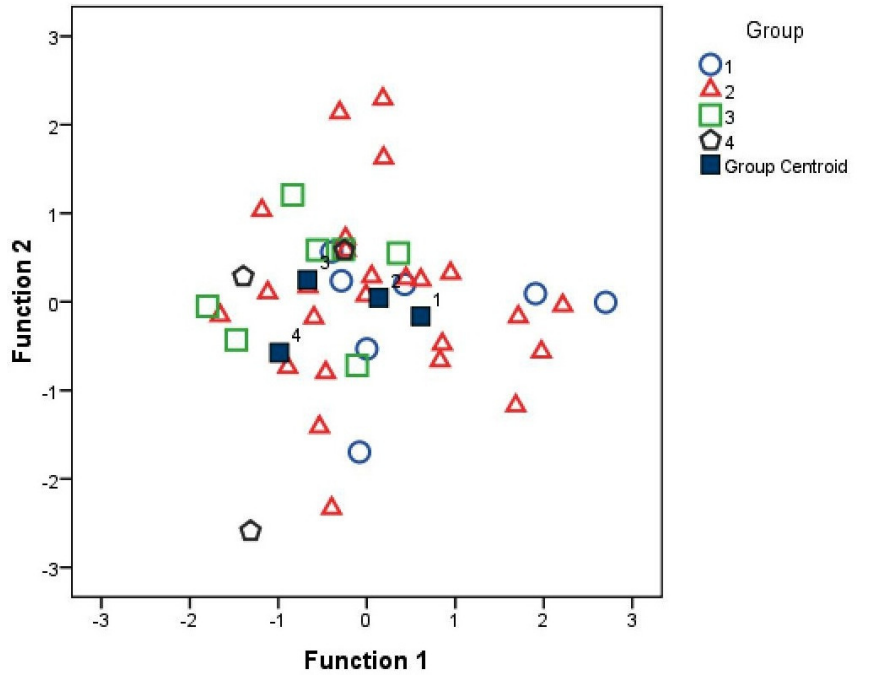


Figure 3.3. Sonograms of Downy Woodpecker pik calls from different locations: a. Ontario; b. Manitoba; c. West Virginia; d. Arkansas and e. Utah. Most calls were visually similar, slight variations can be seen in the length of the call and in the shape of the first harmonic.

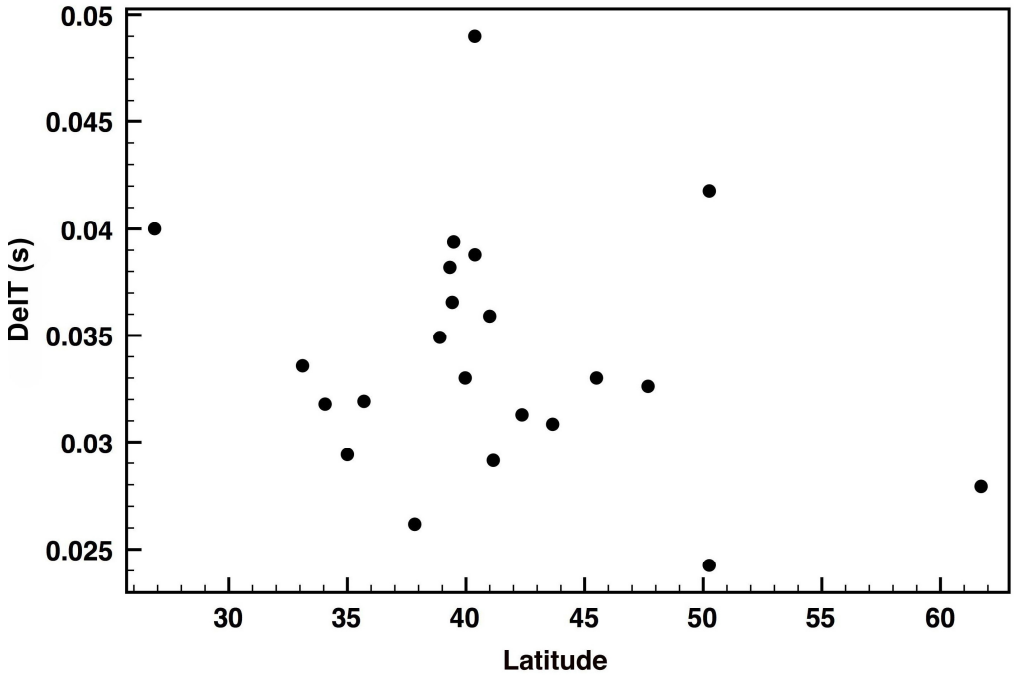


a.

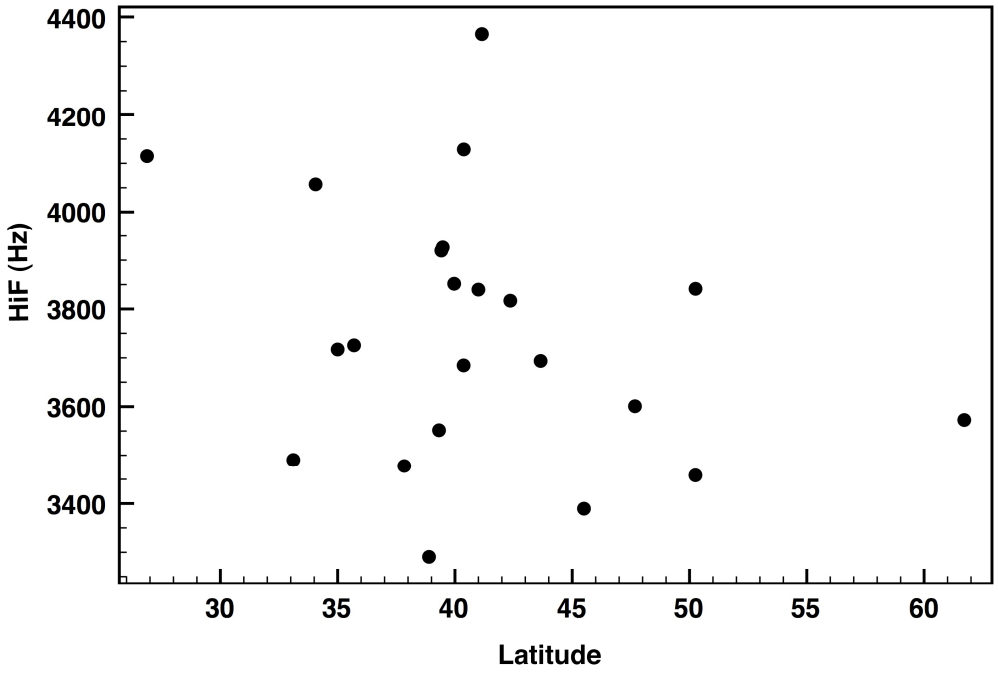


b.

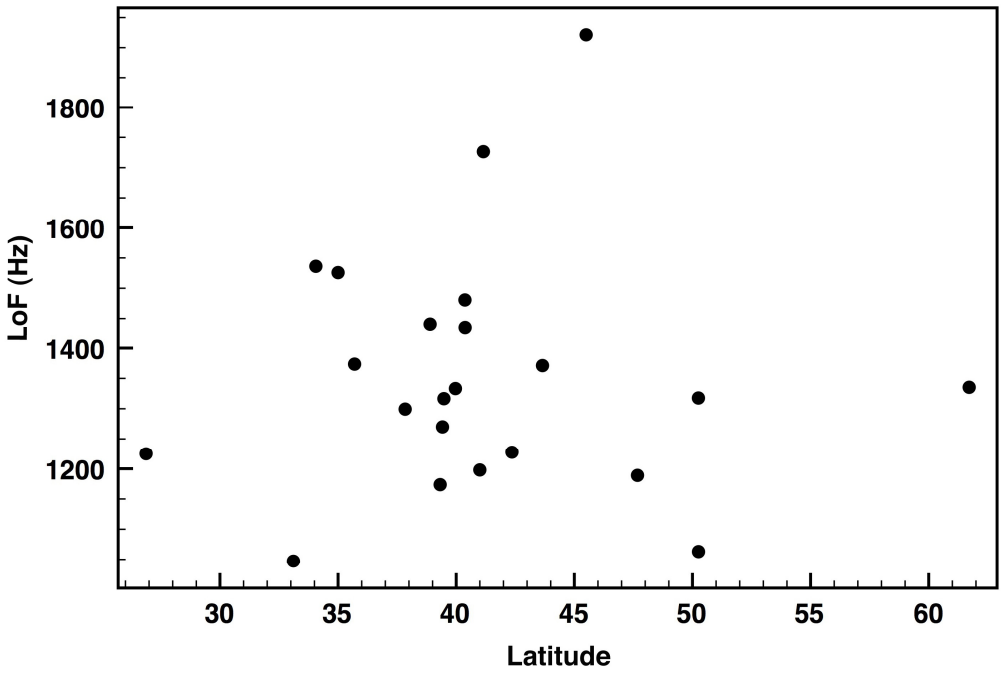
Figure 3.4. a. Distribution of calls belonging to seven individuals in Ontario. Individuals are represented by different symbols. b. Distribution individuals from four subspecies and centroids along the two first canonical discriminant functions. Circles represent *P. p. pubescens*; triangles *P. p. medianus*; squares *P. p. leucurus*, and pentagon *P. p. gairdnerii*.


0.05
0.045
0.04
0.035
0.03
0.025
DelT (s)
30
35
40
45
50
55
60
Latitude

a.


4400
4200
4000
3800
3600
3400
HiF (Hz)
30
35
40
45
50
55
60
Latitude

b.



c.

Figure 3.5. Relationship between latitude and a. call duration (DelT - r^2 = 0.059, P = 0.276); b. maximum frequency (HiF - r^2 = 0.085, P = 0.893); c. minimum terminal frequency (LoF - r^2 = 0.000004, P = 0.979), from 22 localities across the species' range.

CHAPTER FOUR

GENERAL CONCLUSIONS

The integration of several complementary approaches can be a powerful tool to disentangle biological questions in the fields of population genetics and biogeography (Knowles et al. 2007). For instance, studies attempting to reconstruct the demographic and population history of bird species since the last glacial maximum have traditionally used a single approach to reach their conclusions (Milá et al. 2000). In general, that single approach has been genetic information from a single locus, usually mitochondrial DNA. An increasing number of studies have shown that patterns resulting from single locus studies can lead to limited conclusions about past population history. In other words, the history of a locus sometimes reflect or not the history a species (Avise 2000).

One of the most complete studies on Australian Grass Finches in the genus *Poephila* demonstrated how analyses of multiple mitochondrial genes and nuclear introns yielded different divergence times among populations, and therefore different estimates on when populations split from the last common ancestor (Jennings & Edwards 2005). In northwestern North America, most studies have used a number of mitochondrial genes, (reviewed in Shaffer et al. 2010); a minority of these studies have used other genetic markers such as microsatellites, introns or other nuclear genes (probably because the lower resolution power of the nuclear genes and introns, and the small number of species specific developed microsatellite primers). There is an historical bias towards mtDNA

based studies, as such conclusions about potential refugia or dispersal ability of certain taxa might reflect the history of a locus rather than the species itself.

Although it is clear that mitochondrial genes by themselves can provide a lot of information in population genetics and phylogeography (Avise 2009), an explanation based on multiple molecular markers (such as microsatellite loci) is more realistic and biologically meaningful than the history proposed by a single marker. Accordingly, the problems go beyond using only a single type of molecular marker; there are also limitations when only one type of data (genetic) is used in biogeography. The geographical component in studies on biogeography is usually missing, and therefore conclusions are limited and restricted to one approach. The use of ecological niche modeling, acoustic variation and polymorphic genetic data have produced solid explanations of Pleistocene-related divergence and isolation in Satin Bowerbirds (*Ptilonorhynchus violaceus*), Mexican Jays (*Aphelocoma ultramarina*), Winter Wrens (*Troglodytes troglodytes*) and Brazilian frogs (genus *Phyllomedusa*) (Nicholls et al. 2006, McCormack 2008, Toews & Irwin 2008, Carnaval et al. 2009, Knowles 2009).

This is the first time that a study integrated three different approaches (multilocus genetic information, ecological niche modeling and bioacoustics) in order to understand the recent evolutionary history of a Nearctic woodpecker. The genetic information and predicted distribution model for the distribution of the Downy Woodpecker during LGM produced a more solid, complete scenario than Ball & Avise (1992). The former authors predicted that any further study should reveal low genetic differentiation between

different populations of the Downy Woodpecker; however, it totally underestimated the possibility that some geographic areas (eastern Rockies) contained genetically more differentiated populations, as I found using the microsatellite data. The present study also suggests that the geographic distribution of certain mtDNA haplotypes in the Rocky Mountains could be an indication of a more complex history in that part of the species' range.

The modeled distribution for the Downy Woodpecker during the LGM and contemporary distribution information is a limited, but a more solid and realistic approximation to the possible geographic areas occupied by this species ~21 kya. Without this approach, conclusions would have relied mostly on possible habitats or favorable conditions based on two-dimensional maps or patterns of genetic variation. Therefore, the use of maximum entropy for modeling 19 climatic variables combined with more than twenty thousand records from a reliable biological information database is a more powerful alternative. In short, the use of ENM has produced many insights into the understanding of the role of the Pleistocene climatic cycles in the distribution of different North American vertebrates (Waltari et al. 2007).

Assuming that the Downy Woodpecker pik call is innate, and that the number and origin of the samples used in this study is representative of the variation in this species, the results of the current bioacoustic analyses are partially congruent with the multilocus genetic data. It is congruent with the lack of extensive mtDNA differentiation between

populations, but is incongruent with the two differentiated groups supported by the microsatellite data.

The pik call is just one of the three main types of vocalizations in the Downy Woodpecker, hence, the results presented here are a glimpse of the potential geographic variation in acoustic signals in this species. Even more, an increase in sample size (including a better representation of sexes and ages) from different geographic areas (western parts of the range and the northwest) could change the current perspective based in a limited number of samples. For instance, geographic variation has been suggested in the pik call of the Hairy Woodpecker (Winkler & Short 1978), a species with quite a few similarities in plumage and distribution, but with much more population structure and complex phylogeographic history (Graham & Burg pers. comm.).

This new perspective on the demographic and population history of the Downy Woodpecker based on genetics, ecological niche modeling and bioacoustics offers new directions to formulate alternative hypothesis for example on refugial areas on the west of the Rockies and SE US. Another aspect that deserves further attention is that neither the genetic data nor the acoustic analyses of the pik call support any of the currently recognized subspecies (based on quantitative and qualitative plumage and body size variation) according to the AOU (Jackson & Ouellet 2002). There are clear incongruences between the patterns shown by the different available evidence, one could suggest that subspecific variation has not evolved in concert with the variation expressed by neutral genetic markers and the pik call (Ball & Avise 1992, Milá et al. 2007). For this

and most species it is still unknown whether or not climate or local ecological adaptations are responsible for morphological differences such as plumage variation or body size. The understanding of the genetic or environmental factors behind morphological variation in widely distributed organisms represent a potential area to study the origin of geographic variation and hence incipient speciation.

Future research into several aspects of bioacoustics and genetics will contribute to a better understanding of the role of Pleistocene glaciations on the demography of the Downy Woodpecker. One of them will be the use of novel more powerful molecular markers such as single nucleotide polymorphism to assess fine-scale signatures of population demographics or to uncover different evolutionary forces operating at a molecular level that might have shaped the genetic background of current populations (Brumfield et al. 2003). Another aspect is to have a more complete understanding of the male and female dispersal abilities and trends in different areas across the species' range, especially along the Rocky Mountains where more intricate patterns are present due to a more complex topography and possibly more stable climate over the last 2 millions years (Brunsfeld et al. 2001). A comparative study of post breeding season movements using satellite tracking or geolocators can help to answer several important questions about the seasonal movements of this woodpecker, and the possible consequences on the genetic structure of current populations.

More work needs to be done on the geographic variation of the pik call, specifically filling the sampling gaps for far north, west and southern parts of the range,

and also increasing the number of individuals in the already sampled areas with a better representation of sexes and ages (since sample size in this project was a limitation to make more accurate generalizations about geographic variation at a continental scale). Filling these sampling gaps might support either the genetic patterns exhibited by the microsatellites or mtDNA data or both. Pik call samples are missing in places such as BC (Revelstoke), Idaho, Montana and Oregon, where both mtDNA and microsatellites suggest a more complex history of isolation or gene flow. It will be also beneficial to explore how much geographic variation is exhibited in the rattle call and non-vocal drumming; both used in communication during the breeding season.

REFERENCES

Avise JC (2000) *Phylogeography: the history and formation of species*. Cambridge, M.A: Harvard University Press.

Avise JC (2004) *Molecular makers, natural history and evolution (second edition)*. Sinauer, Sunderland, M.A.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3-15.

Ball RM, Avise JC (1992) Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk*, **109**, 626-636.

Ball RM, Freeman S, James FC, Bermingham E, Avise JC (1988) Phylogeographic population structure of Red-winged Blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*, **85**, 1558-1562.

Baptista LF, Kroodsma DE (2001) *Foreword, Avian Bioacoustics, A tribute to Luis Baptista*., Pages 11-52 in J. Del Hoyo, A. Elliott, and J. Sargatal (eds.), Handbook of the Birds of the World. Volume 6. Lynx Edicions, Barcelona.

Beadell JS, Hyseni C, Abila PP, Azabo R, Enyaru JCK, Ouma JO, Mohammed YO, Okedi LM, Aksoy S, Caccobe A (2010) Phylogeography and population structure of *Glossina fuscipes* in Uganda: implications for controls of Tsetse. *PloS Neglected Tropical diseases*, **4**, e636.

Belkhir K, Borsa P, Chickhi L, Raufaste N, Bonhomme F (1996-2001) Genetix v4.02: logiciel sous Windows pour la génétique des populations. Laboratoire génome, populations, interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Benz BW, Robbins MB, Peterson TA (2006) Evolutionary history of woodpeckers and allies (Aves: Picidae): Placing key taxa on the phylogenetic tree. *Molecular Phylogenetics and Evolution*, **40**, 389-399.

Bretagnolle V, Genovis V (1997) Geographic variation in the call of the blue petrel: effects of sex and geographic scale. *Condor*, **99**, 985-989.

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of Tawny Owls (*Strix aluco*). *Molecular Ecology*, **16**, 3423-3437.

Brito, P, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, **135**, 439-455.

Browning RM (1995) Do Downy Woodpeckers migrate? *Journal of Field Ornithology*, **66**, 12-21.

Browning RM (1997) *Taxonomy of* Picoides pubescens *from the Pacific Northwest*. In: The era of Allan R. Phillips: A festschrift (R. W. Dickerman, ed.), pp. 25-33. Horizon Communications, Albuquerque, NM.

Brumfield R, Nickerson DA, Beerli P, Edwards SV (2003) The utility of single nucleotide polymorphism in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249-256.

Brunsfeld S, Sullivan J, Soltis D, Soltis P (2001) *Comparative phylogeography of northwestern North America: A synthesis*. In (J. Silvertown & J. Antonovics, eds.) Integrating ecological and evolutionary processes in a spatial context. Pp. 319-339. Blackwell Science, Oxford.

Burg TM, Gaston AJ, Winker K, Friesen VL (2005) Rapid divergence and postglacial colonization in western North American Steller's Jays (*Cyanocitta stelleri*). *Molecular Ecology*, **14**, 3745-3755.

Burg TM, Gaston AJ, Winker K, Friesen VL (2006) Effects of Pleistocene glaciations on population structure of North American Chestnut-backed Chickadees. *Molecular Ecology*, **15**, 2409-2419.

Burg TM, Croxall JP (2001) Global relationships among Black-browed and Gray-headed Albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. *Molecular Ecology*, **10**, 2647-2660.

Cadena CD, Cuervo AM (2010) Molecules, ecology, morphology and songs in concert: Can we know how many species is *Arremon torquatus* (Aves, Emberizidae)? *Biological Journal of the Linnean Society*, **99**, 152-176.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.

Charif RA, Waack AM, Strickman LM (2008) *Raven Pro 1.3 User's Manual*. Cornell Laboratory of Ornithology, Ithaca, NY.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1659.

Colbeck GJ, Gibbs LH, Marra PP, Hobson K, Webster MS (2008) Phylogeography of a widespread North American migratory songbird (*Setophaga ruticila*). *Journal of Heredity*, **99**, 453-463.

Cracraft J, Donoghue M (2004) *Assembling the tree of life*. Oxford University Press, USA.

Dingle C, Halfwerk W, Slabbekoorn H (2008) Habitat-dependent song divergence at subspecies level in the Gray-breasted Wood-wren. *Journal of Evolutionary Biology*, **21**, 1079-1089.

Dodenhoff DJ (2002) *An analysis of acoustic communication within social system of Downy Woodpeckers* (Picoides pubescens). PhD dissertation. Ohio State University, Ohio.

Dodenhoff DJ, Stark RD, Johnson EV (2001) Do woodpecker drums encode information for species recognition? *The Condor*, **103**,143-150.

Drovetski SV, Zink RM, Rohwer S, Fadeev IV, Nesterov EV, Karagodin I, Koblick EA, Red'kin YA (2004) Complex biogeographic history of a Holarctic passerine. *Proceedings of the Royal Society B*, **271**, 545-551.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Ellegren H, Carlson A, Stengber I (1999) Genetic structure and variability of White-backed Woodpecker (*Dendrocopus leucotos*) populations in northern Europe. *Hereditas*, **130**, 291-299.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, **5**, 435-445.

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

Excoffier L, Laval G, Schneider, S (2005) Arlequin ver. 3.0: An integrated software package for population genetic data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.

Fu YX (1997) Statistical neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147,** 915-925.

Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology,* **8**, 1513-1520.

Gaston KJ (2000) Global patterns in biodiversity. *Nature*, **405**, 220-227.

Godfrey WE (1986) *The birds of Canada*. Rev. ed. Natural Museum of Canada, Ottawa, Ontario.

González C, Ornelas JF, Rodríguez-Gutiérrez C (2011) Selection and geographic isolation influence hummingbird speciation: genetic, acoustic and morphological divergence in the Wedge-tailed Sabrewing (*Campylopterus curvipennis*). *BMC Evolutionary Biology*, **11**, 38.

Greenwood PJ, Harvey PH (1982) The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics*, **13**, 1–21.

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, **4**, 9.

Hey J, Machado CA (2003) The study of structured populations – new hope for a difficult and divided science. *Nature Review Genetics*, **4**, 535-543.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial recolonization of European biota. *Biological Journal of Linnean Society*, **68**, 87–112.

Hewitt, GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. *Nature review genetics*, **10**, 639-650.

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *PNAS*, **99**, 6112-6117.

Hutchison D, Templeton A (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898-1914.

Irwin DE (2000) Song variation in an avian ring species. *Evolution*, **54**, 998- 1010.

Irwin DE, Thimgan MP, Irwin JH (2008) Call divergence is correlated with geographic and genetic distance in Greenish Warblers (*Phylloscopus trochiloides*): A strong role for stochasticity in signal evolution? *Journal of Evolutionary Biology*, **21**, 435-448.

Isler ML, Isler PR, Whitney BM (1998) Use of vocalizations to establish species limits in Antbirds (Passeriformes: Thamnophilidae). *The Auk*, **15**, 577–590.

Jackson JA, Ouellet HR (2002) *Downy Woodpecker* (*Picoides pubescens*). In: The Birds of North America Online (A. Poole, Ed.). Ithaca, NJ. Cornell Lab of Ornithology.

James FC (1970) Geographic size variation in birds and its relationship to climate. *Ecology***, 51**, 365-390.

Jaramillo-Correa JP, Beaulieu J, Khasa DP, Bousquet J (2009) Inferring the past from the present phylogeographic structure of North American forest trees: Seeing the forest for the genes. *Canadian Journal of Forest Research*, **39**, 286–307.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424-429.

Jennings BW, Edwards SV (2005) Speciation history of Australian Grass Finches (*Poephila*) inferred from thirty genes trees. *Evolution*, **59**, 2033-2047.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of the Greater Prairie-chicken. *Molecular Ecology*, 1**2**, 3335-3347.

Kilham L (1962) Reproductive behavior of Downy Woodpeckers. *Condor*, **64**, 126-133.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561-576.

Klicka J, Zink RM (1997) The importance of recent ice ages in speciation: A failed paradigm. *Science*, **277**, 1666-1669.

Klicka J, Zink RM (1999) Pleistocene effects on North American songbird evolution. *Proceeding of the Royal Society B*, **266**, 695-700.

Knopp T, Merilä J (2008) The postglacial recolonization of northern Europe by *Rana arvalis* as revealed by microsatellite and mitochondrial DNA analyses. *Heredity*, **102**, 174-181.

Knowles LL, Carstens BC, Keat ML (2007) Coupling genetic and ecological-niche models to examine how past population distributions contributed to divergence. *Current Biology,* **17**, 940-946.

Knowles, LL (2009) Statistical phylogeography. *Annual Review in Ecology, Evolution and Systematics*, **40**, 593-612.

Koenig WD, Mumme RL (1987) *Population ecology of the cooperatively breeding Acorn Woodpecker*. Monographs in Population Biology 24, Princeton University Press, Princeton, NJ.

Koetz AH, Westcott DA, Congdon BC (2007) Geographical variation in song frequency and structure: the effects of vicariant isolation, habitat type and body size. *Animal Behavior*, **74**, 1573-1583.

Larsson LC, Charlier J, Laikre L, Ryman N (2009) Statistical power for detecting genetic divergence—organelle versus nuclear markers. *Conservation Genetics*, **10**, 1255–1264.

McCormack JE, Bowen BS, Smith TB (2008) Integrating paleoecology and genetics of bird populations in two sky island archipelagos. *BMC Biology*, **6**, 28.

Milá B, Girman DJ, Kimura M, Smith TB (2000) Genetic evidence for the effect of postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society B*, **267**, 1033-1040.

Milá B, Wayne RK (2006) Postglacial population expansion drives the evolution of long-distance migration in a songbird. *Evolution*, **60**, 2403-2409.

Milá B, McCormack JE, Castaneda G., Wayne RK, Smith TB (2007) Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus *Junco*. *Proceedings of the Royal Society B*, **274**, 2653-2660.

Miller EH (1986) Components of variation in nuptial calls of the Least Sandpiper (*Calidris minutilla*, Aves, Scolopacidae). *Systematic Zoology*, **35**, 400-413.

Miller EH (1996) *Acoustic differentiation and systematics of shorebirds*. In: Ecology and Evolution of Acoustic Communication in Birds (D. E. Kroodsma and E. H. Miller, eds.), pp. 241-257. Cornell University Press, Ithaca, New York.

Nicholls JA, Austin JJ, Moritz, C, Goldizen AW (2006) Genetic population structure and call variation in a passerine bird, the Satin Bowerbird, *Ptilonorhynchus violaceus*. *Evolution*, **60**, 1279-1290.

Nims BD, Vargas HF, Merkel J, Parker PG (2008) Low genetic diversity and lack of population structure in the endangered Galápagos penguin (*Spheniscus mendiculus*). *Conservation Genetics*, **9**, 1413-1420.

Oosterhout CV, Hutchinson WF, Willis DPM, Shipley P (2004) MICRO-CHECKER: Software for identify and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.

Otto-Bliesner BL, Brady EC, Clauzet G, Tomas R, Levis S, Kothavala Z (2006) Last glacial maximum and Holocene climate in CCSM3. *Journal of Climate*, **19**, 2526-2544.

Ouellet H (1977) *Biosystematics and ecology of* Picoides villosus *(L.) and* P. pubescens *(L.), (Aves : Picidae).* PhD thesis. McGill University.

Patel R, Mulder RA, Cardoso GC (2010) What makes vocalisation frequency an unreliable signal of body size in birds? A study on Black Swans. *Ethology*, **116**, 554-563.

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modeling*, **190**, 231-259.

Pielou EC (1991) *After the ice age: the return of life to glaciated North America*. The University of Chicago Press, USA.

Podos J, Warren PS (2007) The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, **37**, 403-458.

Pritchard, JK, Wen, X, Falush, D (2000) *Documentation for structure software, version 2.2.* http://www.pritch.bsd.uchicago.edu/software/structure22/readme.pdf.

Ray N, Adams JM (2001) A GIS-based vegetation map of the world at the last glacial maximum (25,000-15,000 BP). *Internet Archaeology*, **11**, 1-44.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): Population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248-249.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality test against population growth. *Molecular Biology and Evolution*, **19**, 2092-2100.

Ridgely RS, Allnutt TF, Brooks T, McNicol DK, Mehlman DW, Young BE, Zook JE (2007) *Digital distribution maps of the birds of the western hemisphere, version 3.0*. NatureServe, Arlington, Virginia, USA.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608-615.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496-2497.

Schlötterer C (2004) The evolution of molecular markers – just a matter of fashion? *Nature Review Genetics*, **5**, 63-69.

Scribner KT, Tablot SL, Pearce JM, Pierson BJ, Bollinger KS, Derksen DV (2003) Phylogeography of Canada Geese (*Branta canadensis*) in Western North America. *The Auk*, **120**, 889-907.

Shafer ABAA, Cullingham CI, Coté SD, Coltman DW (2010) Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589-4621.

Short LL (1982) *Woodpeckers of the world*. Delaware Museum of Natural History, Monograph series no. 4, Greenville, Delaware.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555-562.

Sorenson, MD, Ast JC, Dimcheff DE, Yuri T, Mindell, DP (1999) Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution*, **12**, 105-114.

SPSS for Windows. Rel. September. 2008. Chicago: SPSS Inc.

Stark RD (2002) *An analysis of eastern Nearctic woodpecker drums*. PhD thesis, Ohio State University.

Stockman, AK, Bond JE (2007) Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. *Molecular Ecology*, **16**, 3374-3392.

Sung HC, Miller EH (2007) Variation and individuality in display-flight calls of male Piping Plovers. *Condor*, **109**, 468-475.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596-1599.

Tautz D, Trick M, Dover GA (1986) Cryptic simplicity in DNA is a major source of genetic variation. *Nature*, **322,** 652-656.

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BF, De Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, Van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta MA, Peterson AT, Phillips OL, Williams SE (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.

Tremain SB, Swiston KA, Mennill DJ (2008) Seasonal variation in acoustic signals of Pileated Woodpecker. *Wilson Journal of Ornithology*, **120**, 499-504.

Välimäki K, Jaari S, Piha M, Pakkala T, Merilä J (2008) Isolation and characterization of 17 polymorphic microsatellite loci for the Three-toed Woodpecker (*Picoides tridactilus*). *Molecular Ecology Resources*, **8**, 1152-1154.

Vila M, Robles H, Ciudad C, Oleas PP, Baglione V (2008) Isolation and characterization of 12 microsatellite markers in the Middle-spotted Woodpecker (*Dendrocopus medius*). *Molecular Ecology Resources*, **8**, 415-417.

Waltari E, Hijmans RJ, Peterson AT, Nyári ÁS, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: Comparing phylogeographic and ecological niche model predictions. *PLoS ONE*, **2**, 1-11.

Wang S, Lewis CM Jr, Jakobsson M, Ramachandran S, Ray N, et al (2007) Genetic variation and population structure in Native Americans. *PLoS Genetics* **3**, e185.

Weins JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review in Ecology, Evolution and Systematic*, **36**, 519-539.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society B,* **271**, 1881-1887.

Wells PV (1983) Paleobiogeography of montane islands on the Great Basin since the last glaciopluvial. *Ecological Monographs*, **53**, 341-382.

Winkler H, Short LL (1978) A comparative analysis of acoustic signals in Pied Woodpeckers (Aves: Picoides). *Bulletin of the America Museum of Natural History*, **160**, 1-110.

Wormworth J, Mallon K (2006) *Bird species and climate change: The global status report version 1.0.*

Zink RM (1996) Comparative phylogeography in North American birds.
Evolution, **50**, 308–317.

Zink RM (1997) *Phylogeographic studies of North American birds*. In Avian molecular evolution and systematics (ed. D. P. Mindell), pp. 301-324. San Diego, CA: Academic Press.

APPENDICES

Appendix 1. List of 19 environmental variables from the WorldClim database used in ecological niche modeling.

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (* 100)

BIO4 = Temperature Seasonality (standard deviation *100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

BIO19 = Precipitation of Coldest Quarter

Appendix 2. Sound recordings from personal and institutional archives used in this study.

State/Prov	Location	Lat	Long	# Calls	Institution/Collection	Catalog/Track #	Recorder	Date
AK	Palmer to Sutton Alpine	61.713	-148.886	12	Paulo Pulgarin-R	USA_Alaska_PCPRAKA001	Paulo Pulgarin-R	Jun10
AB	Parkland	50.252	-113.657	8	MLNS Cornell	62775	Lucie R. Gunn	16Nov81
AR	Jack's Bay. White River	34.055	-91.182	9	Xenocanto	XC33615	Andrew Spencer	27Apr09
CO	Bonny State Park. Yuma Co.	40.377	-105.522	9	Nathan Pieplow	NDP2007-33-10-Fs-DOWOfemC-Bonny-9-10	Nathan Pieplow	9Oct07
FL	Charlotte	26.873	-81.860	9	MLNS Cornell	105322	Geoffrey A. Keller	2May94
IN	Newton	41.154	-87.289	8	MLNS Cornell	105643	Geoffrey A. Keller	25May95
KY		37.839	-84.270	6	Stokes field guide to bird songs - E region	Disk 2 - Track 32. page 33	Lang Elliott	17May88
MB	Balmoral	50.256	-97.321	9	Catherine Thexton	Downy woodpecker in Nature sounds CD	Catherine Thexton	1980
MD	Frederick	39.417	-77.208	5	MLNS Cornell	107232	Wilbur L. Hershberger	21Nov99
MD	Frederick	39.417	-77.200	8	MLNS Cornell	94232	Wilbur L. Hershberger	7Mar98
MI	Waterloo state recreation area	42.361	-84.191	6	Xenocanto	XC17149	Allen T. Chantier	6Jul07
MI	Waterloo state recreation area	42.361	-84.191	12	Allen T. Chantier collection	20090616_1135_MI-PHSGA_C1_ATC	Allen T. Chantier	16Jun09
MI	Metro Beach Metropark	42.576	-82.809	5	Xenocanto	XC16881	Allen T. Chantier	21Feb07
NJ	Burlington	39.967	-74.943	2	MLNS Cornell	6920	George B. Reynard	20Nov60
NC	Winstom Salem	35.699	-81.697	11	Paulo Pulgarin-R	USA_NorthCarolina_PCPRNC001	Paulo Pulgarin-R	29Apr10
OH	Blendon Woods Metropolitan Park	40.000	-83.000	8	Borror Lab Bioacoustics	1771	Donald J. Borror	3May56
OH	Blacklick Metropolitan Park	40.000	-83.000	8	Borror Lab Bioacoustics	3110	Donald J. Borror	16Mar58
OH	Blendon Woods Metropolitan Park	40.000	-83.000	11	Borror Lab Bioacoustics	4011	Donald J. Borror	15May59
OH	Blendon Woods Metropolitan Park	40.000	-83.000	6	Borror Lab Bioacoustics	6377	Donald J. Borror	17May63

OH	Fowler Woods Wildlife Area	41.000	-82.000	7	Borror Lab Bioacoustics	17455	Sandra L. Gaunt	22Apr90
ON	Essex	41.945	-82.526	5	MLNS Cornell	62756	William W. H. Gunn	12Dec56
ON	Backus woods. Port Rowan	42.850	-80.267	6	British library sound archive	W1CDR0001171 BD20	Tom Cosburn	31May01
ON	Halton Hills	43.640	-79.932	11	British library sound archive	W1CDR0000688 BD17	Tom Cosburn	19May97
ON	Toronto	43.653	-79.382	11	British library sound archive	W1CDR0000879 BD22	Tom Cosburn	13Feb78
ON	Boyd Conservation	43.812	-79.586	11	British library sound archive	W1CDR0000688 BD16	Tom Cosburn	9May96
ON	Haliburton	45.046	-78.509	7	British library sound archive	W1CDR0000879 BD24	Tom Cosburn	3Jul82
ON	Bruce	45.217	-81.717	7	MLNS Cornell	63938	Donald J. Kerr	9Nov92
OR	Willamette valley	45.500	-122.501	5	Bird Sounds of the Willamette Valley	DoWp_call01	Don Boucher	Unk
PA	Blue Marsh Lake Stilling Basin. Berks Co.	40.370	-76.032	6	Xenocanto	XC33508	James Eckert	26Apr09
SC	Fairlawn Plantation. 50 mi N. of Charleston	33.110	-79.980	7	MLNS Cornell	62757	William W. H. Gunn	6May57
TN	Bear Hollow Mountain WMA. Franklin Co.	34.999	-86.057	6	Xenocanto	XC15230	Chris Parrish	11May05
TN	Sewanee. Franklin	35.218	-85.922	2	Xenocanto	XC15229	Chris Parrish	12Oct06
TN	Knoxville	35.859	-84.095	11	Xenocanto	XC52422	Mike Nelson	Sep09
UT	Sawmill canyon. Beaver	38.525	-113.220	7	Kevin Colvert	DOWO 2005-14 Drum. F calls UT	Kevin J. Colver	May05
UT	Diamond Fork canyon along creek	39.321	-111.094	5	Kevin Colvert	DOWO 2000-02-05 male	Kevin J. Colver	2May00
UT	Diamond Fork canyon along creek	39.321	-111.094	8	Kevin Colvert	DOWO 2005-04-05 Drum. call UT	Kevin J. Colver	5Apr05
UT	Diamond Fork. 7 mi N. of Springville	40.000	-111.000	6	Borror Lab Bioacoustics	30595	Kevin J. Colver	1Apr00
WA		47.674	-122.122	8	Owl and the woodpecker	28 Downy Woodpecker_Stewart	Martyn Steward	Unk
WA	Redmond	47.674	-122.122	6	Birdsongs of the Northwest	51a-Downy_Woodpecker_5	Martyn Steward	Unk
WA	Samish Island. Skagit Co.	48.578	-122.540	4	Xenocanto	XC41856	Tayler Brooks	26Dec09

DC	Siegel Farm. Potomac River. Sycamore	38.895	-77.036	11	British library sound archive	W PICOIDES PUBESCENS R1 C1	Richard Ridgeway	3Jun62
WV	Berkeley	39.483	-78.167	8	MLNS Cornell	77258	Wilbur L. Hershberger	5Jan96