

THE POPULATION GENETICS AND PHYLOGEOGRAPHY OF THE HAIRY
WOODPECKER (*PICOIDES VILLOSUS*)

BRENDAN A. GRAHAM
BSc Biology, University of Regina, 2004

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

[MASTER OF SCIENCE]

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

©Brendan A. Graham, 2011

ABSTRACT

This thesis examines the effects that Pleistocene glaciation had on the population structure and contemporary genetic patterns of the hairy woodpecker (*Picoides villosus*). A combination of molecular markers, revealed reduced levels of gene flow among groups of hairy woodpeckers. Microsatellite analyses suggest barriers to gene flow have influenced contemporary population structure, with higher structure found in western North America where barriers to gene flow are more prevalent. MtDNA analyses revealed three distinct genetic lineages, two in North America and a third in Central America. Results indicate these lineages separated prior to the Wisconsin glaciation (~100 kya) and that contemporary population structure is the result of post-glacial expansion from multiple refugia following deglaciation. Current taxonomy recognizes 17 subspecies (Jackson *et al.*, 2002), but molecular analyses in this study do not support current subspecies designations.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Theresa Burg for all of her advice, help and patience during the last two years. Additionally I would like to thank my committee members: Hester Jiskoot and Tony Russell for their help during the completion of my degree. I would also like to thank Paulo Pulgarin-Restrepo, John Hindley, Kim Dohms and Linda Lait for their friendship and help in the field and the lab. Additional thanks go to Paulo for all his help with GIS and statistics during the completion of my thesis and I owe many thanks to Elysha Koran for not only her help in the field and even more so for her help in the lab. My friend Jeff Foster has helped me many times over the last few years, and I would like to thank him for all of his help with catching woodpecker last spring in Arizona. Finally I would like to thank my family for all of their support along the way and for encouraging and putting up with my interests in birds.

I am indebted to the following institutions for providing me with additional samples, as their help provided me with samples that I would not have been able to get without their generosity: American Museum of Natural History, Burke Museum, Columbia University Museum, Field Museum of Natural History, Louisiana State University Museum of Natural Sciences, University of Michigan Museum, New Brunswick Museum, Royal Alberta Museum, Royal British Columbia Museum, Royal Ontario Museum, Royal Museum of Saskatchewan, Smithsonian Museum, the University of Northern British Columbia and Queens University. I would also like to

thank Jennifer Pierson from Montana State University who provided me with additional samples.

I would like to thank Project feederwatch in Canada and the United States and their many participants who were kind enough to invite us into their yards to band birds and often feed us. Without their help we would not have been able to catch as many birds as we did. Finally I would like to thank all of the field assistants over the last four years who have helped with data collection. NSERC (TMB) and Alberta Innovates (TMB) funded this project.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
CHAPTER 1: General Introduction	
1.1 Background	1
1.2 Study species	3
1.3 Molecular markers	5
1.4 Study design	6
1.5 Thesis overview	7
1.6 Predictions	8
CHAPTER 2: Patterns of microsatellite variation in hairy woodpeckers (<i>Picoides villosus</i>) across their North American range	
2.1 Introduction	15
2.2 Methods	18
2.2.1 Sampling and DNA extraction	18
2.2.2 Microsatellite amplification	20
2.2.3 Statistical analyses	21
2.2.4 Genetic differentiation and population variation	21
2.2.5 Influence of barriers on genetic structure	25
2.3 Results	26
2.4 Discussion	31
2.4.1 Barriers	32
2.4.2 Effect of historic barriers and patterns of recolonization	35
2.5 Conclusions	37
CHAPTER 3: Post-glacial expansion of trees determines population structure for a sedentary woodpecker species	
3.1 Introduction	49
3.2 Methods	52
3.2.1 Sampling and DNA extraction	52
3.2.2 MtDNA amplification	53
3.2.3 Statistical analyses	54
3.2.4 Ecological niche modeling	58
3.3 Results	59
3.4 Discussion	62
3.4.1 Patterns of recolonization	63
3.4.2 Location of refugia	64
3.4.3 Recolonization from multiple refugia	66
3.4.4 Barriers	67
3.5 Conclusions	68
CHAPTER 4: General Discussion	
4.1 Introduction	78
4.1.1 Population structure	79

4.1.2 Refugia	80
4.1.3 Physical barriers	81
4.1.4 Patterns of recolonization	82
4.1.5 Biogeographic patterns	83
4.2 Future work	84
REFERENCES	86
APPENDIX 1: Locality information of samples	98
APPENDIX 2: Delaunay triangulation	110
APPENDIX 3: Summary of expected and observed heterozygosity	111
APPENDIX 4: Variable sites table	112
APPENDIX 5: Haplotype distribution chart	115
APPENDIX 6: BioClim variables for ecological niche modeling	117

LIST OF TABLES

Table 2.1: Microsatellite loci primers and PCR conditions	39
Table 2.2: Summary statistics for each loci per population and overall for each population and locus	40
Table 2.3: Pairwise F_{ST} values for population differentiation	41
Table 3.1: Genetic diversity and demographic statistics for hairy woodpecker populations	70
Table 3.2: MtDNA F_{ST} pairwise differences between populations and P values	71
Table 3.3: MtDNA F_{ST} pairwise differences between groups and P values	72

LIST OF FIGURES

Figure 1.1: Diagram of pioneer and phalanx models of recolonization	9
Figure 1.2: Map of elevation and forest vegetation in North America	10
Figure 1.3: Map of hairy woodpecker subspecies ranges	11
Figure 1.4: Total tree cover and extent of glaciation at 21, 16, 14, 11, 9, 6 and in 0.5 kya North America	12
Figure 2.1: Geographic distribution of sampling sites and patterns of population structure and barriers in North America for microsatellite analysis	42
Figure 2.2 a-c: Barplots for results from STRUCTURE for microsatellite analyses	45
Figure 2.3 a-c: Microsatellite isolation by distance analysis	46
Figure 3.1: Distribution of sampling sites and patterns of population structure for mtDNA analyses	73
Figure 3.2: Haplotype network for mtDNA analysis	75
Figure 3.3: Bayesian skyline plot for hairy woodpecker populations	76
Figure 3.4: Ecological niche modeling predictions for current distribution of hairy woodpeckers and distributions during LGM (~21 kya)	77

List of abbreviations

A - adenine
AK - Alaska
AMNH - American Museum of Natural History
AUC - area under curve
avg - average
AR - allelic richness
AZ - Arizona
BC - British Columbia
BCR - British Columbia Revelstoke
bp - base pairs
C - cytosine
CA - California
CAB - central Alberta
CBC - central British Columbia
CeOR - central Oregon
CMN - Canadian Museum of Nature
CoOR - coastal Oregon
CO - Colorado
CQC - central Quebec
CR - Costa Rica
D - Tajima's D
DNA - deoxyribonucleic acid
EDTA - ethylenediaminetetraacetic acid
EXO - exonuclease
FMC - Field Museum of Chicago
F_S - Fu's F_S
F_{ST} - Wright's fixation index
G - guanine
H - heavy strand of mitochondrial DNA
HCl - hydrochloric acid
H_d - Haplotype diversity
H_e - expected heterozygosity
H_o - observed heterozygosity
ID - Idaho
K - cluster
km - kilometres
kya - thousands of years ago
L - light strand of mitochondrial DNA
LA - Louisiana
LAB - Labrador
LGM - last glacial maximum
LSUMZ - Louisiana State University Museum of Zoology
MCMC - Monte Carlo Markov chain
ME - Maine
MI - Michigan

min - minute
MO - Missouri
MT - Montana
mtDNA - mitochondrial DNA
Mya - millions of years ago
My - million years
N - data missing at site; sample size
Na - number of alleles
NaCl - sodium chloride
NBC - northern British Columbia
NC - North Carolina
NCMNS - North Carolina Museum of Natural Sciences
NEOR - northeast Oregon
NL - Newfoundland
NS - Nova Scotia
ON - Ontario
OR - Oregon
PA - private alleles
PAN - Panama
PCA - principle components analysis
PCR - polymerase chain reaction
QC - Quebec
QCI - Queen Charlotte Islands
QU - Queens University
r - raggedness index, mutation rate
 r^2 - regression coefficient
 R_2 - Ramos-Onsins-Rozas test
RABM - Royal Alberta Museum
RBCM - Royal British Columbia Museum
RNBM - Royal New Brunswick Museum
RSKM - Royal Saskatchewan Museum
s - second
SAB - southern Alberta
SAP - shrimp alkaline phosphatase
SCA - southern California
SCCA - south central California
SD - South Dakota
SDS - sodium dodecyl sulfate
SK - Saskatchewan
SMITH - Smithsonian Museum
SOR - southern Oregon
SSK - southern Saskatchewan
T - thymine
 T_1 - annealing temperature 1
 T_2 - annealing temperature 2
taq - DNA taq polymerase

TE - Tris-HCl EDTA buffer
Tris-HCl - tris(hydroxymethyl)aminomethane
U - unit
UMI - University of Michigan
UNBC - University of Northern British Columbia
UCOM - University of Columbia Museum
US - United States
UT - Utah
UWBM - University of Washington Burke Museum
VI - Vancouver Island
WA - Washington
WI - Wisconsin
w/v - weight per volume
ybp - years before present
°C - degrees Celsius
 π - nucleotide diversity
 τ - tau
 δ - net nucleotide substitution rate

Chapter 1: General Introduction

1.1 Background

Phylogeography aims to understand the processes and mechanisms responsible for the distribution of genetic variation in a geographic context. How factors like environment, geography and geology affect species ecology and evolutionary pathways is one of the main questions phylogeography looks to answer (Avice, 2000). Molecular markers are becoming prominent tools for answering questions about population ecology and natural history, as molecular variation has the ability to detect and retain historical and current patterns of variation within populations. Similarly genetic analyses have been used to help identify geographic areas with high genetic diversity and biodiversity, and whether specific areas and habitats should be priorities for future conservation efforts (Moritz and Faith, 1998).

Many speciation events pre-dated the Pleistocene (1.78 Mya), but for avian taxa it has been contested whether speciation events for sister species occurred before or during the Pleistocene (Avice *et al.*, 1998a; Avice *et al.*, 1998b; Klicka and Zink, 1997; Weir and Schluter, 2004a). It is widely accepted that glaciations did promote genetic diversification for plant, bird and mammal species (Avice *et al.*, 1998a; Hewitt, 2000). During the Pleistocene ice ages, large ice sheets covered northern latitudes of Eurasia and North America, disrupting species' ranges by creating isolated, fragmented populations in ice-free areas known as refugia (Pielou, 1991). While a number of species are known to have gone extinct in North America

during glacial periods (Pielou, 1991), ice-free refugia played an important role by preserving genetic diversity and promoting genetic diversity within extant species through long term isolation of populations (Petit *et al.*, 2003).

Historically, pollen data were the primary source for determining the locations of ice-free refugia and estimating environmental and climatic conditions within these areas (Bennett and Provan, 2008). Within North America multiple refugia have been hypothesized, with pollen data supporting known refugia in Beringia, present day Alaska (Brubaker *et al.*, 2005) and south of the ice sheets in the United States (Whitlock and Bartlein, 1997; Williams, 2003), and contested refugia located along the Pacific and Atlantic Coasts near the Queen Charlotte Islands (QCI also known as Haida Gwaii) and Newfoundland, respectively (Warner *et al.*, 1982b; Pielou, 1991).

Contemporary genetic patterns and distributions are the product of colonization patterns following deglaciation (Johansen and Latta, 2003). Was colonization rapid with long distance dispersal resulting in pockets of isolated populations, preventing additional colonization by subsequent individuals (pioneer model; Figure 1.1) or was it more gradual and continuous resulting in little population structure with newly founded populations being very similar genetically to the source population (phalanx model; Hewitt, 1996)? For all species, colonization of previously glaciated areas was dependent on the availability of habitat, with genetic patterns reflecting the history of vegetation patterns within biogeographic regions (Brunsfeld *et al.*, 2001; Soltis *et al.*, 2006). Within North America we see distinct lineages east and west of the Rocky Mountains in many

species (Boulet and Gibbs, 2006), but genetic patterns in eastern and western North America are very different. Physical barriers, such as mountains (Figure 1.2), are more prominent in western North America, leading to distinct population structuring of taxa on either side of barriers such as the Cascade Mountains (Carstens *et al.*, 2004; Orange *et al.*, 1999). In contrast eastern North America has fewer barriers and population structure is weak, with populations distributed over a broad region sharing genetic similarities (for review see Soltis *et al.*, 2006).

1.2 Study species

Woodpeckers (Picidae) perform important roles within forested ecosystems and studying their population dynamics can provide important information on the environments they occupy (Virkkla, 2006). While previous studies have explored species phylogenies and relationships within the Picidae family (Webb and Moore, 2005; Zink *et al.*, 2002a; Zink *et al.*, 2002b), few have actually explored population structure (Ellegren *et al.*, 1999; Pierson, 2010; Pierson *et al.*, 2010). Past work has focused on identifying and classifying subspecies, but subspecies do not always reflect patterns of genetic variation and are poor indicators of overall population structure (Avice and Nelson, 1989; Crochet *et al.*, 2000; Zink, 2004).

The hairy woodpecker (*Picoides villosus*) is a year-round resident with limited dispersal and displays variation in plumage and morphological traits across its range. Commonly found at all elevations throughout most forested systems, its range extends from Alaska in the north to the highlands of Panama. Currently 17 subspecies (Figure 1.3) of hairy woodpeckers are recognized, with classifications

based on morphological, plumage and behavioural traits (Jackson *et al.*, 2002). Hairy woodpeckers are viewed as important species for forested systems, since they are primary cavity nesters and create cavities for other forest taxa to use (Saab *et al.*, 2004). Studying this species in close detail may help us to further understand the forest ecosystems they inhabit. One study by Topp and Winker (2008) used mitochondrial cytochrome b data from a limited number of samples (n=17) to look at genetic patterns in the Pacific Northwest, but genetic patterns and population structure across the entire range have yet to be studied.

Hairy woodpeckers are an ideal species for studying the effects of glaciation and contemporary genetic patterns, as their range includes areas previously covered by ice (Figure 1.4), as well as known refugia (Alaska and southern North America) and contested refugia (QCI and Newfoundland). Physical barriers, including the Alaska, Cascade and Rocky Mountain ranges and large bodies of water separating mainland subspecies from those found on islands (i.e. QCI and Newfoundland), delineate subspecies' ranges. The presence of distinct plumage and morphological traits (e.g., bill and tarsus length) across the hairy woodpecker's range suggests limited gene flow or rapid phenotypic adaptation. Sedentary populations are good species for exploring genetic patterns and will provide greater insight into historical gene flow, as these resident species will retain genetic patterns longer than migratory species due to limited dispersal. In addition, detailed molecular analyses of a widely distributed species whose range extends across multiple physical barriers will allow examination of the patterns and processes promoting genetic diversification of North American taxa.

1.3 Molecular markers

I used mitochondrial DNA (mtDNA) and neutral, nuclear markers (microsatellites) to answer questions about historical and contemporary patterns of genetic variation and population structure in hairy woodpeckers. Mitochondrial DNA because of its uncomplicated nature (no recombination) allows us to track matrilineal lines, and follow movements across geographic locales over long periods of time (millions of years). The control region is a large, non-coding region evolving at rates as high as 20%/million years and has been found to be highly variable in birds (Baker and Marshall, 1997). These characteristics make the control region a useful marker for studying population structure and exploring demographic events, whereas slower evolving protein coding regions are better suited for looking at phylogenies of different species (Milá *et al.*, 2007b; Zink and Barrowclough, 2008). In contrast to mtDNA, microsatellites are biparentally inherited, highly variable and evolve more quickly, making them ideal for looking at contemporary patterns where fine scale differences have occurred over shorter periods of time (Primmer *et al.*, 1996). Previous studies have demonstrated contrasting patterns between mtDNA and nuclear markers (Beadell *et al.*, 2010; Brito, 2007; Burg and Croxall, 2001; Haavie *et al.*, 2000), and the use of both markers in my study provides greater insight into both long term and short term population genetic structure in hairy woodpeckers.

1.4 Study design

To test patterns of colonization, I sampled individuals from across the contemporary range of the hairy woodpecker, including individuals from known and putative refugia, and areas previously covered by ice sheets. Sampling is more intensive in western North America and reflects the larger number of potential physical barriers (Figure 1.2) and higher morphological variation (Figure 1.3) in hairy woodpeckers.

I used an integrative approach to answer questions about post-glacial expansion and the effect of natural barriers on gene flow by incorporating coalescent theory to analyze genetic data, with information on current distribution, contemporary and historical vegetation distributions (Williams, 2003) and paleoclimate data (Otto-Bliesner *et al.*, 2006). Coalescent models are becoming more widely used for genetic analysis as they provide detailed estimates about population size and growth, and gene flow. These models therefore have the potential to solve complex histories (Knowles, 2009) providing genetic analysis with the necessary tools for rigorous statistical analysis.

Combining genetic analyses with species' distribution data is a powerful method for detecting biogeographical patterns and inferring their causes (Richards *et al.*, 2007). Understanding historical vegetation distribution and paleoclimatic conditions allows us to better understand how historical events have shaped contemporary patterns and distributions (Hugall *et al.*, 2002). For my own study, the incorporation of genetic data with current species' distribution and paleoclimate data will help to identify geographical features across the hairy woodpeckers' range

that act as barriers to gene flow, and to better estimate their historical range and locations which served as Pleistocene refugia (McCormack *et al.*, 2008; Peterson *et al.*, 2004).

1.5 Thesis overview

The first data chapter of my thesis will focus on the contemporary population structure of hairy woodpeckers in North America. Using microsatellite data I will look at the role physical barriers like mountain ranges, large expanses of open water and discontinuities in habitat have had in shaping population structure. If physical barriers are restricting gene flow, then we will expect birds on either side of these barriers to be genetically distinct from each other. I also look at how colonization of previously glaciated areas and historical barriers have affected contemporary genetic patterns.

The second data chapter uses mtDNA data to look at how Pleistocene glaciation patterns affected genetic patterns in hairy woodpeckers. I use paleodistribution models to estimate past distributions during the LGM (Last Glacial Maximum) ~21 kya and compare model predictions to contemporary genetic patterns and range distribution. Incorporating ecological niche modeling will allow me to better estimate locations of ice-free refugia used by hairy woodpeckers during the LGM and test two models of recolonization (pioneer vs. phalanx). A better understanding of refugia locations and the model of recolonization can then be used to determine recolonization routes used by hairy woodpeckers following

deglaciation and understand contemporary genetic patterns for a widely distributed species.

In the final chapter I compare congruencies and discontinuities in genetic patterns between the two different markers. Comparing the results from the two markers allows me to better evaluate overall genetic patterns and population structure for hairy woodpeckers and explain how genetic patterns for a sedentary species compare to overall biogeographic patterns in North America.

1.6 Predictions

I predict that historical recolonization patterns following the phalanx model have shaped contemporary genetic patterns in hairy woodpeckers. As woodpeckers are sedentary, depend on trees and forests for feeding and nesting, movements out of refugia were likely slow and gradual. As hairy woodpeckers' current range includes both known and contested refugia (Pielou, 1991) and this species shows distinct plumage and morphological differences across their range, I predict post-glacial expansion occurred from multiple refugia. I predict that physical barriers, in particular discontinuities in forested habitat, have played a role in shaping population structure and populations separated by breaks in forested habitat will be genetically distinct from each other. Endemic subspecies of hairy woodpeckers are recognized for two offshore islands (QCI and Newfoundland; Jackson *et al.*, 2002). Phenotypic differences likely indicate isolation of populations and I predict restricted gene flow between mainland and offshore island populations.

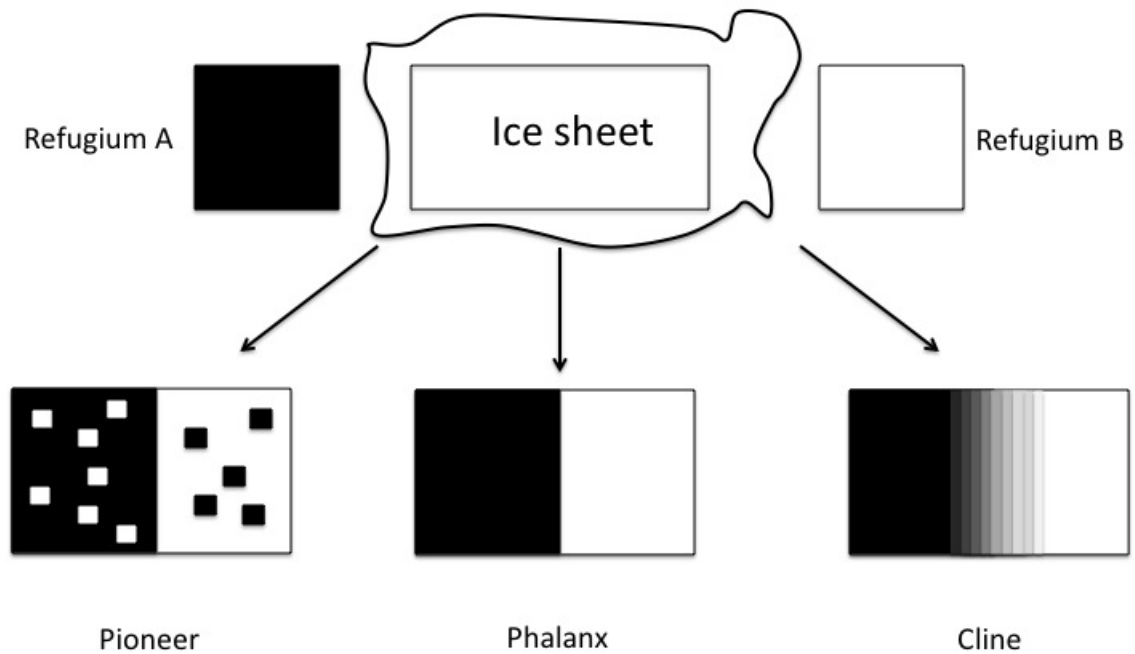


Figure 1.1: Diagram illustrating the two different models of recolonization from two hypothetical glacial refugia separated by an ice sheet during the LGM, as proposed by Hewitt (1996). Recolonization into previously glaciated areas following the pioneer model colonization is rapid with long distance dispersal resulting in multiple isolated populations being established. Black dots in pioneer model represent populations most similar to refugium A, while white dots indicate populations most similar to refugium B. Recolonization following the phalanx results in gradual and short distance dispersal resulting in two genetically separate populations being established. A cline shows gradual distribution following contact between two distinct populations where shades of grey indicate mixing of individuals from refugium A (black) and B (white) (redrawn from Johansen and Latta, 2003).

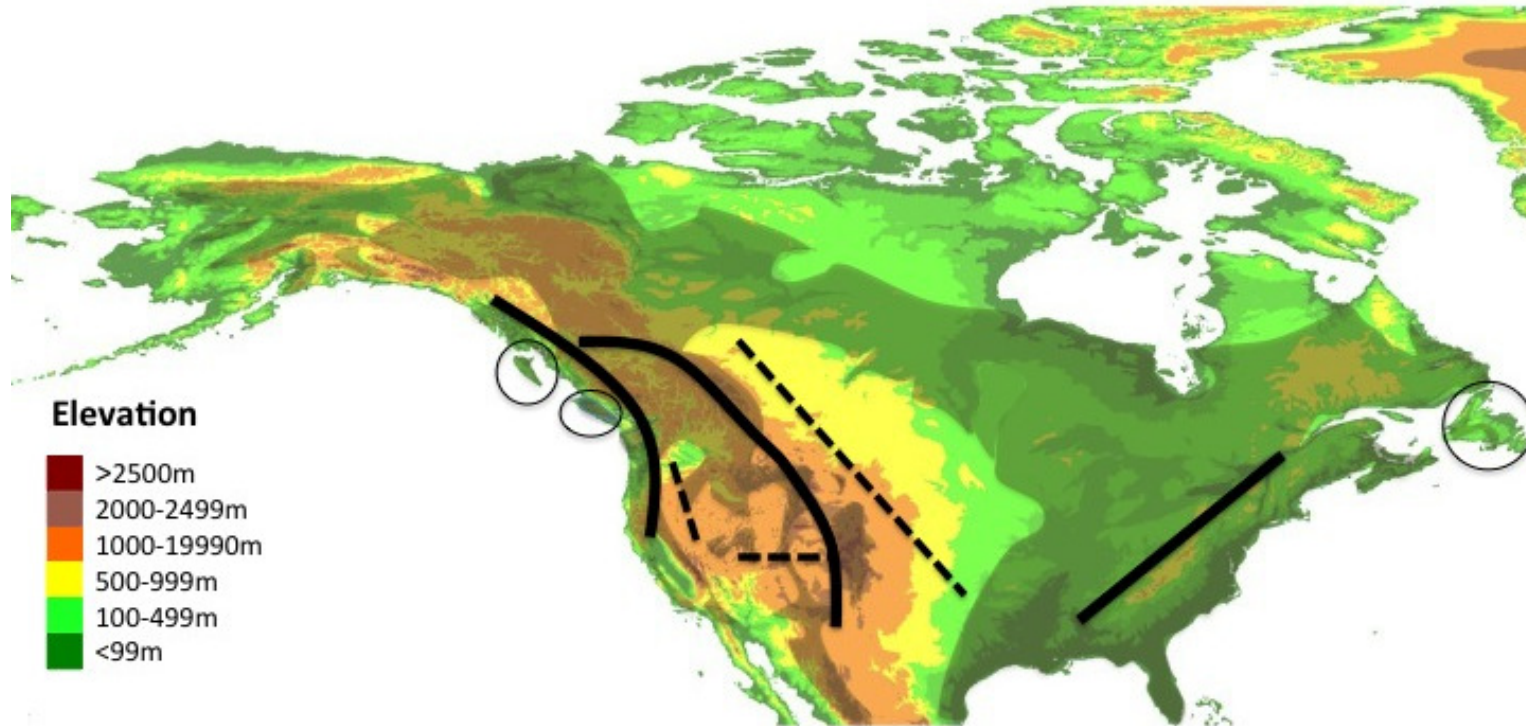


Figure 1.2: Map of North America showing elevation and forest tree species vegetation. Overlaid brown shading represents distribution of forested vegetation (tree species). Higher elevations (solid black lines) may act as potential barriers to gene flow, as well as open expanses of water separating islands from mainland North America. Breaks in contiguous forest (dashed black lines) may also act as barriers to gene flow for forest dependant species. Open expanses of water separating islands from continental North America are also potential barriers to gene flow (represented by open black circles). Map adapted from information available through global forest watch (www.globalforestwatch.org).

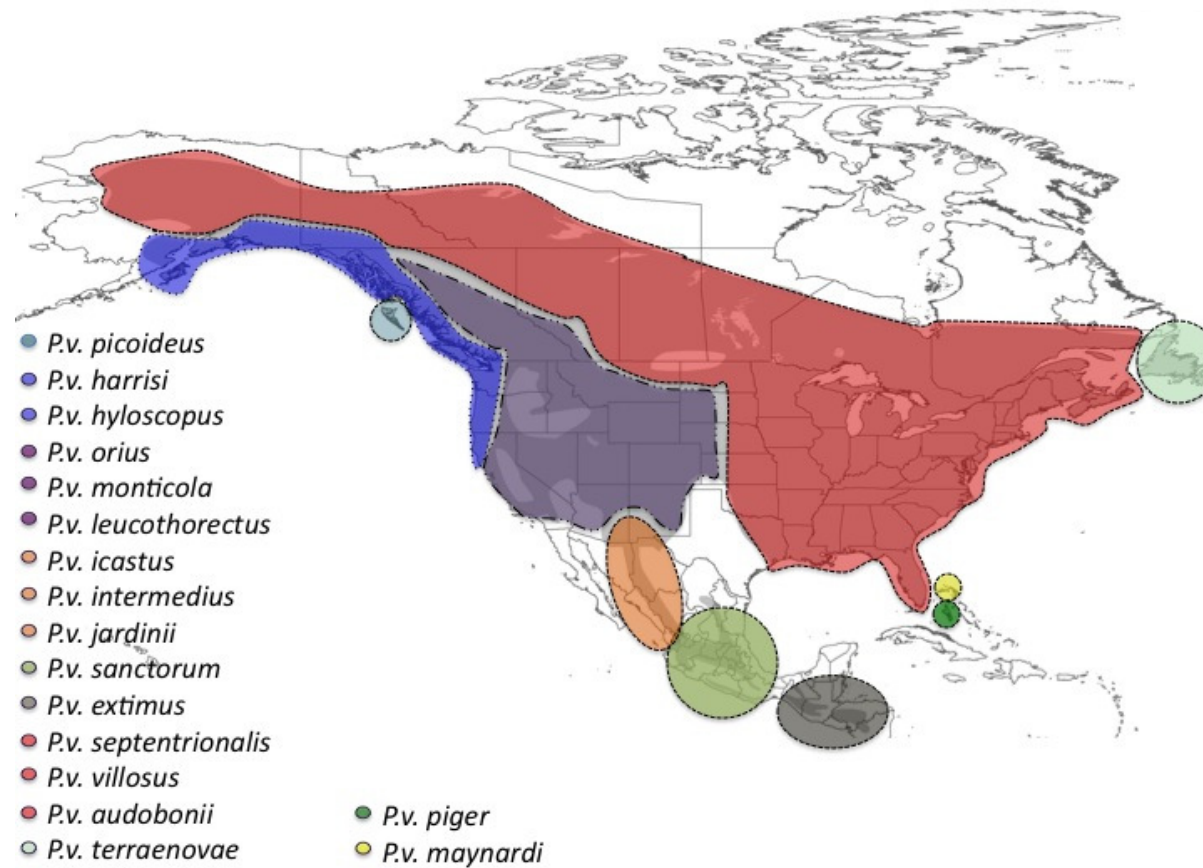


Figure 1.3: Map showing distribution of hairy woodpecker subspecies across their range. As subspecies' ranges overlap, multiple subspecies have been given the same colour since it is difficult to delineate where subspecies' ranges begin and end: Pacific Coast (blue; *P. v. harrisi* and *P. v. hyloscopus*), western (purple; *P. v. orius*, *P. v. monticola*, *P. v. leucothorectus*), eastern (red; *P. v. septentrionalis*, *P. v. villosus*, and *P. v. audobonii*) and Mexican (orange; *P. v. icastus*, *P. v. intermedius* and *P. v. jardiinii*). Range map modified from Ridgely *et al.* (2007) and subspecies' ranges adapted from Jackson *et al.* (2002).

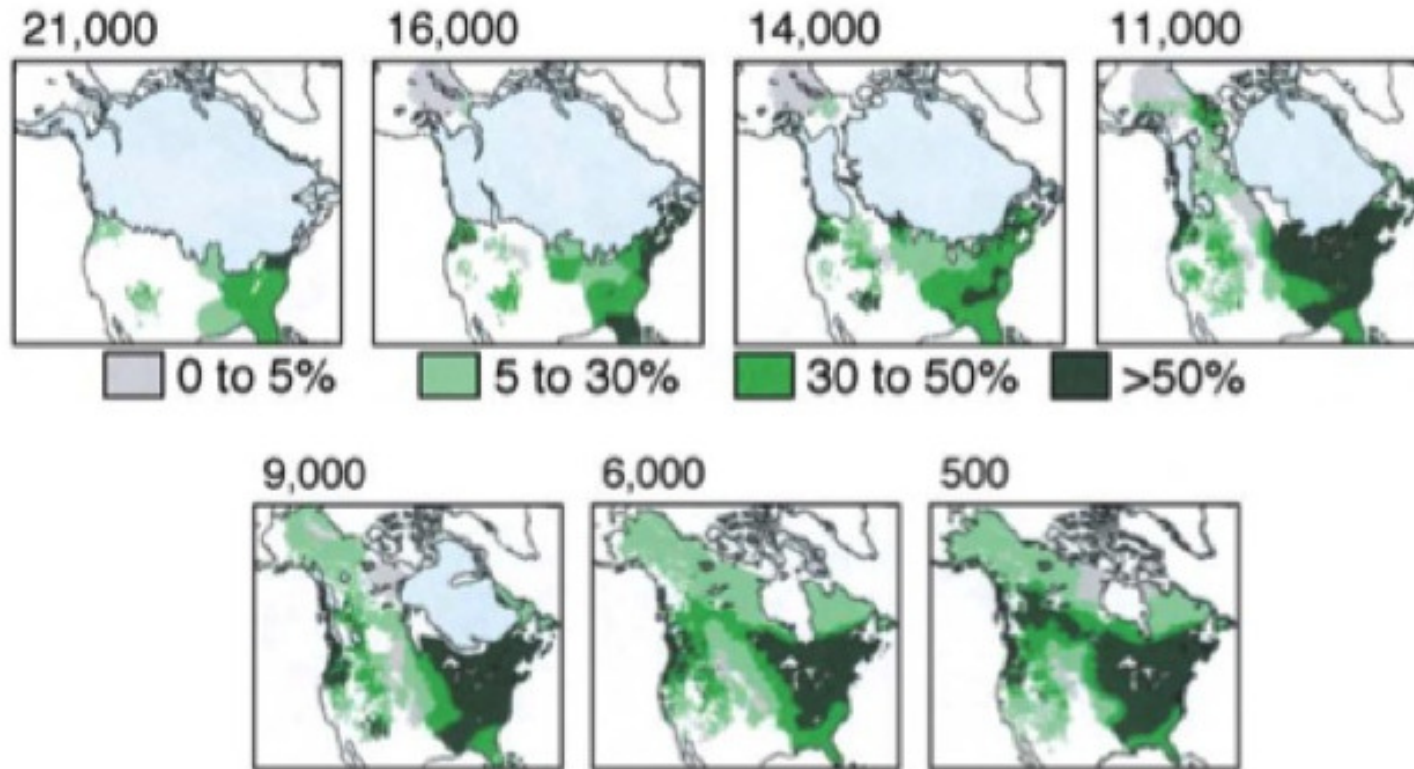


Figure 1.4: Total forest cover and extent of glaciation for North America for 21, 16, 14, 11, 9, 6 and 0.5 kya. Colours (Gray, light green, green and dark green) correspond to total tree cover at each time interval.¹

¹Reprinted from *Global and Planetary Change* 35, J. W. Williams, Variations in tree cover in North America since the last glacial maximum, 1-23, 2002, with permission from Elsevier.

Chapter 2: Patterns of microsatellite variation in hairy woodpeckers (*Picoides villosus*) across their North American range

B.A. Graham^{1,2} and T.M. Burg¹

prepared as manuscript for submission

¹Department of Biological Sciences, University of Lethbridge, 4401 University Drive,
Lethbridge AB, T1K 3M4

²corresponding author: b.graham@uleth.ca

Abstract

Hairy woodpeckers are a common, year round resident with distinct plumage and morphological variation across North America. We genotyped 314 individuals at six microsatellite loci to explore patterns of population genetic variation and test whether physical barriers are influencing genetic patterns across their range. We found two main genetic clusters representing a distinct east - west split that corresponds to the eastern edge of the Great Plains in the south and a western contact zone in the northern and central Rocky Mountains. Additional structure was found within each group; in the east: boreal-eastern US, and Alaska; and in the west: Queen Charlotte Islands, Cascades, interior west and southeastern Rocky Mountains. Genetic differentiation in the mountainous regions does not seem to be the result of mountains acting as a barrier, but is associated with breaks in contiguous forest habitat restricting gene flow between areas. Within the main east group we found very little genetic structuring due to the presence of contiguous forested habitat. Our results suggest gene flow is restricted between: the Queen Charlotte Islands, Cascade Mountains, interior west, southeastern Rocky Mountains and eastern North America. Historical barriers like ice sheets and recolonization patterns following the last glacial maximum (LGM) have also affected patterns of genetic variation. Contemporary genetic patterns are reflective of recolonization patterns following the phalanx model from ice-free refugia following deglaciation.

keywords: hairy woodpecker, post-glacial colonization, multiple refugia, dispersal barriers, contiguous forest, microsatellite, paleogeography

2.1 Introduction

Advances in the availability of molecular markers have aided our understanding of how landscapes affect population genetic patterns, variation and gene flow throughout a species' distribution (Manel *et al.*, 2003). Past studies have shown that barriers to gene flow caused by topographical features (Keyghobadi *et al.*, 1999), unsuitable habitat (Piertney *et al.*, 1998) and human mediated land changes (Gerlach and Musolf, 2000) all affect contemporary genetic patterns. Exploring environmental conditions and landscape features will provide further insight into understanding the processes that influence gene flow and how genetic variation within species are produced.

In North America previous studies have examined the role of past glacial events in shaping population structure and promoting genetic diversification for both plant (reviewed in Jarmillo-Correa *et al.*, 2009) and animal species (Avisé *et al.*, 1998b; Burg *et al.*, 2006; Hewitt, 2000). During the last glacial maximum (LGM), most of northern North America north of 48°N was covered by ice sheets (Pielou, 1991). The last glacial was disruptive and fragmented species' ranges, restricting individuals and populations to ice-free areas known as refugia (Klicka and Zink, 1999). Putative refugia during the last glacial maximum included Alaska and southern North America, while the Queen Charlotte Islands (QCI), also known as Haida Gwaii, and Newfoundland are contested to have been ice-free (Pielou, 1991). Populations expanded from these refugia following the melting of the ice sheets. How populations colonized previously glaciated areas following deglaciation remains an important question, as the mode of recolonization affects contemporary

genetic patterns (Johansen and Latta, 2003). Was colonization rapid with long distance dispersal ahead of the leading edge, resulting in pockets of isolated populations (pioneer model) or was it more gradual and continuous with newly founded populations being very similar genetically to the source population (phalanx model; Hewitt, 1996)?

The hairy woodpecker (*Picoides villosus*) is a common year-round resident with distance limited dispersal (Jackson *et al.*, 2002) found in most forest and woodland habitats at both high and low elevations, whose range extends north-south from Alaska to the highlands of Panama; and east-west from the Pacific Coast to the Atlantic Coast (Figures 1.2 and 2.1). Currently seventeen subspecies classified by plumage, morphological and behavioural characteristics are recognized, with eleven subspecies occurring in North America (Jackson *et al.*, 2002). Phenotypic variation in plumage varies across its range. The presence of distinct plumage and morphological traits suggests limited dispersal or rapid phenotypic adaptation. Past studies have explored phenotypic variation (Miller *et al.*, 1999), habitat selection (Ripper *et al.*, 2007) and genetic variation (Topp and Winker, 2008; Pierson *et al.*, 2010) in hairy woodpeckers. Genetic studies have been limited to small portions of the hairy woodpecker's range and it remains to be seen if genetic patterns are indicative of rangewide genetic patterns or exclusive to the examined areas.

The hairy woodpecker's range is delineated by a number of physical barriers including the Cascade and Rocky Mountain ranges, and large bodies of water separating continental subspecies from those found on islands (i.e. QCI and Newfoundland; Jackson *et al.*, 2002). Studies of migratory birds have shown genetic

differences between populations on either side of the Rocky Mountains (Boulet and Gibbs, 2006; Hull *et al.*, 2008), but processes creating phylogeographic patterns remain poorly understood because of the confounding effects of long distance migration and barrier-mediated dispersal. Unlike migratory species, hairy woodpeckers remain in their breeding areas year round and do not undergo true migration (Ouellet, 1977). Banding data suggest sedentary behavior with 97% of all banding recoveries having been recaptured <40 km from their original capture location (Jackson *et al.*, 2002). Studying sedentary populations will provide greater insight into historical gene flow, as resident birds will retain genetic patterns longer due to limited dispersal. Detailed molecular analyses will provide information about the population structure and genetics of a widespread, geographically variable species, which will not only provide greater insight into the evolutionary patterns and processes of this species, but will also help increase our understanding of the genetic patterns of other organisms in North America.

We used genetic analyses to examine the rangewide population genetic patterns in hairy woodpeckers. With a prevalence of physical barriers throughout the range of the hairy woodpecker (Figure 1.2), we explored the role that contemporary physical barriers have had in shaping the observed genetic patterns. If physical barriers like mountains and waterways are restricting gene flow, then we predict populations on either side of physical barriers will be genetically distinct from each other. Due to North America's glacial history, we also address how historical barriers like paleo ice sheets and recolonization patterns following deglaciation have potentially influenced contemporary genetic patterns for hairy

woodpeckers. If recolonization of previously glaciated areas followed the pioneer model, we will expect to see genetically isolated populations, whereas if hairy woodpeckers moved via short distance dispersal only, phalanx model, then we will see little genetic structure in previously glaciated areas (phalanx model; Hewitt, 1996). Additionally post-glacial population expansion could produce genetic structure if hairy woodpeckers were isolated in multiple refugia as has been shown for other bird species (Soltis *et al.*, 1997; Weir and Schluter, 2004b). As woodpeckers are primarily sedentary and show morphological variation (Jackson *et al.*, 2002), we predict decreased gene flow between populations and high population structure. We also predict that expansion from ice-free refugia followed the phalanx model due to the sedentary nature of hairy woodpeckers and their reliance on forests for food and nesting, woodpeckers likely moved short distances following deglaciation.

2.2 Methods

2.2.1 Sampling and DNA extraction

Blood and tissue samples were obtained from birds from 30 sampling sites (hereafter referred to as 30 populations) covering the North American range of the hairy woodpecker (Figure 2.1, Appendix 1). We attempted to restrict sampling to a 30 km radius for each sampling location. We used mist nets to capture birds during the breeding season from 2007 to 2010 and collected blood samples (~100 μ l of blood) from the brachial vein. All blood samples were preserved in 95% ethanol. Tissue samples were also obtained from collections (American Museum of Natural

History, Burke, Field, New Brunswick, Royal British Columbia, Royal Alberta, Royal Ontario, Royal Saskatchewan and Smithsonian Museums and University of Northern British Columbia and Acadia University; Appendix 1) to supplement our field sampling and add new sampling sites. All museum samples were from birds collected during the breeding season over the last 20 years to ensure that we analyzed contemporary genetic patterns. Supplemental museum samples were obtained from the same geographic points as field sampling points where available. When they were not available, we used samples from adjacent counties, and grouped these samples with field samples, attempting to group samples over as small an area as possible. Additional sites (e.g. Michigan) were comprised completely of museum samples using the same criteria (breeding season, small geographic area and last 20 years). By limiting our samples to birds from the breeding season, we tried to avoid sampling birds atypical of each population, since hairy woodpeckers have been reported to move short distances (although no studies have documented exactly how far they move) during the late fall and early winter (Jackson *et al.*, 2002).

DNA was extracted from blood and tissue samples using a modified chelex protocol (Walsh, 1991). We incubated 10 µl of blood-ethanol mix at 60°C for 30 min. to allow all ethanol to evaporate. Three-hundred microlitres of extraction buffer (0.1 M Tris-HCl pH 8.0, 0.05 M EDTA, 0.2 M NaCl and 1% SDS) containing 5% w/v chelex, 500 µg of proteinase K and 250 µg of RNase was added and samples were placed on a rotating wheel and incubated overnight at 50°C. Extracted DNA was added to and

preserved in 300 µl of low TE (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA) 5% w/v chelex.

2.2.2 Microsatellite amplification

We used six microsatellite primer sets previously isolated from other woodpecker species (Table 2.1) to genotype 314 hairy woodpeckers. All polymerase chain reactions (PCR) were conducted in 10 µl reactions using 5x Goflexi Clear Buffer (Promega), 0.2 mM dNTP, 2.5 mM MgCl₂, 1 µM of each forward, reverse and M13 fluorescently labeled primers and 0.5 U of Crimson Taq (New England BioLabs). The only exception was Ptri 3 where the concentration of MgCl₂ was 2.0 mM. The following PCR conditions were used to amplify DNA: one cycle for one minute at 94°C, 45 s at T₁, one minute at 72°C; seven cycles of one minute at 94°C, 30 s at T₁ and 45 s at 72°C; 31 cycles for 30 s at 94°C, 30 s at T₂ and 45 s at 72°C; and ended with 1 cycle of 72°C for five minutes and 20 s at 4°C. For PCR amplification, five of the six loci (Dlu1, Dlu5, DMC111, DMC115, DMC118) were optimized using T₁=50°C and T₂=52°C, while Ptri3 was optimized using T₁=45°C and T₂=48°C. PCR products were run on a 6% acrylamide gel on a Li-Cor 4300 DNA analyzer. To maintain consistent sizing and scoring of alleles, we ran controls with known size standards on every load. All gels were scored independently by two observers and a subset of samples from each gel was re-run to ensure consistent amplification and scoring. For all analyses we only used individuals that had genotypic information for at least four loci.

2.2.3 Statistical analyses

Deviations from Hardy-Weinberg and linkage equilibrium between loci were analyzed using the program GENEPOP version 4.0.10 (Raymond and Rousset, 1995) and genotyping errors, presence of null alleles and allelic dropout were checked using MICROCHECKER version 2.23 (Van Oosterhout *et al.*, 2004). Observed heterozygosity (H_o), expected heterozygosity (H_e), total number of alleles (N_a) and number of private alleles (PA) were calculated using GENALEX version 6.41 (Goudet, 1995; Peakall and Smouse, 2006), while allelic richness was calculated using the program FSTAT version 2.9.2.3 (Goudet, 1995). Population bottlenecks result in a loss of alleles and temporarily high levels of observed heterozygosity, and can influence contemporary genetic patterns. To determine if any of the populations have recently undergone a bottleneck, a Wilcoxon sign-rank test was performed with BOTTLENECK version 1.2.02. We used the Wilcoxon-signed rank test as it is the most powerful test to use when analyzing populations with fewer than 20 individuals (Piry *et al.*, 1999).

2.2.4 Genetic variation and population differentiation

We used pairwise F_{ST} values and two Bayesian clustering models (STRUCTURE version 2.3.3; Pritchard *et al.*, (2000) and BAPS version 5; Corander *et al.*, (2003)) to look at population structure and genetic variation. Pairwise F_{ST} values are capable of detecting differences when sample sizes are low ($n=10$; Table 2.3; Harding, 1996) and genetic differences are small. Bayesian clustering models perform best when F_{ST} values are ≥ 0.05 and can overlook weak population structure (F_{ST} values are ≤ 0.03 ;

Latch *et al.*, 2006). One disadvantage of F_{ST} is it uses pre-defined groupings, in this case sampling sites. Since pairwise F_{ST} values are more robust and less conservative than Bayesian clustering models, they may show population structure that Bayesian clustering models are not able to detect (Weir and Cockerham, 1984). Pairwise F_{ST} values were calculated using the program GENETIX version 4.0.10 (Belkhir, 1999) and significance determined using 10 000 permutations. All tests of population differentiation using pairwise F_{ST} were corrected for using the false discovery rate correction method (Benjamini and Hochberg, 1995). Recent studies have suggested that Bonferroni corrections are too severe and often overestimate the number of populations that are not significantly different (Garamszegi, 2006). Corrections for multiple comparisons (i.e. the false discovery rate) lower the critical value of significance to prevent type 1 errors (Benjamini and Hochberg, 1995). We used only populations with ten or more individuals ($n=14$; Harding, 1996), with the exception of the Queen Charlotte Islands ($n=9$) which we included as it contains an endemic subspecies (*Picoides villosus picoideus*). Because the theoretical maximum of 1 for F_{ST} is only valid when there are two alleles and microsatellites have higher variation, we calculated the global, theoretical maximum F_{ST} (Hedrick, 1999).

The program STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000) is a Bayesian clustering model that uses multi-locus genotypes to examine population structure and assign individuals to genetic clusters (K). Assignments are based on individual genotypes and not on population allele frequencies. For this reason we included samples from the less well sampled populations (BCR $n=2$; CeOR $n=2$; CoOR $n=4$; LAB $n=4$; SCCA $n=2$; SK $n=3$; SSK $n=2$; Figure 2.1) in the initial run, because

assignments are not biased by small sample size. Individual genotypes are sampled and assigned to a cluster one at a time and group membership to a K is determined by the proportion of an individuals' genome (Q) that originates or belongs to a given K (Pritchard *et al.*, 2000). The model is run starting with the assumption that there is a single genetic cluster ($K=1$) and subsequent runs are conducted at $K>1$ to determine the true K for the given dataset. For our study we used the admixture model, correlated alleles and sampling locations as a prior in the model as suggested by Pritchard *et al.* (2000). These priors were used based on the assumptions that: limited gene flow is present (admixture model) and allele distributions will be restricted within a geographic area (correlated alleles, sampling location) due to limited dispersal by individuals. For all STRUCTURE runs, we used a burn in length of 100 000 Monte Carlo Markov Chains (MCMC) followed by 500 000 MCMC runs for $K=1$ to 10 for the initial runs and $K=1$ to 7 for all additional runs. MCMC is an algorithm sampling method widely used in Bayesian statistics (as well as other disciplines) to sample probability distributions, with multiple and longer runs conducted to improve the quality of results (Diaconis, 2008). Each K was run for 10 iterations, as this is the minimum number of iterations suggested by Evanno *et al.* (2005). MCMC burn in and run lengths were based on multiple preliminary runs used to determine the length of burn in and runs required. To determine the number of genetic clusters (K) present we used two methods ΔK (Evanno *et al.*, 2005) and Bayes factor (Pritchard *et al.*, 2000). The ΔK statistic (determined by the change in log probability ($\ln(K|D)$) between runs) is plotted for each K . The ΔK values were calculated using the online program *Structure Harvester*

(http://taylor0.biology.ucla.edu/struct_harvest/). Bayes factor also uses posterior probability ($\ln(K|D)$) to estimate the true K (Pritchard *et al.*, 2000). Specifically it measures the probability for each value of K ($e^{-\ln(K|D)} / \sum e^{-\ln(K|D)}$) and the run with the highest Bayes factor (maximum of 1) is the true K . Following the initial STRUCTURE run, two main groups were detected. To test if additional clusters were present within the groups, a second set of runs with the same settings were done using individuals from each of the two main groups separately. Several populations with small sample sizes were equally assigned between the two main groups and we dropped these populations for subsequent runs to avoid assigning them to the wrong cluster as they had $\leq 60\%$ assignment to a single group.

A second Bayesian clustering program, BAPS version 5 (Corander *et al.*, 2003), was used. BAPS differs from STRUCTURE in that it originally assumes one panmictic population and assigns groups of individuals to genetic clusters using geographic information arriving at its conclusions based on joint posterior probabilities (Corander *et al.*, 2003). In addition BAPS implements a spatial model (Corander *et al.*, 2008) that utilizes coordinate points (i.e. longitude and latitude of the sample) and plots the spatial pattern of the genetic variation. BAPS has been shown to be conservative in the assignment of individuals (Latch *et al.*, 2006) and so we compared results with STRUCTURE and F_{ST} to determine the patterns of genetic variation across the entire range. We used the no admixture spatial cluster model to plot the spatial genetic patterns using the default settings (Corander *et al.*, 2008) (Figure 2.1). We used 23 populations (see Figure 2.1) for BAPS using only

populations with at least 5 individuals (based on preliminary runs, as the program was inconsistent with its assignment to populations with fewer than 5 individuals).

2.2.5 Influence of barriers on genetic structure

To study how barriers affect genetic structure, we used three approaches. First a Mantel test was conducted to test for isolation by distance (Wright, 1946) by plotting $F_{ST}/(1-F_{ST})$ against straight line geographic distances between each population using GenAlEx version 6.41 (Peakall and Smouse, 2006). Straight line geographic distances were calculated using the Geographic Distance Matrix Generator (http://biodiversityinformatics.amnh.org/open_source/gdmg/). Tests for isolation by distance included all populations used for F_{ST} analyses (Table 2.3) and tested for isolation by distance across the whole range and within geographic regions (east vs. west as determined by groupings based on F_{ST}). The possibility of barriers was tested using the program BARRIER 2.2 (Manni, 2004). BARRIER connects sampling locations based on latitude and longitude (centre of sampling site) using Delaunay triangulations (Appendix 2) and barriers are determined using Monmonier's distance algorithm (draws lines through population-pairs connected from Delaunay triangulations and Voronoi tessellations, starting with the population pair with largest genetic distance (i.e. F_{ST} values), Appendix 2). F_{ST} matrices were generated for each locus and for all six loci combined for 23 populations (Figure 2.1). The strength of each barrier was determined based on the number of loci supporting it. We kept all barriers up to the fifth order (barriers are ranked in order of strength, with one representing a strongly supported barrier and ten a barrier

with weak support) with at least five loci supporting it and ignored all barriers greater than the sixth order and those with fewer than five loci supporting them.

Finally to test whether physical barriers act as barriers to gene flow, we paired populations separated by contemporary physical barriers (e.g. sampling sites on either side of the Rocky Mountain Range: UT and CO) and compared F_{ST} values between populations (Table 2.3).

The program BAYESASS+ (Bayesian inference of recent migration using multilocus genotypes) version 1.3 uses MCMC to estimate recent migration rates (Wilson and Rannala, 2003). We used BAYESASS+ to determine recent rates of migration between populations ($n=23$; Figure 2.1) and to further determine how isolated populations are from each other using the default options as advised by Wilson and Rannala (2003).

2.3 Results

We genotyped 314 hairy woodpeckers from 30 sampling locations across North America using six microsatellite loci. Two loci (Dlu1 and Dlu5) showed a high probability for null alleles ($P=0.20$; $P=0.16$, respectively). These same two loci showed significant deviations from Hardy-Weinberg equilibrium for twelve sampling sites following corrections for multiple tests ($P<0.025$; Table 2.2). Removing Dlu5 from the analysis resulted in all sampling sites being in Hardy-Weinberg equilibrium, with the exception of CO, MT and NS ($P\leq 0.025$). Including Dlu1 and Dlu5 in all genetic analyses, gave the same results as when they were excluded. No tests for linkage disequilibrium were significant and no population

showed significant evidence for a recent bottleneck ($P \geq 0.05$ for excess heterozygosity).

Observed heterozygosity was less than the expected heterozygosity and ranged from 0.10 to 1.00 (Appendix 3) depending on the population and locus. We found 8 to 15 alleles/locus (Table 2.1), with an average of 6.4 alleles per locus/population and allelic richness (which accounts for unequal sampling size) ranged from 1.3 to 5.5 (Table 2.2). No population had more than one private allele and there were no obvious geographic patterns associated with private alleles (Table 2.2).

Pairwise F_{ST} values showed population structure (Table 2.3) with AK and QCI being significantly different from all other populations. CO is significantly different from all populations except UT, while NL was significantly different from most populations with the exception of several eastern populations. MI was significantly different from all populations with exception to three populations (NC, NS and ON) found in the east. The global F_{ST} is 0.04, which is high given a theoretical, global maximum F_{ST} of 0.05.

Plotting ΔK for STRUCTURE runs at each K using the full dataset produced three peaks at K=2, K=6 and K=8. K=2 produced the highest peak and is indicative of the overall hierarchical structure ($\text{Pr} \ln(K|D) = -7326.2$). At both K=6 and K=8 the posterior probability (range = -7265.8 to -7800.3, Bayes factor ≤ 0.01) became more variable and had a lower ΔK value than at K=2. At K=2 we found distinct east - west differentiation following the eastern extent of the Great Plains in the south and approximately the Rocky Mountains in the north (all populations west of this line

are hereby referred to as the west group and all populations east of the line are referred to as the east group; Figure 2.1). Several sampling sites (BCR, CBC, NEOR, SAB and SD, Figure 2.2) showed high admixture between the east and west groups (<60% assignment to either the east or west genetic cluster) suggesting a zone of contact between the two groups in British Columbia, the Pacific Northwest and further south on the Great Plains (Figure 2.2a). WI also showed high admixture, but this is due to all samples missing information at one locus (DMC 111). We included Wisconsin in the east as when we analyzed the populations without DMC 111 it was strongly assigned to the east group.

Within the east group we found distinct population substructure (Figure 2.2b). At $K=2$, Alaska was genetically distinct from all other populations in the boreal-eastern US group ($\text{Pr } \ln(K|D)=-3450.3$; Bayes factor=0.99). Additional runs excluding AK from the analyses were performed to see if further population substructure was present within the boreal-eastern US subgroup, but we found no further population substructure. We found more population substructure within the west group than in the east group. Posterior probability was highest at $K=4$ ($\text{Pr } \ln(K|D)=-3776.96$) and probabilities with Bayes factor (0.92) confirmed four as the true K . The four distinct west subgroups are: the Queen Charlotte Islands, a Cascade Mountain group along the west coast; and two interior groups, one in the interior west and a second in the southeast Rocky Mountains (Figure 2.2c). The majority of populations were assigned $\geq 60\%$ to specific genetic clusters, however individuals in SOR, UT and VI populations showed high admixture (<60% assignment to one

genetic cluster; Figure 2.2c) and were assigned equally to more than one genetic cluster.

Results from BAPS differed slightly from those obtained using STRUCTURE, recognizing only four distinct clusters (Figure 2.1). While BAPS results were concordant with the subgroups identified by STRUCTURE in the east group (Alaska and boreal-eastern US), it only found two subgroups in the west group (QCI and all other west populations). BAPS also assigned CBC to the west group, whereas this population was assigned to the east group in STRUCTURE. CBC showed high admixture between the east and west group in our STRUCTURE runs and so it is not surprising that BAPS assigned CBC to the west group.

Both Bayesian clustering analyses failed to separate NL from other boreal-east populations, despite F_{ST} values between NL and some boreal-east populations being significant. Bayesian models likely failed to distinguish Newfoundland from other populations, as they perform best when $F_{ST} \geq 0.05$ (Latch *et al.*, 2006) and F_{ST} values for Newfoundland were under the predicted threshold (0.011 to 0.035).

Testing for isolation by distance (IBD) across the whole North American range showed a weak but significant relationship ($r^2=0.11$, $P=0.02$; Figure 2.3a). As QCI and AK are both significantly different from all other populations (Table 2.3) and located at the edge of the range (i.e. large geographic distances to other sampling sites), we removed both populations from the analysis to see if these two populations were influencing IBD for the whole range. Patterns of IBD strengthened when we did this ($r^2=0.20$; $P<0.01$; Figure 2.3a), but the relationship between genetic and geographic distance was still weak. CBC was included in the west as it

was more similar to western populations than east populations based on F_{ST} (Table 2.3). Tests for isolation by distance were significant within the east ($r^2=0.92$, $P=0.02$; Figure 2.3b), but not within the west group ($r^2=0.11$, $P=0.11$; Figure 2.3c). We performed the test for the east removing AK to see if this population was influencing IBD patterns in the east. When Alaska, was removed we found a weak and non-significant relationship in the east group ($r^2=0.19$, $P=0.12$; Figure 2.3b). Results show that genetic differences, as measured by $F_{ST}/(1-F_{ST})$, do not increase with geographic distances within the east and west group (except when Alaska is included). Weak patterns of IBD suggest that hairy woodpeckers expanded from ice-free refugia following the phalanx model (Figure 1.1) as individuals were genetically similar over large distances suggesting slow gradual movement from a common refugial source.

The program BARRIER identified five main barriers to gene flow (Figure 2.1). When we compared pairwise F_{ST} values between populations separated by barriers, they were significantly different ($P=0.001$ to 0.010), with the exception of the barrier separating North Carolina from Missouri ($P=0.328$). Alaska is more genetically similar to birds in central British Columbia (CBC) and Alberta (as identified by both Bayesian clustering methods), but nearby sampling sites (CAB, CBC, NBC) are separated from Alaska by several mountain ranges, notably the Alaska and Yukon-Tanana Uplands Mountain ranges. We were unable to calculate the potential for barriers between Alaska and the nearest continental sampling sites (CAB, CBC, NBC) due to the large geographic distances between our Alaska and British Columbia and Alberta sampling sites and the nature of Delaunay

triangulation as employed in the program BARRIER. Without sampling sites between Alaska and these areas it is difficult to tell if Alaska is completely isolated due to physical barriers like the Alaskan Mountain Ranges.

Estimates for migration using BAYESASS+ showed little to no gene flow between either AK or QCI and any other sampling site supporting pairwise F_{ST} values which identified these populations as being significantly different from all other populations. Gene flow was detected between sites separated by mountains, with significant asymmetrical migration rates from MT to CBC (0.11 migration rate; 95 % confidence intervals 0.00, 0.28), NBC (0.08; 0.00, 0.22) and ID (0.07; 0.00, 0.20); WA to SOR (0.05; 0.00, 0.19) and UT (0.05; 0.00, 0.16); and from CO to AZ (0.08; 0.003, 0.23) and UT (0.11; 0.02, 0.24). In the boreal-eastern US group gene flow was present throughout with high asymmetrical gene flow from NS to NC (0.08; 0.00, 0.25) and ON (0.16; 0.05, 0.28).

2.4 Discussion

Microsatellite analyses revealed distinct population structure across the North American range of hairy woodpeckers. We found a main east - west division with additional population structure within both groups, contrary to other widely distributed bird species like the red-tailed hawk (Hull *et al.*, 2008), which only shows population substructure in the western part of its range. The east group can be divided into two sub-groups: Alaska and boreal-eastern US, while the west is divided into as many as four subgroups: QCI, interior west, Cascade and southeastern Rocky Mountain. Greater population structure in the west reflects

barriers like mountains and large expanses of open water being more prevalent than in the east, where there are relatively few barriers.

Currently two endemic subspecies of hairy woodpeckers are recognized: *P. v. picoideus* in the Queen Charlotte Islands and *P. v. terraenovae* in Newfoundland. Genetic differentiation of hairy woodpeckers from the Queen Charlotte Islands supports previous work documenting morphology, plumage and genetic differences (Miller *et al.*, 1999; Topp and Winker, 2008) and previous work showing distinct genetic differentiation for other bird species found in QCI (Burg *et al.*, 2005, 2006; Topp and Winker, 2008). To date, no extensive genetic studies have been conducted on the Newfoundland subspecies of hairy woodpeckers. Our results suggest that while these birds are genetically different from most populations (Table 2.3), genetic differentiation is unlikely to be the result of long-term isolation. Firstly individuals from Newfoundland are not significantly different from nearby continental populations and secondly they do not contain any private alleles (Table 2.2). Previous studies have demonstrated similar patterns for other plant and animal species, where Newfoundland populations are genetically similar to continental populations despite reduced gene flow (Paetkau and Strobeck, 1996; Rajora *et al.*, 1998).

2.4.1 Barriers

Several pairs of populations are separated by large bodies of water from other nearby populations. QCI is isolated from northern BC by a large expanse of open water, Hecate Strait (~80 km), while the Strait of Georgia separates Vancouver

Island (VI) from the mainland coast (<35 km) and the Strait of Belle Isle separates the island of NL from Labrador (<20 km). QCI individuals have decreased genetic diversity and are genetically distinct from individuals in nearby northern BC populations on the mainland. By comparison, individuals from VI (based on Bayesian clustering program results, where estimates are conservative) and NL (F_{ST} , allelic richness and Bayesian clustering program results) are not genetically distinct from their nearest continental populations (WA and NS respectively). Our results suggest large open expanses of water like that between QCI and the mainland act as barriers to gene flow, whereas smaller expanses of open water do not. Within the Strait of Georgia there is a series of small islands that reduces the distance between VI and the mainland and may aid in gene flow between VI and mainland populations.

The hairy woodpecker's range is delineated by a number of prominent mountain ranges, including the younger and taller Rocky and Cascade Mountain ranges in the west and the older and lower Appalachian Mountain range in the east. While our results do suggest that distributions of subgroups in the west coincide with or are specific to mountainous regions, there is no evidence to suggest that mountains themselves are acting as barriers to gene flow. Despite tall mountains separating Alberta and British Columbia, gene flow is occurring between these areas. If mountains were acting as barriers to gene flow, we would expect to see individuals west of the mountains being significantly different from individuals east of the mountains. Hairy woodpeckers are viewed as generalists, as they have the ability to live in a wide range of forested habitats and at a range of elevations. Their

ability to survive at high elevations (up to 1900 m in British Columbia and 3500 m in New Mexico) and disperse short distances (Jackson *et al.*, 2002) likely allows them to move through valleys and mountain passes within mountainous regions.

Breaks in genetic clusters within North America are concordant with discontinuities in contemporary forested habitat. The east - west division follows the Great Plains, where grassland species are dominant and contiguous forest is rare (Barker and Whitman, 1988). Furthermore populations with admixture between the east and west groups occur in the Pacific Northwest and central British Columbia where boreal forest tree species meet western montane forest tree species. Areas like the Columbia and Harney Basins in Washington and Oregon separate the Cascades from the central Rocky Mountains, and hairy woodpeckers, due to the absence of forested vegetation, inhabit neither of these basins. Similar patterns are present for mesic forest dwelling frogs where interior populations are genetically distinct from coastal populations due to discontinuities in forested habitat (Nielson *et al.*, 2001). Genetic differences between WA and MT are significant ($F_{ST}=0.026$; $P<0.004$) over a relatively short distance (~500 km) suggesting the Columbia Basin may be acting as a barrier to gene flow.

The Wyoming Basin, an area characterized by shrub steppe and grasslands (Driese *et al.*, 1997), separates the southeastern Rocky Mountains from Montana and the northern Rocky Mountains. The Wyoming Basin may act as a barrier to gene flow, as populations of hairy woodpeckers on either side of the Wyoming Basin (MT and CO) are genetically distinct from each other (Table 2.3). Breaks in forest cover likely restrict gene flow across the Wyoming Basin for hairy woodpeckers, as the

Wyoming Basin is well documented as an area where genetic breaks occur for a variety of habitat specialist species (Arbogast *et al.*, 2001; DeChaine and Martin, 2005; Demboski and Cook, 2001; Galbreath *et al.*, 2010).

2.4.2 Effect of paleogeographic barriers and patterns of recolonization

Our study suggests contemporary genetic patterns have been influenced by past glacial events and patterns of recolonization following the LGM. The presence of few private alleles found in the individuals sampled and private alleles not being common to specific populations or geographic areas (e.g., north vs. south or glaciated vs. unglaciated areas) and weak patterns of IBD suggest woodpeckers dispersed slowly over short distances as they recolonized previously glaciated areas following the phalanx model (Hewitt, 1996). The presence of multiple, genetically distinct groups suggests that hairy woodpeckers may have recolonized previously glaciated areas from multiple refugia, as has been reported for other boreal bird species (Weir and Schluter, 2004b). Vegetation coverage and distribution of tree species became fragmented and isolated during the LGM (Williams, 2003) and with hairy woodpeckers preference for mature forested areas (Ripper *et al.*, 2007), individuals would have been restricted to the available habitat. Models using palynological data predict that tree species were likely restricted to refugia along the Pacific Coast, in the southeastern Rockies and to a greater extent in the eastern US (Williams, 2003). Genetic discontinuities and patterns for hairy woodpeckers reflect similar patterns seen in other mesic forest dwelling vertebrate species (Burg *et al.*, 2006; Carstens *et al.*, 2004; Nielson *et al.*, 2001) and have likely been

influenced by post-glacial expansion of tree species from ice-free refugia (Mack *et al.*, 1978b; Whitlock and Bartlein, 1997).

Boreal forest tree species were isolated in at least one refugium in the eastern US (Soltis *et al.*, 2006) and a potential Beringia refugium during the LGM (Brubaker *et al.*, 2005; Jackson *et al.*, 2000), and began to expand north and east, respectively, during deglaciation. Colonization of previously glaciated areas by boreal tree species was relatively fast in comparison to post-glacial expansion by tree species isolated in western refugia (Williams, 2003). Distribution of and genetic structure within the east group suggests post-glacial colonization from a single refugium (in the eastern US) following expansion of the boreal forest. Under this scenario range expansion into Alaska by hairy woodpeckers following the LGM likely occurred more recently as westward expansion of *Picea mariana*, *P. glauca* and *Pinus banksiana* (three common boreal forest species) from eastern refugia was relatively slow at high latitudes and elevations (McLeod and MacDonald, 1997). Allelic patterns do not support hairy woodpeckers being isolated in a northern Beringia refugium and current genetic patterns may reflect a founder effect, where loss of genetic variation occurs as a result of a new population being established by a small number of individuals.

Individuals from Vancouver Island share similar genetic signatures with birds from QCI (based on STRUCTURE analyses) and, as large expanses of open water currently isolate these islands from each other, genetic similarities may represent shared histories during the last glacial maximum when sea levels were lower (~125 m) and QCI and VI were connected through the mainland (Pielou, 1991). However,

given the decrease in genetic diversity we cannot rule out QCI being colonized more recently and genetic differentiation resulting from a founder effect. While we are unable to determine whether a coastal refugium was located as far north as the Queen Charlotte Islands or Hecate Strait, an area known to have acted as refugium for many plant and vertebrate species; (Brunsfield *et al.*, 2001; Burg *et al.*, 2005; Byun *et al.*, 1997; Warner *et al.*, 1982a) or further south, our results indicate populations in the Queen Charlotte Islands have been isolated from other populations in the west (Figure 2.1).

Breaks between the Cascade Mountains subgroup and the interior west and southeast Rockies subgroups reflect similar patterns in sedentary birds that show distinct morphology and plumage differences (Barrowclough *et al.*, 2004) and may suggest the presence of multiple ice-free refugia during the LGM in the geographic west. However, biogeographic patterns in the Pacific Northwest are quite complex (Brunsfield *et al.*, 2001) and differences between Cascade Mountains, interior west and southeast Rocky Mountain subgroups are likely reflective of the fine scale resolution of microsatellites where recent isolation may produce genetic differentiation (Beadell *et al.*, 2010). Further analyses with non-recombinant markers such as mitochondrial DNA (mtDNA) are necessary to better understand patterns of colonization in the west following deglaciation.

2.5 Conclusions

Distinct population structure is present in hairy woodpeckers within their North American range. We found an east - west split with the observed genetic

patterns reflective of similar patterns seen in both plant (Godbout *et al.*, 2005) and other vertebrate species (Turmelle *et al.*, 2011). Greater population substructure was found in the west, where contemporary barriers and breaks in contiguous forested habitat are more prevalent and multiple ice-free refugia are hypothesized (Brunsfeld *et al.*, 2001). In the east we found considerably less genetic structure where there are relatively few barriers to dispersal and post-glacial expansion likely occurred from a single refugium (Soltis *et al.*, 2006). While microsatellites are useful for exploring recent patterns of genetic differentiation, they lack the ability to fully understand past processes and future work should include mtDNA to explore these questions. Future work should include individuals from Central America to explore population differentiation between North American and Central American populations, as it will help to better understand genetic patterns in a widely distributed species.

Table 2.1: Repeat motif, primer sequence and allele size ranges for microsatellite loci used to genotype hairy woodpecker samples. For all loci primer annealing was optimized at 50/52°C with 2.5 mM MgCl₂, except Ptri3 optimized at 45/48°C with 2.0 mM MgCl₂.

Locus	Repeat type	Sequence (5' to 3')	Size
DIU1F ¹	(TG) ₁₇	CAC ACT GAA CAT ACC ATG TG	158-174
DIU1R		TAA AGA CCC TAA ACT TGC ACA	
DIU5F ¹	(GT) ₁₀ *	CTG ACC AAA GTG GAA AAG TAA	167-182
DIU5R		TCC TAC TAC CAT TTC TAG AAC	
DMC111F ²	(CATC) ₁₀	CGT ATG GAC CAG AAC ATA ATG	208-268
DMC111R		TGG GCT TTT AAG TCT TGT TG	
DMC115F ³	(ATCC) ₁₀	TGT CAG AGA TGG TTC ATG GGT GCA CT	292-332
DMC115R		CCA CTG GTG GCT CAG TTG CAC A	
DMD118F ²	(TGGA) ₁₀ (TAGA) ₁₃	CCC ATA TCC AGA GTT AGT TCT G	180-236
DMD118R		TCC TAG AGT CTT CAA CCT GAT C	
Ptri3F ⁴	(AGAT) ₁₁	GCA AAA GCC AGT TCC TGT GCA TGG	292-348
Ptri3R		GTT TCT TCA CCA TCA TTT TCC AGA CAG AA	

¹(Ellegren *et al.*, 1999); ²(Vila *et al.*, 2008); ³modified from (Vila *et al.*, 2008); ⁴modified from (Välimäki *et al.*, 2008);

*Single base pair insertion/deletion in flanking region for hairy woodpecker

Table 2.2: Summary statistics for each locus-population pair and combined for each population and locus; Na=number of alleles; He=expected heterozygosity; AR= allelic richness; n=sample size; and PA=number of private alleles. Asterisk (*) denotes missing data. Refer to Figure 2.1 for location of sampling sites.

Pop	Dlu1			Dlu5			DMC111			DMC115			DMD118			Ptri3			Overall				
	Na	He	AR	Na	He	AR	Na	He	AR	Na	He	AR	Na	He	AR	Na	He	AR	n	Na	PA	He	AR
AK	2	0.10	1.33	7	0.79	4.23	4	0.57	2.69	5	0.63	2.86	4	0.51	2.47	7	0.79	4.01	10	29	0.0	0.56	2.89
QCI	2	0.41	1.93	8	0.85	4.67	3	0.41	2.25	4	0.63	2.94	3	0.49	2.41	5	0.76	3.71	9	25	0.0	0.59	3.03
VI	2	0.42	1.97	4	0.70	3.43	4	0.56	3.25	4	0.56	3.25	6	0.80	4.50	7	0.82	4.83	5	27	1.0	0.64	3.56
WA	7	0.66	3.14	8	0.76	3.73	8	0.81	3.95	6	0.81	3.92	12	0.90	4.88	10	0.86	4.45	21	51	0.0	0.80	4.04
SOR	7	0.74	3.68	7	0.79	3.90	7	0.69	3.31	7	0.81	3.99	10	0.86	4.52	11	0.83	4.11	21	49	1.0	0.79	3.93
NEOR	5	0.74	3.93	4	0.66	3.33	5	0.74	3.93	4	0.66	3.46	7	0.84	5.00	6	0.76	4.26	5	31	0.0	0.73	4.00
ID	3	0.62	2.77	6	0.69	3.67	5	0.78	3.90	4	0.72	3.49	7	0.84	4.63	7	0.79	4.23	7	32	0.0	0.74	3.79
MT	7	0.69	3.41	9	0.77	3.84	7	0.78	3.67	8	0.80	3.96	13	0.89	4.73	12	0.88	4.66	31	56	1.0	0.80	4.05
NBC	5	0.68	3.19	7	0.78	4.08	6	0.75	3.73	7	0.81	4.15	8	0.85	4.53	5	0.71	3.44	10	38	0.0	0.76	3.85
CBC	7	0.71	3.37	9	0.75	3.77	6	0.73	3.39	7	0.78	3.77	11	0.87	4.61	10	0.88	4.71	21	50	0.0	0.79	3.94
CAB	3	0.40	2.27	4	0.58	3.07	5	0.61	3.13	6	0.77	3.95	7	0.83	4.52	6	0.81	4.36	7	31	0.0	0.67	3.59
SAB	5	0.61	3.01	6	0.73	3.82	5	0.63	3.16	5	0.79	4.14	7	0.81	4.35	8	0.80	4.33	8	36	0.0	0.73	3.78
AZ	2	0.34	1.85	7	0.81	4.52	4	0.46	2.55	4	0.74	3.55	6	0.80	4.17	3	0.60	2.66	7	26	0.0	0.62	3.18
UT	3	0.52	2.21	9	0.82	4.24	6	0.76	3.73	7	0.80	3.96	9	0.82	4.22	7	0.79	3.91	15	41	0.0	0.75	3.76
CO	6	0.69	3.19	9	0.82	4.16	9	0.71	3.44	9	0.71	3.37	10	0.87	4.54	10	0.80	4.04	29	53	0.0	0.77	3.77
SD	3	0.53	2.71	4	0.70	3.43	2	0.22	1.75	3	0.58	2.60	8	0.86	5.33	6	0.76	4.26	5	26	0.0	0.61	3.43
WI	3	0.46	2.47	5	0.76	3.92	*	*	*	6	0.76	4.26	10	0.89	5.55	7	0.82	4.83	6	31	0.0	0.62	4.25
MI	5	0.76	3.61	6	0.77	3.80	6	0.73	3.59	4	0.60	2.70	11	0.86	4.65	11	0.84	4.43	15	43	0.0	0.76	3.79
MO	3	0.53	2.71	5	0.78	4.39	6	0.80	4.50	4	0.66	3.46	4	0.72	4.00	6	0.81	4.93	5	34	0.0	0.72	4.02
NC	6	0.76	3.59	8	0.76	3.77	7	0.67	3.10	5	0.74	3.81	10	0.86	4.62	5	0.76	3.92	15	41	1.0	0.76	3.77
ON	5	0.61	3.08	7	0.71	3.41	7	0.79	3.92	*	*	*	10	0.86	4.54	5	0.75	3.79	19	35	0.0	0.62	3.70
NS	7	0.77	3.69	11	0.78	3.99	10	0.77	3.68	8	0.77	3.64	11	0.89	4.79	11	0.87	4.54	30	58	1.0	0.81	4.05
NL	5	0.71	3.30	8	0.83	4.41	6	0.79	3.85	7	0.80	4.04	9	0.79	4.09	9	0.84	4.40	13	44	0.0	0.79	4.00
Total/Avg.	8	0.59	2.89	15	0.76	3.89	12	0.64	3.38	11	0.69	3.60	15	0.81	4.42	15	0.80	4.21	314	76	5.0	0.71	3.75

Table 2.3: Pairwise F_{ST} values for population differentiation (for populations with $N \geq 9$): F_{ST} values listed below diagonal and P values listed above diagonal. Bold indicates that populations are significant at $P < 0.025$ (critical P value for Benjamini-Hochberg correction).

	AK	QCI	WA	SOR	MT	NBC	CBC	UT	CO	MI	NC	ON	NS	NL
AK		<0.001	0.004	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.001	0.004	<0.001	<0.001	<0.001
QCI	0.243		<0.001	0.002	<0.001	0.004	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
WA	0.139	0.089		0.119	0.004	0.062	0.004	0.115	<0.001	<0.001	0.006	<0.001	<0.001	<0.001
SOR	0.172	0.066	0.017		0.006	0.475	0.004	0.056	0.004	<0.001	0.001	<0.001	<0.001	0.001
MT	0.115	0.104	0.026	0.022		0.166	0.069	0.051	0.010	<0.001	0.009	<0.001	<0.001	<0.001
NBC	0.124	0.067	-0.012	-0.001	0.010		0.367	0.173	0.005	0.008	0.008	0.065	0.053	0.004
CBC	0.081	0.115	0.028	0.027	0.010	0.002		0.006	<0.001	0.011	0.281	0.007	0.163	0.005
UT	0.129	0.097	0.035	0.021	0.014	0.014	0.029		0.184	<0.001	<0.001	<0.001	<0.001	<0.001
CO	0.172	0.121	0.059	0.027	0.017	0.040	0.042	0.008		<0.001	<0.001	<0.001	<0.001	<0.001
MI	0.126	0.166	0.088	0.057	0.052	0.044	0.025	0.086	0.094		0.149	0.097	0.164	0.008
NC	0.125	0.146	0.067	0.037	0.023	0.042	0.004	0.056	0.055	0.013		0.260	0.178	0.145
ON	0.116	0.201	0.111	0.066	0.045	0.031	0.030	0.097	0.110	0.021	0.013		0.152	0.033
NS	0.108	0.122	0.043	0.034	0.035	0.019	0.006	0.053	0.068	0.008	0.007	0.008		0.095
NL	0.124	0.156	0.081	0.054	0.041	0.055	0.032	0.064	0.083	0.035	0.012	0.033	0.011	

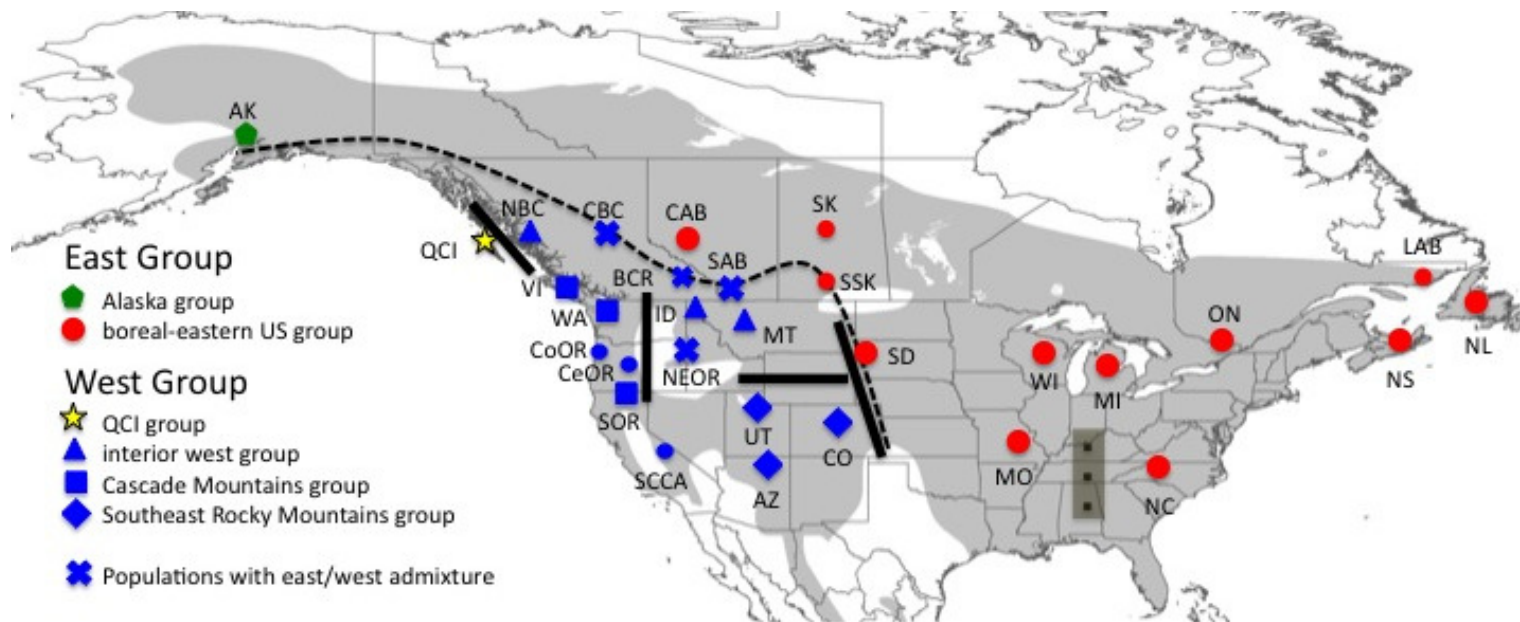
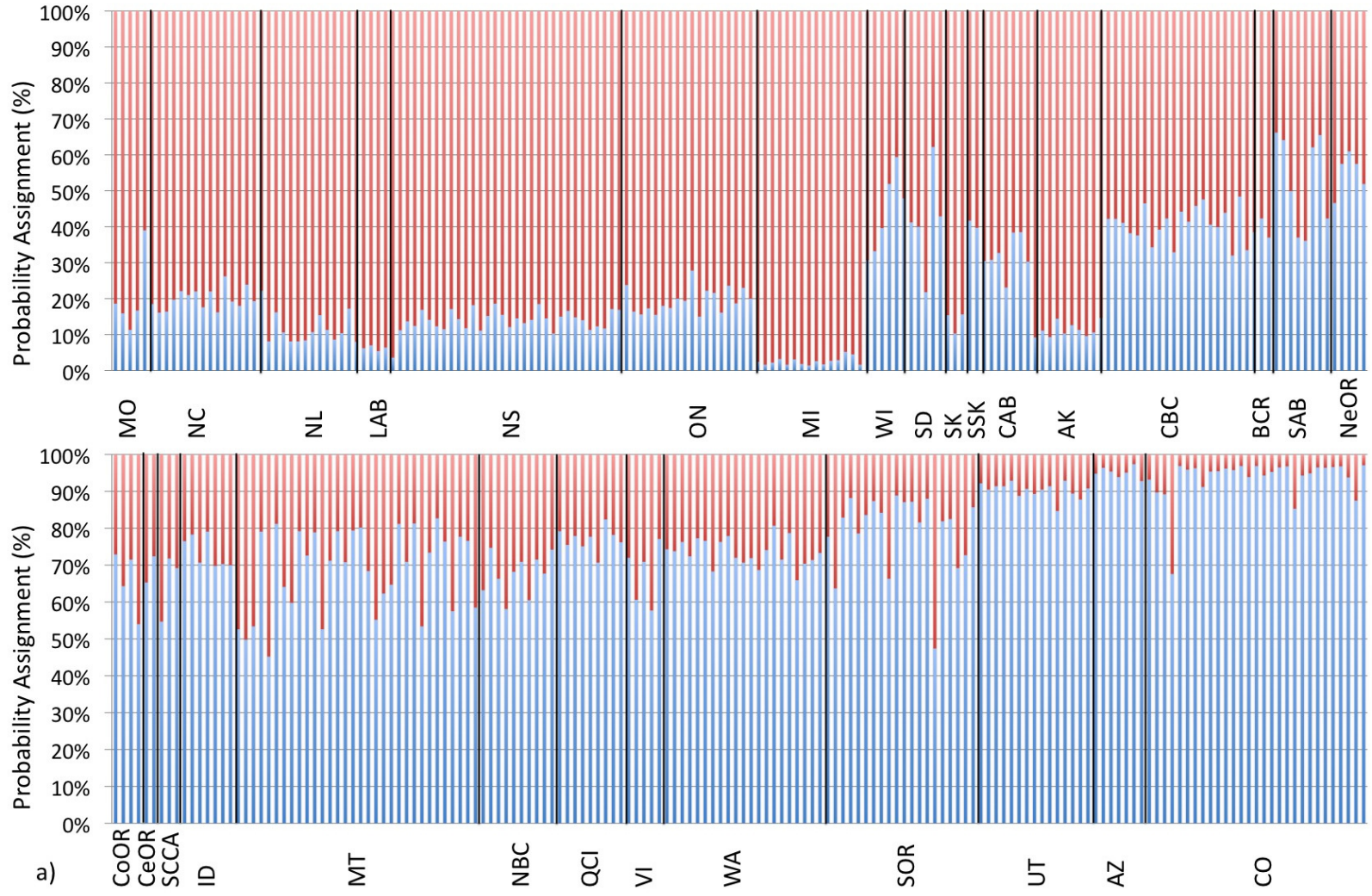
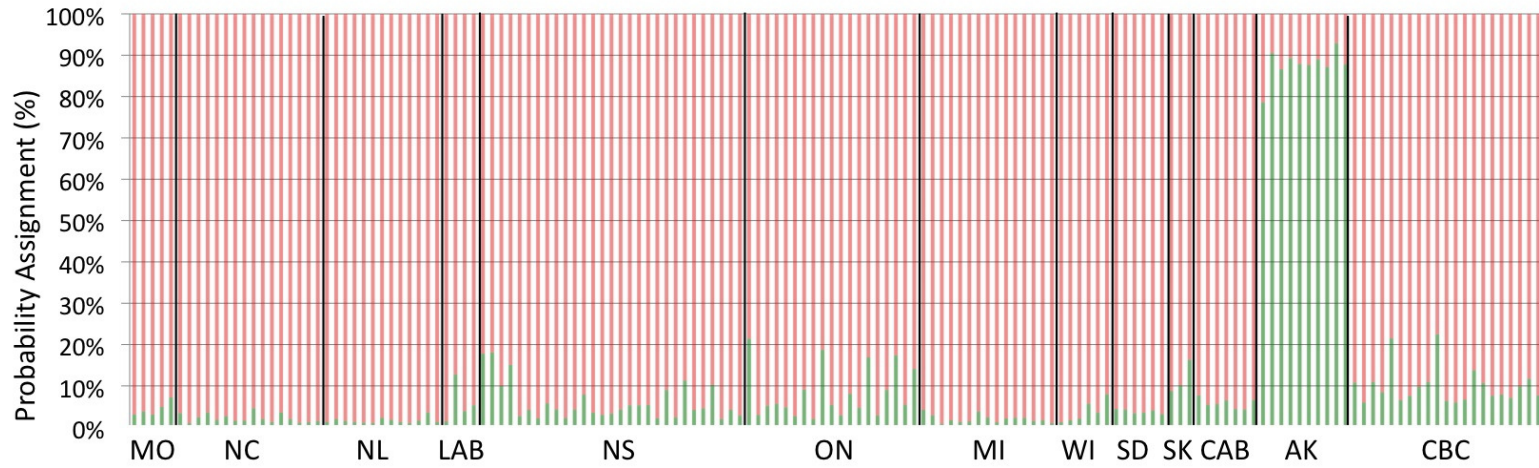
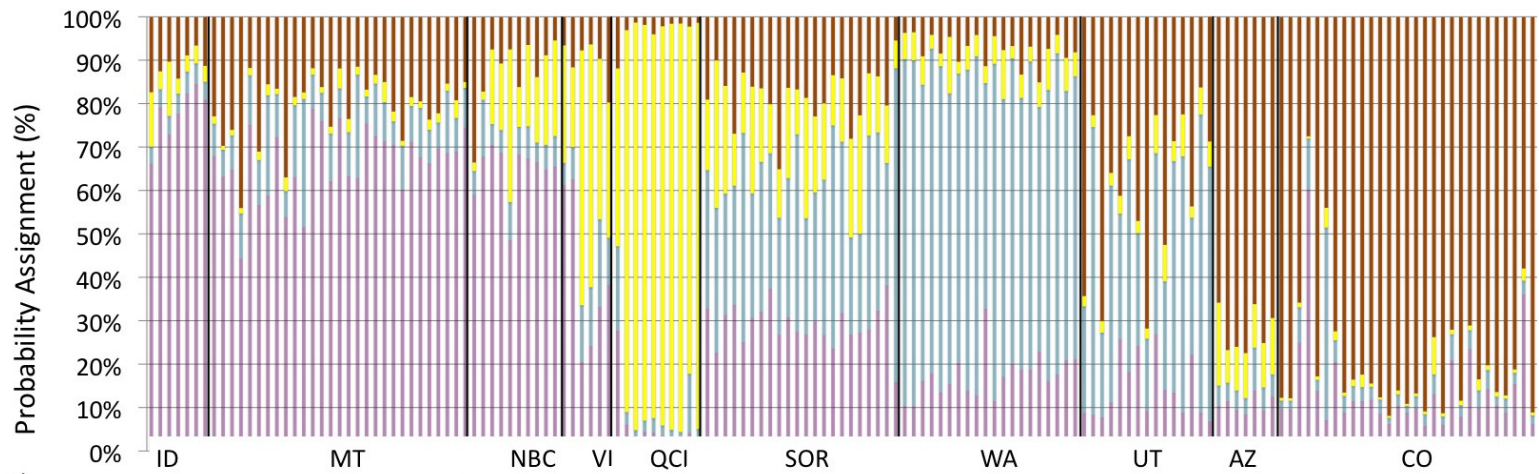


Figure 2.1: Geographic distribution and population structure of hairy woodpeckers in North America based on microsatellite data. Sample sizes are provided in Table 2.2. Shaded grey area represents the current range distribution of hairy woodpeckers (Ridgely *et al.*, 2007). Colours show assignment of populations to groups based on BAPS analyses of microsatellite data. Green=Alaska subgroup; red=boreal-eastern US subgroup; yellow=Queen Charlotte Island subgroup; dark blue=west subgroup. Black dashed line represents the east-west split based on STRUCTURE analyses (K=2). Key on left denotes groups within the east (K=2) and west (K=4) based on additional STRUCTURE analyses. Smaller circles represent populations ($n \leq 4$) that were only included for the initial STRUCTURE analysis (with colour denoting whether they belong to the east or west group) and excluded from all other analyses (SK was only included in BAPS and STRUCTURE analyses). Populations were considered to be east-west admixed if they were assigned at <60% to one group. Solid black lines represent barriers to gene flow (supported by F_{ST} pairwise values) as identified by BARRIER. Boxed dotted black line represents barriers to gene flow identified by BARRIER but not supported by pairwise F_{ST} values.



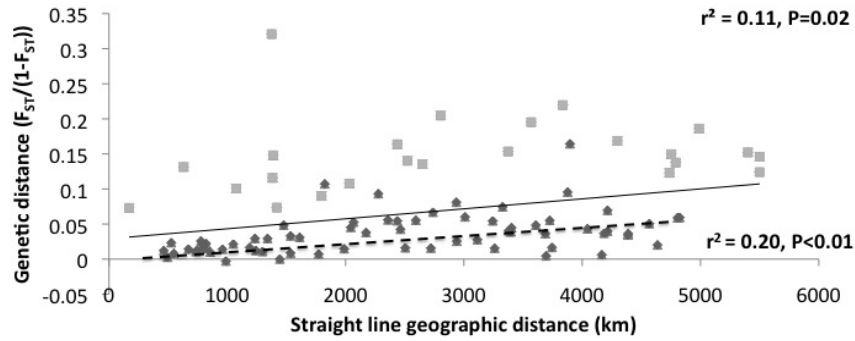


b)

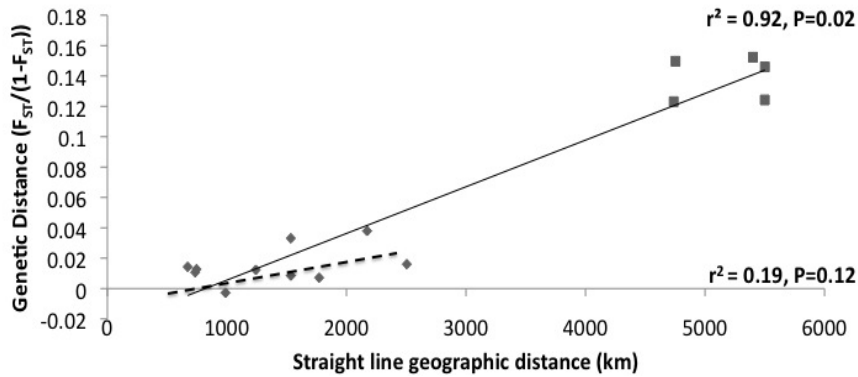


c)

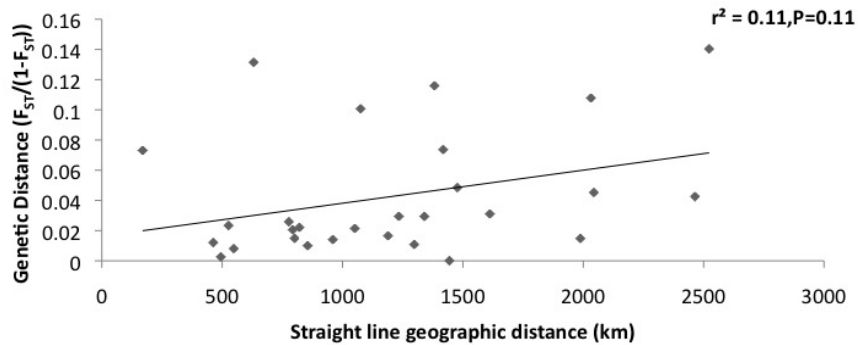
Figure 2.2: Histograms showing population clusters based on STRUCTURE analyses for a) K=2 across whole range (east in red, west in blue), b) K=2 for boreal-east group (red=boreal-eastern US subgroup; green=Alaska subgroup). c) K=4 for west group (purple=interior west subgroup, yellow=QCI subgroup, blue=Cascade Mountains subgroup and grey=southeast Rocky Mountain subgroup). Names along x-axis of histograms represent population names (see Figure 2.1 for locations), while values along the y-axis shows the group assignment probability to each group/subgroup. Each bar represents a single individual.



a)



b)



c)

Figure 2.3: Graphical representation of tests of isolation by distance measuring geographic distance (km) versus genetic difference ($F_{ST}/(1-F_{ST})$) versus for **a)** whole range; squares denote Alaska and Queen Charlotte Islands points; diamonds are points for the other 12 populations; black line denotes slope for whole range ($r^2=0.11$; $P=0.02$), dashed line denotes slope for range minus Alaska and Queen Charlotte Islands points ($r^2=0.20$; $P<0.01$) **b)** within the east group (six populations); squares denote AK points, black line denotes slope through eastern range ($r^2=0.92$; $P=0.02$), dashed line denotes slope for east group minus Alaska ($r^2=0.19$; $P=0.12$) and **c)** within the west group (eight populations); black line denotes slope through western group ($r^2=0.11$; $P=0.11$).

**Chapter 3: Post-glacial expansion of trees determines population structure for
a sedentary woodpecker species**

B.A. Graham^{1,2} and T.M. Burg¹

prepared as manuscript for submission

¹Department of Biological Sciences, University of Lethbridge, 4401 University Dr,
Lethbridge AB, T1K 3M4

²corresponding author: b.graham@uleth.ca

Abstract

The hairy woodpecker is a common year round resident found from the tree line in North America to the highlands of Central America. We sequenced an 825 bp fragment of the avian mtDNA control region for 322 individuals from 34 populations to examine how genetic patterns and population structure have been affected by North America's glacial history. We found three distinct groups: two in North America (boreal-east and west; the west is divided into two groups: Pacific Coast and southeast Rocky Mountains) and one in Central America. Using ecological niche modeling in combination with genetic analyses, we determined that hairy woodpeckers were isolated in multiple refugia in North America (corresponding to the three genetic clusters) during the Wisconsin glaciation. Population genetic patterns indicate that post-glacial expansion followed the phalanx model and closely corresponds to recolonization patterns found in tree species. Post-glacial expansion of hairy woodpeckers from western refugia was impeded by slower recolonization rates of mesic tree species along the Pacific Coast and historical barriers (glaciated mountain passes) preventing northern expansion from the southeast Rocky Mountains. In comparison, expansion of boreal tree species from the eastern refugium into previously glaciated areas was more rapid, allowing boreal-east haplotypes to become more widely distributed than the other two North American genetic groups.

Keywords: mtDNA, post-glacial expansion, hairy woodpecker, Pleistocene, mesic, boreal, phalanx model, paleogeography

3.1 Introduction

Large portions of North America were covered in ice during the last glacial maximum (LGM) ~21,000 years ago (Pielou, 1991). Ice sheets during the Pleistocene fragmented species' ranges, restricting surviving individuals and populations to ice-free areas known as refugia. Long-term isolation in refugia during glaciations has been shown to promote genetic diversification of plants, birds and mammals (Avice *et al.*, 1998a; Avice *et al.*, 1998b; Jarmillo-Correa *et al.*, 2009; Weir and Schluter, 2004b).

Known refugia during the LGM include Alaska and southern North America, while the Queen Charlotte Islands (QCI), also known as Haida Gwaii, and Newfoundland are contested to have been ice-free (Pielou, 1991). Three refugia south of the ice sheets are reported (Pacific Coast, Rockies and Taiga), from which plant and animal species expanded following deglaciation (Brunsfeld *et al.*, 2001; Weir and Schluter, 2004a; Williams, 2003). The genetic patterns seen in contemporary populations are strongly influenced by how populations expanded from refugia following deglaciation (Johansen and Latta, 2003). Two alternate patterns of recolonization have been proposed by Hewitt (1996). Following the pioneer model individuals disperse long distances rapidly, establishing pockets of isolated populations ahead of the leading edge, thereby preventing further colonization events (Figure 1.1). The second model is the phalanx model, where colonization from the refugium is gradual and continuous resulting in little genetic structure with more recently founded populations being similar genetically to source populations (Figure 1.1).

Post-glacial colonization by forest animal species began as vegetation returned to previously glaciated areas (Pielou, 1991). Following the LGM, post-glacial expansion of forest dwelling taxa into previously glaciated areas would have mirrored recolonization patterns of the tree species they depended on while surviving in glacial refugia. The result would be similar distributions of forest species and the trees they depend on with genetic breaks reflecting their natural history. Understanding historical conditions and habitats in refugia may help to better understand current genetic patterns, densities and distributions.

The hairy woodpecker (*Picoides villosus*) is a year-round resident with limited dispersal and displays plumage and morphological trait variation across its range (Jackson *et al.*, 2002). Currently 17 subspecies are recognized with subspecies classification based on behavioural and phenotypic differences (Jackson *et al.*, 2002). The hairy woodpecker's range extends from Alaska in the north to the highlands of Panama and is found in most forest and woodland habitats at all elevations. Physical barriers, including the Cascade and Rocky Mountains and large bodies of water separating mainland subspecies from those found on islands (i.e. QCI and Newfoundland) delineate subspecies' ranges. The presence of distinct plumage and morphological traits across the range suggests limited dispersal or rapid phenotypic adaptation. Due to its preference for mature forested areas (Ripper *et al.*, 2007), post-glacial range expansion of hairy woodpeckers would have likely followed post-glacial colonization patterns of trees as they recolonized previously glaciated areas. Sedentary species will retain genetic patterns longer than migratory species due to limited dispersal, potentially providing greater insight

into historical gene flow. Detailed molecular analyses of a widely distributed species will provide further insights into the patterns and processes promoting genetic diversification of North American taxa.

The purpose of this study is to examine the effects Pleistocene glaciation have had on the genetic patterns and population structure of hairy woodpeckers. As hairy woodpeckers' contemporary distribution includes both known (Alaska and southern United States) and putative refugia (Queen Charlotte Islands and Newfoundland), populations have potentially been isolated in multiple refugia. We used an integrative approach, combining molecular techniques and ecological niche modeling to determine the location of ice-free refugia used by hairy woodpeckers during the LGM. Secondly we explored how recolonization patterns affected contemporary genetic patterns. Rapid expansion following the phalanx model has been shown to produce little genetic structure (Ball *et al.*, 1988), whereas rapid long-distance dispersal following the pioneer model (Hewitt, 1996) has been shown to produce genetic structure (Burg *et al.*, 2006). Due to the presence of distinct morphological and plumage variation, we predict that hairy woodpeckers were isolated in multiple refugia. If patterns are the result of historical separation and not recent adaptation, then divergence of morphological and plumage characteristics occurred during prolonged isolation in separate refugia. If the model of expansion is tied to dispersal patterns then hairy woodpeckers likely expanded from refugia following the phalanx model, since they are not normally long-distance dispersers.

3.2 Methods

3.2.1 Sampling and DNA extraction

Blood and tissue samples were obtained from 34 sample sites (from here on referred to as populations) covering the range of the hairy woodpecker (Figure 3.1, Appendix 1). We used mist nets to capture birds from 2007 to 2010 and collected blood samples (~100 µl of blood) from the brachial vein. All birds were banded with unique USGS bands to avoid resampling the same bird and samples were preserved in 95% ethanol. Tissue samples were also obtained from museums (American Museum of Natural History, Burke, Field Museum of Natural History, Louisiana Museum of Zoology, New Brunswick Museum, Royal British Columbia Museum, Royal Alberta Museum, Royal Ontario Museum, Royal Saskatchewan Museum and Smithsonian) to supplement our field sampling. For field samples we attempted to obtain samples within a 30 km radius for each sampling site. All museum samples were collected during the last 20 years to ensure that we analyzed contemporary genetic patterns. Museum samples were obtained from the same geographic points as field samples where available. When they were not available we obtained samples from adjacent counties or areas, and grouped these samples with field samples, attempting to group samples over as small an area as possible. All blood and tissue samples used came from samples collected during the breeding season. By limiting our samples to birds from the breeding season we tried to avoid sampling birds atypical of each population, since hairy woodpeckers have been reported to move short distances during the late fall and early winter (Jackson *et al.*, 2002).

DNA was extracted from blood and tissue samples using a modified chelex protocol (Walsh, 1991). We incubated 10 µl of blood ethanol mix at 60°C for 30 min. to allow all ethanol to evaporate. Three-hundred microlitres of extraction buffer (0.1 M Tris-HCl pH 8.0, 0.05 M EDTA, 0.2 M NaCl and 1% SDS) containing 5% w/v chelex, 500 µg of proteinase K and 250 µg of RNase was added, samples were placed on a rotating wheel and incubated overnight at 50°C. Extracted DNA was added to and preserved in 300 µl of low TE (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA) 5% w/v chelex.

3.2.2 MtDNA amplification

The mitochondrial control region (CR) was amplified using LThr (5' CAT TGG TCT TGT AAR CCA AAG 3') and H16824 (5' TGA TGG GAT TTT AGA GGA TGT G 3') primers. Nested primers were designed to amplify DNA from degraded tissue samples and full sequences were obtained using two primer sets: LThr and H16454 (5' GAC CAG TAA TGG CCC TGA GA 3') and L16388 (5' GCT TCA GGC CCA TAC TTT CC 3') with H16824. All polymerase chain reactions (PCR) were conducted in 25 µl reactions using 5x Crimson Taq (Mg-free) reaction buffer (New England Biolabs), 0.2 mM dNTP, 2.5 mM MgCl₂, 1 µM of each an L strand and H strand primer and 1.0 U of Crimson Taq (New England BioLabs). DNA was amplified using the following conditions: one cycle for two minutes at 94°C, 45 s at 54°C, one minute at 72°C; 37 cycles of 30 s at 94°C, 45 s at 54°C and one minute at 72°C; ending with 1 cycle of 72°C for five minutes and 20 s at 4°C. PCR products were run on a 3% agarose gel to confirm DNA amplification.

For sequencing, five microlitres of PCR product were incubated with 0.1 units of SAP and 0.1 units of exonuclease at 37°C for 15 minutes and enzymes deactivated by heating to 80°C for 15 minutes. One microlitre of purified PCR product was used for sequencing in a 10 µl reaction using 2.5x McLab sequencing buffer sequencing buffer, 1.5 units of BigDye Terminator Mix (version 3.1) and 0.5 µM of primer (LThr or L16388). Cycle sequencing was done using the following conditions: one cycle at 96°C for 45 s; 25 cycles at 96°C for 30 s; 50°C for 15 s; four minutes at 60°C and 22°C for 20 s. Sequencing reactions were purified using a standard sodium acetate ethanol precipitation protocol (Sambrook *et al.*, 1989) and run on an Applied Biosystems 3130 DNA Analyzer (University of Lethbridge) or 3730xl DNA Analyzer (Génome Québec Innovation Centre). Sequences were aligned using MEGA version 4.0 (Tamura *et al.*, 2007) and double-checked by visual inspection of the chromatograms.

3.2.3 Statistical analyses

We measured genetic variation within populations and haplogroups (Nei and Tajima, 1981) by calculating haplotype diversity (H_d) and nucleotide diversity (π) using DnaSP version 5.0 (Librado and Rozas, 2009). Pairwise F_{ST} values were used to examine population structure and assess genetic differentiation between populations and haplogroups. Individuals were assigned to a haplogroup based on results from our maximum likelihood tree. F_{ST} values are directly related to variation in alleles, with values approaching zero inferring a close relationship between two populations, whereas values approaching one are indicative of large

differences between populations (Holsinger and Weir, 2009). For analysis of genetic differentiation between populations using F_{ST} values we only included populations with five or more samples. Pairwise nucleotide differences can be used to calculate divergence times between sequences and populations. Pairwise F_{ST} and nucleotide differences were calculated using Arlequin version 3.11 (Excoffier *et al.*, 2005).

To visualize population structure and evaluate relationships among haplotypes, we constructed a statistical parsimony network using the program TCS version 1.21 (Clement *et al.*, 2000). We created an unrooted maximum likelihood tree to further assess the relationship among individual haplotypes. Before creating the maximum likelihood tree we ran the program JMODELTEST version 0.1.1 (Guindon and Gascuel, 2003; Posada, 2008) to determine the nucleotide evolution model that best fit our data set. The program chose a generalized time-reversible model plus gamma plus invariants (GTR+G+I) as the best fit and we confirmed this model over other suggested model using Bayes factor. Our tree was created in MEGA 5.0 and FIGTREE version 1.3.1 (Rambaut and Drummond, 2006) using 500 bootstrap replicates (Tamura *et al.*, 2011).

We used three demographic tests, Tajima's D (Tajima, 1989), Fu's F_S (Fu, 1997) and R_2 (Ramos-Osnins and Rozas, 2002), to test whether populations of hairy woodpeckers have been stable or undergone recent demographic expansion. While the three statistics measure similar indices, they differ in sensitivity (Ramos-Osnins and Rozas, 2002). Tajima's D, is relatively conservative but performs well with large sample sizes. Fu's F_S and R_2 are less conservative and have higher statistical power than D. F_S performs best when sample sizes are high ($n=50$) and R_2 is best when

sample sizes are limited ($n=10$). Demographic tests suggest recent population expansion when values are significant ($P \leq 0.02$; Ramos-Osnins and Rozas, 2002) and for D , and F_s if values are negative. All tests of demographic expansion were performed in DNASP using 10 000 coalescent simulations. To characterize and visualize population demography (in addition to the demographic tests D , F_s and R_2), we plotted mismatch distributions. Mismatch distribution differs from other neutrality tests as it explores the distribution frequency of the number of nucleotide sites that differ between sequences. Mismatch distributions can be used to infer population history, with populations showing a Poisson distribution indicating a recent expansion or bottleneck, while a bimodal curve indicates the presence of more than one genetic lineage (Slatkin and Hudson, 1991). Raggedness index (r) measures whether mismatch distributions fit a model of population expansion (Harpending, 1994), with significant values indicating populations have been stable over time and have not undergone recent demographic expansion. All mismatch distributions and r values were calculated using Arlequin 3.11 (Excoffier *et al.*, 2005).

We used Bayesian skyline plots (Drummond *et al.*, 2005) to explore demographic history. Bayesian skyline plots are very powerful and use standard Monte-Carlo Markov chain sampling methods to estimate changes in effective population size over time. Skyline plots differ from other models as estimates are based on the sampled gene sequences rather than estimated genealogies (Drummond *et al.*, 2003). Bayesian skyline plots were executed in the BEAST (Bayesian evolutionary analysis by sampling trees) software package that includes

BEAST version 1.6.1, BEAUTi version 1.6.1 and TRACER version 1.5 (Drummond and Rambaut, 2007), and were run using the GTR+G+I, nucleotide substitution model and a strict molecular clock with a substitution rate of 5.57×10^{-8} /My/lineage (My=million years) for 10 million generations, and sampled every 1000 iterations with a burn in of 10%. To calculate nucleotide substitution rates we used a mutation rate of 11.1%/My and a generation time of two years. Mutation rates were estimated from avian mitochondrial control region values calculated for other avian species (domain I = 20%/My; domain II=5%/My; domain III=23%/My; Baker and Marshall, 1997). The mutation rate (11.1%/My) was the average based on the proportion of each domain sequenced (domain I=288 bp; domain II=496 bp; domain III=41 bp).

To determine when populations began expanding we calculated tau (τ). τ is equal to $2ut$, where u equals $2\mu k$, with μ equal to the mutation rate (11.1%/My) and k equal to the sequence length and t equaling the time of expansion (Rogers and Harpending, 1992). Estimates for divergence times were calculated using a standard population genetic technique (Morris-Pocock *et al.*, 2010) where $T=\delta/r$, with δ equal to the net number of nucleotide substitutions per site between populations and r equal to the mutation rate (11.1%/My) (Rogers and Harpending, 1992). All τ and net nucleotide differences were calculated using DnaSP version 5.0 (Librado and Rozas, 2009).

3.2.4 Ecological niche modeling

We reconstructed past distributions of hairy woodpeckers using MAXENT version 3.3.3e (Phillips *et al.*, 2006) in order to predict potential refugia occupied by hairy woodpeckers in North America during the LGM. Ecological niche modeling predictions have been shown to correspond with phylogeographic patterns, suggesting that both methods are complimentary to each other (Waltari *et al.*, 2007). Model predictions are based on 19 climate variables (Appendix 6) downloaded from the WorldClim dataset (Hijmans *et al.*, 2005), and all analyses used the default convergence threshold and maximum number of iterations (500), with 25% of all localities used for model training. Distribution modeling for the LGM used paleoclimate data drawn from the Community Climate System model (Otto-Bliesner *et al.*, 2006). Model suitability is determined using the AUC (area under the curve) where values range between 0 and 1. Values closer to 1 indicate a good model fit. Current and historical distributions of hairy woodpeckers were created using 7295 locality points, combining samples used for genetic analyses with records downloaded from the Global Biodiversity Information Facility (GBIF) data portal (<http://data.gbif.org>). MAXENT combines current distribution records with climate conditions and then infers past distributions by identifying areas in the past where these climate conditions were found. All duplicate points were omitted by the program for distribution predictions to prevent any sampling bias due to specific locations in the data set. To determine if historical distribution models were suitable, we compared AUC values between the training and testing models and looked at the omission plot. Model predictions are considered strong, if differences

between testing and training model values are above 0.8 (Phillips *et al.*, 2006). The model also produces a predicted omission plot, with large deviations from this line indicating statistically weak models. Paleodata are unavailable for Central America, so we used ecological niche modeling to make predictions about past distributions for North America only.

3.3 Results

We sequenced an 825 bp sequence of the avian mitochondrial control region for 322 hairy woodpeckers from 34 sampling localities (Figure 3.1). Within the 825 bp sequence we found 50 variable sites, 32 of which were parsimony informative and identified 120 different haplotypes (Figure 3.2, Appendices 4 and 5). Haplogroups were also differentiated based on our maximum likelihood tree (Figure 3.1 inset), which recognized three distinct lineages corresponding with geographic areas: Central America, boreal-east and western North America. Western North America can be further differentiated into two subgroups: Pacific Coast and southeast Rockies. Overall haplotype diversity was high ($H_d=0.97$; Table 3.1) with moderate levels of nucleotide diversity ($\pi=0.009$). Haplotype diversity was high ($H_d=0.92$ to 0.94) with moderate nucleotide diversity ($\pi=0.003$) for each of the three North American haplogroups, but was low for Central America ($H_d=0.17$, $\pi=0.0002$), where 11 of 12 individuals shared the same haplotype. Within individual populations we saw similar patterns for both haplotype and nucleotide diversity, with the exception of Alaska where haplotype diversity was low ($H_d=0.36$).

Pairwise F_{ST} values revealed high levels of genetic structure across the range (Table 3.2). Of the 253 pairwise F_{ST} values, 212 were significant. Central America is significantly different from all North American groups and populations and pairwise F_{ST} values were high ranging from 0.57 to 0.93 (Tables 3.2 and 3.3). Within the boreal-east group, both Alaska and Newfoundland are significantly different from all other populations, while NBC is significantly different from most populations (Table 3.2). All four groups were significantly different from each other ($P < 0.01$; Table 3.3).

Demographic tests suggest the Central American group ($F_S = -0.48$, $P = 0.13$; $R_2 = 0.28$, $P = 0.70$) has been stable while North American groups ($F_S = -15.05$ to -96.16 , $P < 0.01$; Table 3.1) have undergone recent demographic expansion. Within the boreal-east haplogroup, both D (-1.68 , $P = 0.02$) and F_S (-96.16 ; $P < 0.01$) were significant (Table 3.1) and demographic tests were significant for multiple populations (MT ($F_S = -5.10$, $P < 0.01$; $R_2 = 0.09$, $P = 0.02$), ON ($F_S = -4.79$, $P < 0.01$; $R_2 = 0.08$, $P < 0.01$), CBC ($F_S = -5.20$, $P = 0.01$), NS ($F_S = -13.48$, $P < 0.01$), and MI ($F_S = -6.62$, $P < 0.01$)) suggesting recent demographic expansion. In the west, results suggest that populations in CO ($F_S = -5.18$; $P = 0.01$) and SOR ($F_S = -5.54$; $P = 0.02$) have both undergone recent expansion.

Mismatch distributions (not shown) support demographic tests and suggest recent population expansion. Despite finding a bimodal curve for all populations combined, southern Oregon and Washington mismatch distributions and raggedness index (r) were not significant ($P > 0.05$). Non-significant mismatch distributions combined with a non-significant raggedness index indicate we cannot reject the null hypothesis of recent demographic expansions. We found a unimodal

curve for all three North American groups and all other populations we tested, and as probabilities for mismatch distributions and r were also non-significant ($P > 0.05$) we cannot reject the null hypothesis of recent demographic expansion for these groups or populations either.

Bayesian skyline plots (Figure 3.3) suggest that overall, populations remained stable during the Wisconsin with population expansion occurring following the onset of deglaciation. Results support demographic tests (D , F_s and R_2 ; Table 3.1 and mismatch), which suggest recent demographic expansion.

Tests for expansion times indicate the southeastern Rocky Mountain group began expanding ~ 7 kya, with expansion from the Pacific Coast and boreal-east groups occurring ~ 6.6 and 8 kya respectively. As expansion times represent an average for each of the four haplogroups, we are likely underestimating expansion times, especially since expansion into the central Rocky Mountains (ID and MT; ~ 3.2 and 3.7 kya respectively), Washington (~ 4.7 kya), southern California (~ 2.6 kya), Utah (~ 2 kya) and Newfoundland (~ 4.5 kya) occurred more recently. While demographic tests indicate recent population expansion, divergence times indicate the three main genetic lineages separated from each other prior to or at the onset of Wisconsin glaciation. Central American populations diverged from North America ~ 99 to 104 kya, while the boreal-east group split from the western group ~ 96 (for the southeast Rocky Mountain group) to 113 kya (for the Pacific Coast group). Divisions within the western group have occurred more recently with the southeast Rockies and Pacific Coastal groups having split during the last glacial ~ 34 kya.

The current distribution predicted by MAXENT (Figure 3.4a) matched the known distribution of hairy woodpeckers (Jackson *et al.*, 2002). The current distribution model performed well as it had a high AUC value (0.902) and curves for both the training and test sample omission were close to the predicted omission curve. Predictions of distributions 21 kya showed large range retraction, with hairy woodpeckers in North America being restricted to three main areas south of the ice sheets (Figure 3.4b). Along the Pacific Coast potential habitat extended primarily from southern California to Washington with smaller amounts of habitat along the British Columbian Coast. In the southeast Rockies predicted habitat was centralized in New Mexico and Arizona and extended south into Mexico through the Sierra Madre Occidental Mountain Range. In the southeast United States potential habitat was present from eastern Texas in the southeast to Virginia in the northeast (Figure 3.4b). The model also predicted potential habitat in Alaska, but predictions were lower (<0.50 probability) when compared to predictions for areas south of the ice sheets (0.53 to 0.85). The model predicted potential habitat in the Hecate Strait (QCI putative refugium), but did not predict potential habitat in or near another putative refugium in Newfoundland.

3.4 Discussion

During the LGM hairy woodpecker populations were relatively stable (Figure 3.3), with recent population expansion in North America (Table 3.1) occurring following deglaciation (~12 kya). Differentiation of the three genetic lineages (boreal-east, western and Central America) occurred relatively late in the

Pleistocene (~96 to 110 kya), with divergence in the western lineage (between Pacific Coast and southeast Rockies) occurring more recently (~34 kya) during the last glacial. Most studies have suggested species divergence occurring relatively early during the Pleistocene (Awise *et al.*, 1998a), although mid and late Pleistocene divergence has been shown for North American wolves (~150 kya, Wilson *et al.*, 2000) and sparrow (*Spizella*) species (~30 kya, Klicka *et al.*, 1999). Our results support Pleistocene glacial cycles promoting genetic divergence for hairy woodpeckers with genetic patterns and population structure resembling that found for some other sedentary bird species like the brown creeper (Manthey *et al.*, 2011), but not others like the downy (Ball and Awise, 1992; Zink *et al.*, 2002b) and three toed woodpecker (Zink *et al.*, 2002b).

3.4.1 Patterns of recolonization

Patterns of genetic differentiation suggest recent expansion from ice-free refugia followed the phalanx model (Hewitt, 1996). While genetic differences between populations of hairy woodpeckers are significant we did not find pockets of isolated populations, which is what we would have expected if hairy woodpeckers had expanded under the pioneer model (Figure 1.1). There does seem to be a cline in the Cascade Mountains where admixture of boreal-east and Pacific Coast haplotypes occurs. Despite North America being colonized following the phalanx model, we found distinct genetic structure across the hairy woodpeckers' North American range. The presence of distinct population structure across the range

suggests that hairy woodpeckers expanded from multiple refugia as has been shown for other high latitude bird species (Weir and Schluter, 2004b).

Post-glacial expansion in North America was exclusively from North American refugia, as we found no evidence for post-glacial expansion into North America from Central America, as has been shown for the widely distributed dark-eyed junco (*Junco hyemalis*, Milá *et al.*, 2007a).

3.4.2 Location of refugia

Ecological niche modeling predicted considerable range retraction for hairy woodpeckers during the last glacial (Figure 3.4) Results suggest that individuals were potentially isolated in three main areas south of the ice sheet, along the Pacific Coast (Hecate Strait), and north of the ice sheet in Beringia, but did not predict habitat in Newfoundland (an area contested to have been ice-free). To determine if individuals were potentially isolated in all of these areas, we can combine results with genetic patterns to suggest which areas likely acted as a refugium for hairy woodpeckers.

Within North America we found two lineages (east and west) and molecular data coupled with ecological niche modeling results suggest hairy woodpeckers were isolated south of the ice sheets during the last glacial in both the eastern and western United States (Boulet and Gibbs, 2006). Within the east genetic patterns suggest hairy woodpeckers were likely isolated in a single refugium, despite multiple refugia being suggested for eastern tree species (de Lafontaine *et al.*, 2010; Godbout *et al.*, 2010; Godbout *et al.*, 2005). Multiple refugia are hypothesized for

animal and plant species (Brunsfield *et al.*, 2001; Gugger *et al.*, 2010) in the western United States and genetic population structure in the west (Table 3.3; Figure 3.2) supports hairy woodpeckers being isolated in two separate western refugia. Ecological niche modeling results predict western refugia were located along the Pacific Coast and in the southeast Rockies, supporting past studies suggesting refugia for both plant and animal species were located in these regions (Arbogast *et al.*, 2001; Byun *et al.*, 1997; Gugger *et al.*, 2010).

Potential habitat for hairy woodpeckers during the last glacial may have extended as far north as the Queen Charlotte Islands based on our ecological niche modeling predictions. Forested habitat has been suggested to have been present on the Queen Charlotte Islands during the LGM (Warner *et al.*, 1982b) and previous work suggesting a QCI refugium for forest dwelling mammal (Byun *et al.*, 1997) and bird species (Burg *et al.*, 2005, 2006) support this. However, our results do not suggest that QCI acted as a refugium for hairy woodpeckers during the last glacial. Had QCI acted as a refugium, we would have expected QCI to have unique haplotypes (which it did not; Figure 3.2) and to have been significantly different from other Pacific Coast populations (based on pairwise F_{ST} values; Table 3.2). As QCI individuals share haplotypes with other birds along the Pacific Coast (Figure 3.2), it does suggest QCI was either part of the Pacific Coast refugium or that QCI individuals shared a common refugium with other Pacific Coast individuals.

Ecological niche modeling predicted potential habitat in Alaska (Brubaker *et al.*, 2005), but not in Newfoundland during the LGM. However genetic patterns do not support either area acting as a refugium. While Alaska and Newfoundland

populations have distinct haplotypes, haplotypes show few differences from other boreal-east haplotypes. If either area had acted as a refugium we would predict: 1) divergent haplotypes restricted to Alaska or Newfoundland (refugia with no subsequent dispersal) or 2) higher haplotype diversity in AK or NL with subsets of AK or NL haplotypes found in previously glaciated areas (refugia with subsequent gene flow). Alaska shows reduced genetic variation ($H_d=0.36$), suggesting either a recent bottleneck or founder effect. Reduced variation could also be characteristic of isolation in a refugium; however, the absence of divergent haplotypes suggests populations in Alaska have not been isolated for a long period of time. Pairwise F_{ST} values support reduced gene flow between Newfoundland and other populations (as has been shown for other plant and animal Newfoundland populations; Gill and Mostrom, 1993; Godbout *et al.*, 2010), but do not support a Newfoundland refugium for hairy woodpeckers despite the presence of a distinct subspecies.

3.4.3 Recolonization from multiple refugia

Boreal tree species emerged from eastern refugia (de Lafontaine *et al.*, 2010; Godbout *et al.*, 2005), and the wide distribution of boreal-east haplotypes across northern latitudes indicates that post-glacial expansion by hairy woodpeckers closely followed that of the boreal forest. Colonization of western Canada, the central Rockies (ID, MT, NEOR, SAB) and Washington was likely a recent, single event, as individuals in these areas share three main boreal-east haplotypes (D, G, and Y), with relatively few singleton haplotypes (Figure 3.2; Appendix 5). Expansion into the central Rocky Mountains likely occurred from northern Canadian

populations since the central Rockies share haplotypes with populations in British Columbia. If post-glacial expansion had occurred directly from the east, then South Dakota would share haplotypes with the central Rockies. However, South Dakota is included in the southeastern Rockies group suggesting this area was colonized as populations from the southeast Rockies moved north.

Contemporary and palaeo biogeographic barriers likely restricted expansion of hairy woodpeckers from the southeastern Rocky Mountain refugium into northern latitudes. Currently contiguous forest is fragmented by the Wyoming Basin (Driese *et al.*, 1997), while during the last glacial, forested habitat would have been fragmented by paleogeographic barriers (Hafner and Sullivan, 1995; Porter *et al.*, 1983). Patterns of genetic differentiation between the central Rocky Mountain and southeastern Rocky Mountain populations are observed for both forest (Aagaard *et al.*, 1995) and forest dwelling species (Arbogast *et al.*, 2001), as well as for other habitat specialist species (DeChaine and Martin, 2005; Demboski and Cook, 2001; Galbreath *et al.*, 2010). Isolation of forested habitats has been used to explain genetic patterns within this region (Arbogast, 1999; Hafner and Sullivan, 1995) and suggests that post-glacial expansion from the southeastern Rockies refugium into the central Rockies and Canada was slowed or prevented.

3.4.4 Barriers

Geographical barriers, such as large expanses of open water separating QCI from the mainland coast, and post-glacial expansion patterns of mesic forest species slowed eastward movement of Pacific Coast haplotypes. Mesic tree species like

Douglas fir (*Pseudotsuga menziesii*), red cedar (*Thuja plicata*) and western hemlock (*Tsuga heterophylla*) have only recently expanded into eastern Washington and Idaho (~3 to 1.5 kya) (Mack *et al.*, 1978a; Mack *et al.*, 1978b; O'Connell *et al.*, 2008; Whitlock and Bartlein, 1997). By comparison boreal tree species advancing from eastern refugia likely became established in western North America around seven to ten thousand years ago. Genetic patterns for hairy woodpeckers in the Pacific Northwest are concordant with patterns for other mesic dwelling taxa and overall biogeographic patterns in this region (Brunsfeld *et al.*, 2001), where Pacific Coast and interior splits are quite common (Burg *et al.*, 2006; Carstens *et al.*, 2004; Nielson *et al.*, 2001; O'Connell *et al.*, 2008).

3.5 Conclusions

Analyses of mtDNA control region revealed four geographic groups of hairy woodpeckers across their range, three in North America and one in Central America, with phylogeographic structure resembling that of other widely distributed vertebrate species (Arbogast *et al.*, 2001; Turmelle *et al.*, 2011). Within the North American groups we found further population structure, with higher levels of population differentiation in terms of haplotype frequency differences (i.e. pairwise F_{ST}) being present. In North America hairy woodpecker populations were isolated in multiple refugia (eastern US, southeast Rockies, Pacific Coast) with population structure resulting from post-glacial expansion from multiple refugia following the phalanx model. Post-glacial expansion of hairy woodpeckers followed recolonization patterns of tree species and population structure is indicative of past

and present biogeographic patterns. Future research should further explore Central American populations to understand the processes contributing to divergence between North America and Central America. As we sampled relatively few individuals in Central America we can only make broad inferences about the population structure and history of individuals in that region. Additionally alternate molecular markers (i.e. protein coding mtDNA or z-linked nuclear genes) should be used to further explore the processes that contributed to phenotypic differences and patterns for hairy woodpeckers.

Table 3.1: Genetic diversity and demographic statistics for hairy woodpecker sampling sites and TCS groups; n=sample size; N_h=number of haplotypes; D=Tajima's D; S=Segregating site; F_S=Fu's F_S; R₂=Rasmus-Osnins-Rozas test; H_d=haplotype diversity; π=nucleotide diversity; and - indicates test not performed due to insufficient data for the given population. Individuals from southern and central SK were pooled (SK). Groups refer to geographic regions based on TCS network (Figure 3.2). Refer to Figure 3.1 for location.

Locality	n	N _h	D	P	S	F _S	P	R ₂	P	H _d	p
AK	10	2	0.02	0.58	4	3.03	0.92	0.18	0.38	0.36	0.0020
NL	13	8	-0.56	0.30	9	-2.85	0.04	0.12	0.11	0.86	0.0030
SD	5	5	0.00	0.61	5	-0.33	0.31	0.24	0.34	0.90	0.0030
CO	20	12	-0.49	0.35	12	-5.18	0.01	0.09	0.06	0.93	0.0030
AZ	7	6	0.20	0.59	9	-1.70	0.10	0.19	0.38	0.95	0.0050
UT	15	9	-0.86	0.21	19	-1.36	0.22	0.11	0.10	0.89	0.0060
SCCA	3	2	-	-	3	-	-	-0.47	1.00	0.67	0.0020
SCA	16	7	-0.76	0.24	9	-1.50	0.16	0.12	0.23	0.75	0.0030
QCI	9	5	-1.42	0.08	15	0.78	0.67	0.21	0.84	0.81	0.0050
CeOR	2	2	0.00	1.00	13	2.56	0.57	0.50	1.00	1.00	-
CoOR	4	4	-0.65	0.34	16	0.09	0.28	0.26	0.53	1.00	0.0100
SOR	22	17	1.10	0.91	22	-5.54	0.02	0.17	0.80	0.97	0.0100
VI	5	5	1.22	0.87	16	-0.57	0.19	0.25	0.66	1.00	0.0110
WA	21	10	1.03	0.90	20	0.48	0.61	0.17	0.89	0.78	0.0090
SAB	7	4	-1.65	0.01	14	1.25	0.72	0.29	0.96	0.71	0.0050
MT	20	9	-1.34	0.09	8	-5.10	<0.01	0.09	0.02	0.78	0.0020
ID	11	5	-0.49	0.33	5	-1.08	0.16	0.16	0.23	0.78	0.0020
NEOR	6	5	-0.14	0.45	5	-1.97	0.05	0.19	0.19	0.93	0.0030
BCR	2	1	-	-	-	-	-	-	-	-	-
CBC	18	11	-0.01	0.53	9	-5.20	0.01	0.14	0.52	0.86	0.0030
NBC	10	6	-1.28	0.11	16	-0.08	0.47	0.20	0.74	0.78	0.0050
CAB	7	4	0.80	0.81	4	-0.23	0.38	0.22	0.46	0.81	0.0020
SK	5	5	0.29	0.66	8	-1.72	0.06	0.18	0.16	1.00	0.0050
LAB	4	3	-0.07	0.58	4	0.25	0.45	0.34	0.67	0.83	0.0030
NS	19	16	-0.40	0.38	12	-13.48	<0.01	0.12	0.29	0.98	0.0040
ME	1	1	-	-	-	-	-	-	-	-	-
MO	3	3	0.00	0.77	6	0.13	0.26	0.14	0.09	1.00	0.0050
NC	7	7	-0.36	0.39	10	-3.71	0.01	0.14	0.07	1.00	0.0050
LA	3	3	0.00	0.81	5	-0.08	0.26	0.25	0.22	1.00	-
WI	6	6	0.52	0.68	6	-3.39	0.01	0.20	0.28	1.00	0.0030
MI	14	11	-0.69	0.29	11	-6.62	<0.01	0.10	0.03	0.96	0.0030
ON	15	10	-1.56	0.05	14	-4.79	<0.01	0.08	<0.01	0.86	0.0030
C. America	12	2	-1.14	0.17	1	-0.48	0.13	0.28	0.70	0.17	0.0002
boreal-east	208	70	-1.68	0.02	41	-96.16	<0.01	0.04	0.04	0.94	0.0030
Pac Coast	53	23	-1.09	0.15	27	-17.09	<0.01	0.07	0.14	0.92	0.0030
SE Rockies	49	25	-1.34	0.07	26	-15.05	<0.01	0.06	0.07	0.94	0.0030
Overall	322	120	-0.14	0.52	50	-117.98	<0.01	0.07	0.51	0.97	0.0090

Table 3.2: Population structure for 24 hairy woodpecker populations based on mtDNA data; pairwise F_{ST} below diagonal, P values following corrections above diagonal. Bold indicates that values are significant at the critical value ($P=0.04$), following Benjamini-Hochberg corrections for multiple tests.

	AK	NL	SD	CO	AZ	UT	SCA	QCI	SOR	VI	WA	SAB	MT	ID	NEOR	CBC	NBC	CAB	NS	NC	WI	MI	ON	C Am
AK		<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.02	0.02	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
NL	0.44		<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.01	<0.01	<0.01	0.01	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
SD	0.83	0.76		0.04	0.02	0.25	<0.01	<0.01	0.02	0.01	0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
CO	0.81	0.77	0.12		0.16	0.03	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
AZ	0.82	0.76	0.18	0.04		0.07	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
UT	0.70	0.65	0.01	0.06	0.11		<0.01	<0.01	<0.01	0.02	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
SCA	0.86	0.82	0.57	0.52	0.53	0.48		0.06	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
QCI	0.78	0.74	0.49	0.52	0.47	0.43	0.07		0.10	0.20	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
SOR	0.48	0.42	0.22	0.33	0.27	0.21	0.20	0.10		0.52	0.04	0.01	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.01	0.01	<0.01	<0.01
VI	0.58	0.52	0.29	0.44	0.36	0.26	0.34	0.10	-0.03		0.12	0.01	<0.01	<0.01	<0.01	<0.01	0.02	0.01	<0.01	0.02	0.01	<0.01	<0.01	<0.01
WA	0.36	0.27	0.37	0.49	0.43	0.34	0.49	0.38	0.09	0.15		0.23	0.01	0.03	0.17	<0.01	0.01	0.04	0.01	0.10	0.16	0.01	0.01	<0.01
SAB	0.55	0.39	0.64	0.70	0.67	0.54	0.76	0.64	0.27	0.39	0.01		0.21	0.04	0.15	0.01	<0.01	<0.01	0.02	0.03	0.26	<0.01	0.01	<0.01
MT	0.65	0.45	0.84	0.82	0.84	0.72	0.86	0.81	0.47	0.67	0.19	0.04		0.23	0.34	<0.01	<0.01	<0.01	<0.01	<0.01	0.11	<0.01	<0.01	<0.01
ID	0.57	0.33	0.82	0.80	0.81	0.69	0.85	0.78	0.42	0.59	0.16	0.11	0.02		0.70	0.35	<0.01	0.06	0.02	0.01	0.54	0.02	0.43	<0.01
NEOR	0.57	0.33	0.78	0.78	0.77	0.64	0.83	0.72	0.35	0.47	0.09	0.03	0.01	-0.55		0.19	0.01	0.06	0.07	0.02	0.73	0.03	0.25	<0.01
CBC	0.40	0.28	0.76	0.77	0.77	0.66	0.81	0.74	0.43	0.55	0.18	0.13	0.11	0.01	0.05		<0.01	0.30	0.03	0.03	0.75	0.22	0.52	<0.01
NBC	0.17	0.20	0.65	0.71	0.67	0.56	0.75	0.62	0.32	0.32	0.18	0.29	0.42	0.28	0.25	0.17		0.22	<0.01	0.03	0.09	0.03	<0.01	<0.01
CAB	0.34	0.20	0.78	0.78	0.78	0.64	0.84	0.74	0.39	0.47	0.19	0.28	0.36	0.17	0.20	0.02	0.03		0.06	0.16	0.28	0.59	0.31	<0.01
NS	0.37	0.17	0.70	0.73	0.73	0.61	0.78	0.70	0.38	0.48	0.15	0.11	0.16	0.10	0.08	0.07	0.13	0.09		0.05	0.65	0.06	<0.01	<0.01
NC	0.41	0.24	0.64	0.70	0.67	0.53	0.77	0.64	0.29	0.34	0.10	0.14	0.30	0.20	0.17	0.11	0.13	0.08	0.09		0.15	0.19	0.03	<0.01
WI	0.41	0.24	0.73	0.75	0.74	0.60	0.81	0.70	0.34	0.42	0.09	0.03	0.07	-0.02	-0.02	-0.04	0.09	0.03	-0.03	0.07		0.50	0.50	<0.01
MI	0.31	0.21	0.73	0.75	0.75	0.63	0.80	0.71	0.39	0.48	0.18	0.18	0.22	0.11	0.14	0.02	0.09	-0.02	0.05	0.04	-0.01		0.13	<0.01
ON	0.42	0.28	0.76	0.77	0.77	0.66	0.81	0.74	0.42	0.53	0.18	0.15	0.14	0.00	0.03	-0.01	0.20	0.01	0.11	0.10	-0.01	0.03		<0.01
C Am	0.93	0.86	0.91	0.82	0.86	0.75	0.88	0.82	0.57	0.74	0.62	0.86	0.92	0.93	0.93	0.85	0.80	0.93	0.81	0.85	0.90	0.85	0.87	

Table 3.3: Pairwise F_{ST} among the four mtDNA groups identified in TCS network (Figure 3.2). Pairwise F_{ST} below diagonal, P values above diagonal. All values are significant at $P=0.05$.

	boreal-east	Pacific Coast	southeast Rocky Mtn.	Central America
boreal-east		<0.01	<0.01	<0.01
Pacific Coast	0.79		<0.01	<0.01
southeast Rocky Mtn.	0.76	0.54		<0.01
Central America	0.78	0.82	0.78	

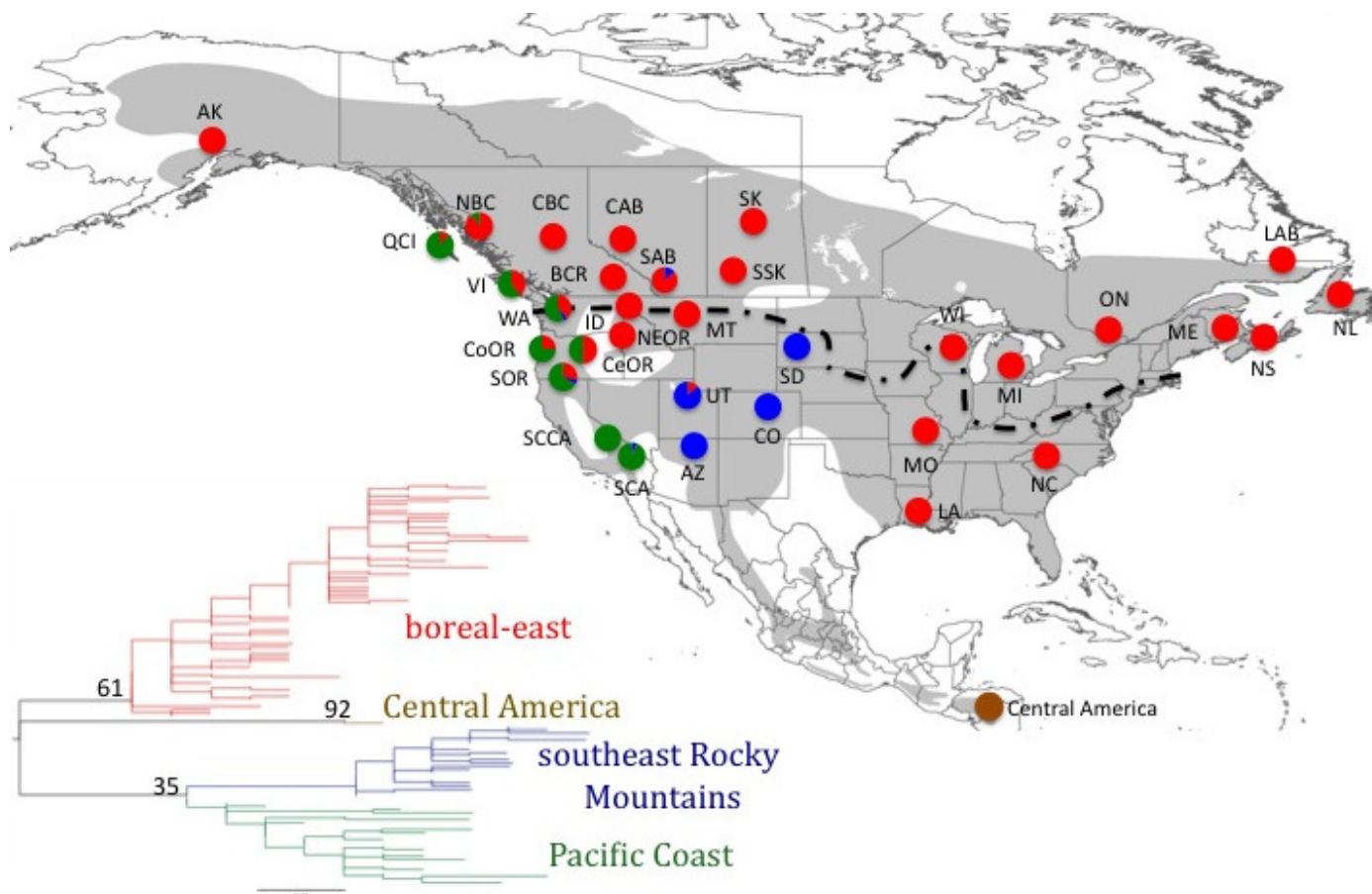


Figure 3.1: Geographic distribution of sampling sites for mtDNA analyses of hairy woodpeckers. Shaded area represents the current distribution of hairy woodpeckers throughout North and Central America (Ridgely *et al.*, 2007). Pie charts show population assignment to the four mtDNA groups based on TCS (Figure 3.2). Red=boreal-eastern North America group; green=Pacific Coastal group; blue=southeast Rocky Mountain group; brown=Central America group. Inset shows an unrooted maximum likelihood tree. Numbers at major nodes represent bootstrap values and colours correspond to TCS mtDNA groups. Dashed black line represents the extent of ice-sheets at the last glacial maximum (~21 kya).

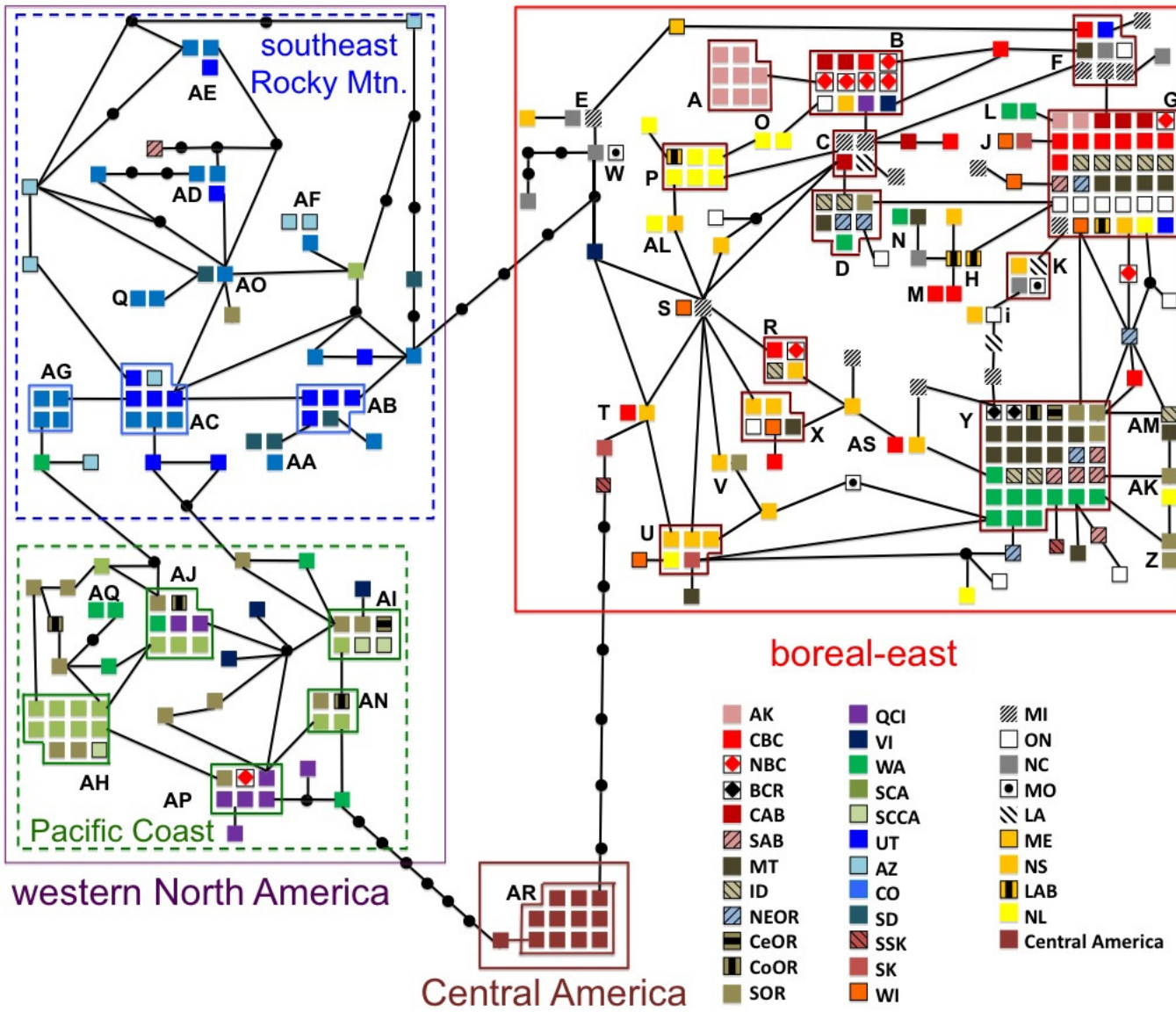


Figure 3.2: Haplotype network of the 120 mtDNA control region haplotypes (Appendices 4 and 5) for hairy woodpeckers across their range. Squares represent a single individual, with lines between haplotypes representing a single mutation and black dots representing inferred/missing haplotypes. Letter codes indicate the name of each haplotype as found in Appendix 4. Solid boxes show the three genetic lineages, with dashed boxes around the Pacific Coast and southeast Rocky Mountain groups showing subdivision within the western lineage as determined using a maximum likelihood tree (Figure 3.1 inset). Locations of sampling sites shown in Figure 3.1.

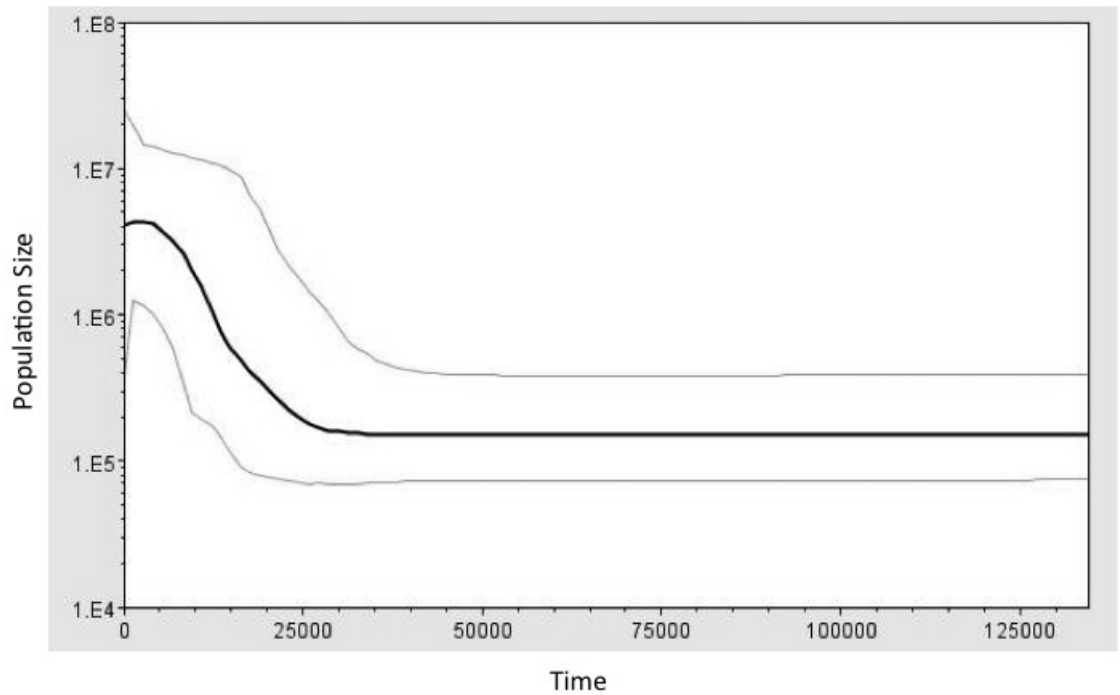
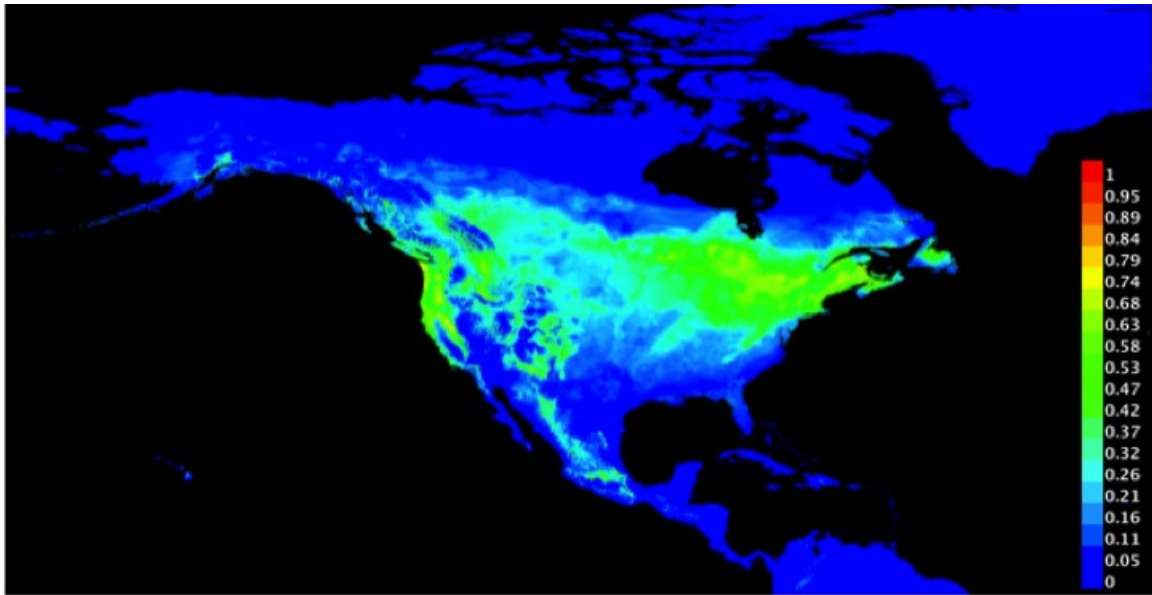
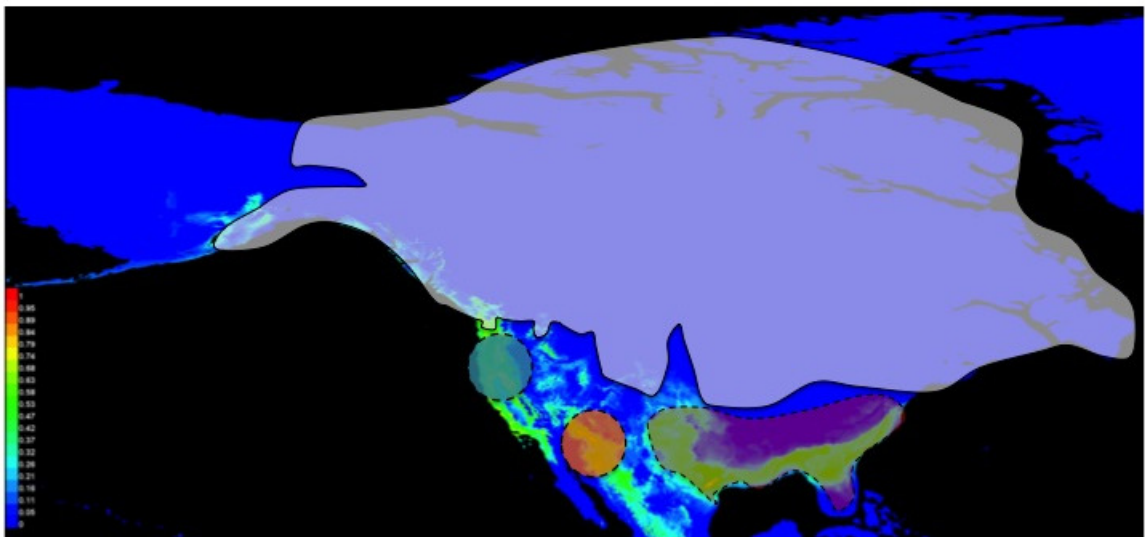


Figure 3.3: Bayesian skyline plot for hairy woodpecker populations (n=322). Time along x-axis denotes years before present with zero representing the current time, while effective population size (number of individuals) is along the y-axis. Grey lines above and below the black line represent the 95% upper and lower confidence limits. The straight line prior to expansion indicates hairy woodpecker populations were stable through the last glacial, with increases in population size occurring around 20 kya, coinciding with the approximate time for the onset of deglaciation.



a)



b)

Figure 3.4: Ecological niche modeling predictions. a) Current distribution based on climate data as predicted using the program MAXENT; b) Historical distribution (~21 kya) for hairy woodpeckers based on paleoclimate data as predicted using the program MAXENT. The three dotted lines in the US indicate location of forested refugia redrawn from Williams (2002). Darker colours (dark blue) indicate low prediction rates for presence of hairy woodpeckers, while brighter colours (yellow to red) indicate higher prediction rates for presence of hairy woodpeckers. Grey overlay represents ice sheet at last glacial maximum (~21 kya).

Chapter 4: General Discussion

4.1 Introduction

MtDNA and microsatellite analyses revealed well-defined population structure for hairy woodpeckers in North America. MtDNA patterns show Central and North American populations are genetically distinct from each other, and diverged at the beginning of the Wisconsin glacial (~96-113 kya). Eastern and western lineages in North America also split around this time with two separate lineages (Pacific Coast and southeast Rockies) in western North America splitting during the last glaciation (~34 kya). While the majority of the genetic splits predate the LGM, the most recent glacial cycles promoted genetic diversity in the hairy woodpecker and have had a role in shaping population structure and genetic patterns. Taxonomic classification currently recognizes 17 subspecies of hairy woodpeckers based on morphology and plumage, but genetic analyses of hairy woodpeckers across their range do not support current classification (Zink, 2004). Four different subspecies are recognized in the boreal-east region of North America, yet birds within this region were all genetically similar with the exception of Alaska. Across North America genetic patterns and population structure for hairy woodpeckers resemble patterns for other vertebrate (Arbogast *et al.*, 2001; Byun *et al.*, 1997) and plant species (Jarmillo-Correa *et al.*, 2009) and are concordant with genetic patterns found within biogeographic regions (Brunsfeld *et al.*, 2001; Soltis *et al.*, 2006).

4.1.1 Population structure

Both microsatellite and mtDNA markers show east - west population differentiation in North American taxa (Boulet and Gibbs, 2006) with microsatellite data often revealing greater population substructure than mtDNA, as has been reported for other species (Beadell *et al.*, 2010; Brito, 2007). Microsatellite analyses found as many as six hairy woodpecker subgroups in North America, while mtDNA revealed two major lineages in North America (including one that shows geographic clustering and can be subdivided into two groups); and a third in Central America. East - west differentiation follows the eastern extent of the Great Plains for microsatellite data, while mtDNA western haplotypes are primarily restricted to the western United States with northern latitudes in Canada containing mostly boreal-east haplotypes (with the exception of the Queen Charlotte Islands). Discordance between markers has been shown to occur in the presence of hybridization or contact zones between genetic groups (BrITO, 2007). Admixture of western and boreal-east haplotypes and genotypes occurs in the Pacific Northwest for my study, an area where hybridization for other species and contact zones are well documented (Swenson and Howard, 2005).

Genetic patterns in the Pacific Northwest are complex for many species due to genetic discontinuities between coastal and interior populations (Nielson *et al.*, 2001). Expansion from multiple refugia, north to south and south to north colonization along the coast, east to west colonization in the interior; and expansion from a single refugium with recent expansion have all been hypothesized as causes of genetic patterns in the Pacific Northwest (Brunsfield *et al.*, 2001). Genetic patterns

show that admixture for hairy woodpeckers in the Pacific Northwest is the result of this area being colonized from two different refugia, corresponding with patterns observed in vole (*Sorex*; Demboski and Cook, 2001) and squirrel (*Tamsciurius*; Arbogast, 1999) species.

4.1.2 Refugia

MtDNA revealed two different hairy woodpecker groups in North America and a third in Central America. Combining genetic patterns with ecological niche modeling, showed hairy woodpeckers were isolated in three North American ice-free refugia during the LGM. Location of refugia based on ecological niche modeling predictions are concordant with predicted locations of ice-free refugia in North America for tree species (Whitlock and Bartlein, 1997; Williams, 2003): in the eastern United States, southeast Rocky Mountains and along the Pacific Coast.

Previous studies have used single marker studies to test whether QCI acted as a refugium during the LGM (Burg *et al.*, 2005; Topp and Winker, 2008). For hairy woodpeckers, the Queen Charlotte Islands share haplotypes with other areas along the Pacific Coast, suggesting a shared history. Microsatellite patterns confirm the Queen Charlotte Islands are genetically unique compared to all other sites. This reflects isolation of individuals on the Queen Charlotte Islands from nearby populations on the mainland. While microsatellite patterns may reflect recent genetic differentiation (Krosby and Rohwer, 2009), the combination of multiple markers and ecological niche modeling suggests the Queen Charlotte Islands were

either part of a coastal refugium for hairy woodpeckers during the LGM or share a common refugial source with other populations along the Pacific Coast.

Boreal forests potentially expanded out of multiple refugia (Beringia and potentially more than one in the eastern US) following deglaciation (de Lafontaine *et al.*, 2010; Godbout *et al.*, 2010). MtDNA patterns for hairy woodpeckers make it difficult to determine where refugia were located in the eastern US, but suggest surviving lineages expanded from a common refugium. Genetic patterns correspond with patterns observed for other bird species (Ball and Avise, 1992), but contrast patterns observed in other animals, where expansion from multiple refugia located east and west of the Appalachian Mountains and Mississippi River has been proposed (reviewed in Soltis *et al.*, 2006). Additionally multiple refugia have been proposed for tree species in the southeastern Rocky Mountains (Mitton *et al.*, 2000). However, mtDNA patterns for hairy woodpeckers suggest expansion out of a single refugium within this region as haplotypes are shared between populations within this region.

4.1.3 Physical barriers

Results based on both mtDNA and microsatellite data do not suggest physical barriers such as the Cascade and Rocky Mountain ranges act as barriers to gene flow for hairy woodpeckers. As hairy woodpeckers are found at a range of elevations, perhaps gene flow across the mountains is not surprising. The fact that birds in British Columbia are genetically more similar to populations east of the Rocky Mountains, also suggests mountains are not impeding gene flow. On the other hand,

large expanses of open water like the Hecate Strait: separating the Queen Charlotte Islands from Northern British Columbia do act as barriers, as confirmed by estimates of gene flow for microsatellite data, which is supported by mtDNA patterns. Breaks in contiguous forest and paleogeographic barriers (e.g. glaciers) have had a greater influence on genetic patterns for hairy woodpeckers. Both mtDNA and microsatellites show genetic differentiation of populations separated by the Great Plains in the contiguous US and in the western US where contemporary barriers like the Columbia, Harney and Wyoming Basins cause discontinuities in forested habitat (Driese *et al.*, 1997). In addition paleogeographic barriers, such as montane glaciers, likely isolated populations in the southeast Rocky Mountains from the central Rocky Mountains, and would have delayed northern post-glacial expansion following deglaciation. MtDNA and microsatellite markers support historical barriers and habitat discontinuities acting as barriers to gene flow for hairy woodpeckers.

4.1.4 Patterns of recolonization

Post-glacial expansion and recolonization of previously glaciated areas followed the phalanx model (Hewitt, 1996). Within genetic groups, individuals separated by large geographic distances share mtDNA haplotypes and are genetically similar. Microsatellite data show concordant patterns despite greater population structure. Multiple analyses using BAPS and STRUCTURE show east - west differentiation, suggesting gradual post-glacial colonization of eastern and western North America.

At first glance the two markers seem to contradict each other, but analyzing microsatellites in combination with mtDNA provides greater insight into recent demographic processes. Contrasting patterns show the benefits and strengths of using a multi-locus approach to solve complex demographic processes (Carstens and Knowles, 2007). Routes of colonization are clear for mtDNA, as mtDNA retains historical patterns longer and analyses show that eastern populations expanded into the central Rockies and central Rockies were colonized by several common haplotypes and expansion occurred recently (~3.7 to 3.2 kya). Microsatellite data confirm mtDNA results, as central Rocky Mountain populations (ID, MT, NEOR and SAB) are unique from all other populations except northern Rocky Mountain populations (CBC; Table 3.2), the area where the central Rockies were likely colonized from.

4.1.5 Biogeographic patterns

Genetic patterns for hairy woodpeckers closely resemble genetic patterns found in North American trees. Concordance of tree and woodpecker genetic patterns shows sedentary species like woodpeckers are a good species for testing hypotheses about post-glacial expansion. Pollen data suggest mesic tree species expanded into the central Rockies recently (3 to 1.5 kya) (Mack *et al.*, 1978a; Mack *et al.*, 1978b) and expansion times for hairy woodpeckers into the central Rockies support these results. Expansion times and genetic patterns for boreal-east birds coincide with genetic patterns for eastern tree species like jack pine (*Pinus banksiana*), which colonized northern latitudes and western North America from an

eastern refugium ~10 kya (Godbout *et al.*, 2010; Godbout *et al.*, 2005). Close association of genetic patterns between woodpecker and tree species confirms woodpeckers followed the phalanx model of expansion and expanded out of multiple refugia (Hewitt, 1996).

4.2 Future Work

Future work should more closely examine Central American populations. Very little is known about hairy woodpeckers in Central America or in the Bahamas (Jackson *et al.*, 2002). More samples are required to better understand genetic patterns and the natural history of woodpeckers at the southern extremity of their range and their relationship to North American populations. Very little is known about the acoustic patterns of hairy woodpeckers across their range, and subsequent studies should be undertaken to see if acoustic variation and patterns are congruent with genetic patterns. Past studies have shown that differences in songs for birds are correlated with genetic differences, suggesting song dialects can arise as gene flow is restricted (Isler *et al.*, 2005; MacDougall-Shackleton and MacDougall-Shackleton, 2001). As hairy woodpeckers show genetic and phenotypic variability across their range, analyses of song may reveal further geographic variability for this species.

Genetic analysis has provided greater insight into the natural history and population structure of hairy woodpeckers. However further research looking at behavioural aspects and dispersal are required to learn more about individual species like the hairy woodpecker and for woodpeckers as a whole (Virkkla, 2006).

There is considerable bias towards studies in the breeding season where researchers have explored the effects of forestry practices on woodpeckers and habitat selection (Ripper *et al.*, 2007). Within the breeding season further studies are required to look at nesting and brooding behaviour, juvenile recruitment and dispersal for woodpeckers (Jackson *et al.*, 2002). While woodpeckers are viewed as sedentary, past work has suggested some individuals may make short distance winter migrations (Ouellet, 1977). Genetic patterns suggest very little movement by hairy woodpeckers, but in areas like Canada and the eastern United States it would be hard to detect movements since birds in this area are all genetically similar. Studies using radio-telemetry or geolocators (Stutchbury *et al.*, 2009) will provide greater insight into the movement patterns of hairy woodpeckers and determine how far juveniles disperse and if individuals undertake short distance migration.

References:

- Aagaard JE, Vollmer SS, Sorensen FC, Strauss SH (1995) Mitochondrial DNA products among RAPD profiles are frequent and strongly differentiated between races of Douglas-fir. *Molecular Ecology* **4**, 441-446.
- Arbogast BS (1999) Mitochondrial DNA phylogeography of the new world flying squirrels (*Glaucomys*): implications for Pleistocene biogeography. *Journal of Mammalogy* **80**, 142-155.
- Arbogast BS, Brown RA, Weigl PD (2001) Genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurus*). *Journal of Mammalogy* **82**, 302-319.
- Avise J, Walker D, Johns G (1998a) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265**, 1707-1712.
- Avise JC (2000) *Phylogeography: The history and formation of species* Harvard University Press, Cambridge, MA.
- Avise JC, Nelson WS (1989) Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* **243**, 646-648.
- Avise JC, Walker D, Johns GC (1998b) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265**, 1707-1712.
- Baker A, Marshall H (1997) Mitochondrial control region sequences as tools for understanding evolution. In: *Avian Molecular Evolution and Systematics* (ed. Mindell D), pp. 49-80. Academic Press, San Diego.
- Ball RM, Freeman S, James FC, Bermingham E, Avise JC (1988) Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 1558-1562.
- Ball RM, Jr., Avise JC (1992) Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk* **109**, 626-636.
- Barker WT, Whitman WC (1988) Vegetation of the Northern Great Plains. *Rangelands* **10**, 266-272.
- Barrowclough GF, Groth JG, Mertz LA, Gutiérrez RJ (2004) Phylogeographic structure, gene flow and species status in blue grouse (*Dendragapus obscurus*). *Molecular Ecology* **13**, 1911-1922.
- Beadell JS, Hyseni C, Abila PP, Azabo R, Enyaru JCK, Ouma JO, Mohammed YO, Okedi LM, Aksoy S, Caccone A (2010) Phylogeography and population structure of *Glossina fuscipes fuscipes* in Uganda: Implications for control of tsetse. *PLoS Negl Trop Dis* **4**, e636.
- Belkhir K (1999) genetix, version 4.0. A windows program for population genetic analysis. In: *Laboratoire Génome, Populations, Interactions, Université Montpellier, Montpellier, France*.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* **57**, 289-300.

- Bennett KD, Provan J (2008) What do we mean by 'refugia'? *Quaternary Science Reviews* **27**, 2449-2455.
- Boulet M, Gibbs HL (2006) Lineage origin and expansion of a neotropical migrant songbird after recent glaciation events. *Molecular Ecology* **15**, 2505-2525.
- 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.
- Brubaker LB, Anderson PM, Edwards ME, Lozhkin AV (2005) Beringia as a glacial refugium for boreal trees and shrubs: new perspectives from mapped pollen data. *Journal of Biogeography* **32**, 833-848.
- Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) A comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecology and Evolution in a Spatial Context* (eds. Silvertown J, Antonovics J), pp. 319-340. Blackwell Science, Oxford.
- Burg TM, Croxall JP (2001) Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. *Molecular Ecology* **10**, 2647-2660.
- 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.
- Byun SA, Koop BF, Reimchen TE (1997) North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. *Evolution* **51**, 1647-1653.
- Carstens BC, Knowles LL (2007) Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology* **16**, 619-627.
- Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004) Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Systematic Biology* **53**, 781-792.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.
- Corander J, Siren J, Arjas E (2008) Bayesian spatial modelling of genetic population structure. *Computational Statistics* **23**, 111-129.
- Corander J, Waldmann P, Sillanpaa M (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* **163**, 367-374.
- Crochet PA, Bonhomme F, Lebreton JD (2000) Molecular phylogeny and plumage evolution in gulls (Larini). *Journal of Evolutionary Biology* **13**, 47-57.
- de Lafontaine G, Turgeon J, Payette S (2010) Phylogeography of white spruce (*Picea glauca*) in eastern North America reveals contrasting ecological trajectories. *Journal of Biogeography* **37**, 741-751.
- DeChaine EG, Martin AP (2005) Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains. *American Journal of Botany* **92**, 477-486.

- Demboski JR, Cook JA (2001) Phylogeography of the dusky shrew, *Sorex monticolus* (Insectivora, Soricidae): insight into deep and shallow history in northwestern North America. *Molecular Ecology* **10**, 1227-1240.
- Diaconis P (2008) The Markov chain Monte Carlo revolution. *American Mathematical Society* **46**, 179-205.
- Driese KL, Reiners WA, Merrill EH, Gerow KG (1997) A digital land cover map of Wyoming, USA: a tool for vegetation analysis. *Journal of Vegetation Science* **8**, 133-146.
- Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.
- Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG (2003) Measurably evolving populations. *Trends in Ecology & Evolution* **18**, 481-488.
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **22**, 1185-1192.
- Ellegren H, Carlson A, Stenberg I (1999) Genetic structure and variability of white-backed woodpecker (*Dendrocopos leucotos*) populations in northern Europe. *Hereditas* **130**, 291-299.
- 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 genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.
- Galbreath KE, Hafner DJ, Zamudio KR, Agnew K (2010) Isolation and introgression in the Intermountain west: contrasting gene genealogies reveal the complex biogeographic history of the American pika (*Ochotona princeps*). *Journal of Biogeography* **37**, 344-362.
- Garamszegi LsZ (2006) Comparing effect sizes across variables: generalization without the need for Bonferroni correction. *Behavioral Ecology* **17**, 682-687.
- Gerlach G, Musolf K (2000) Fragmentation of Landscape as a Cause for Genetic Subdivision in Bank Voles. *Conservation Biology* **14**, 1066-1074.
- Gill FB, Mostrom AM (1993) Speciation in North American chickadees: I. Patterns of mtDNA divergence. *Evolution* **47**, 195-212.
- Godbout J, Beaulieu J, Bousquet J (2010) Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany* **97**, 1903-1912.
- Godbout J, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology* **14**, 3497-3512.
- Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.
- Goudet J (1999) PCA-GEN, Lausanne, Switzerland.

- Gugger PF, Sugita S, Cavender-Bares J (2010) Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Molecular Ecology* **19**, 1877-1897.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.
- Haavie J, Sætre GP, Moum T (2000) Discrepancies in population differentiation at microsatellites, mitochondrial DNA and plumage colour in the pied flycatcher — inferring evolutionary processes. *Molecular Ecology* **9**, 1137-1148.
- Hafner DJ, Sullivan RM (1995) Historical and ecological biogeography of nearctic pikas (Lagomorpha: Ochotonidae). *Journal of Mammalogy* **76**, 302-321.
- Harding RM (1996) Lines of descent from mitochondrial Eve: An evolutionary look at coalescence. In: *Progress in population genetics and human evolution* (eds. Donnelly P, Tavaré S), p. 15. Springer-Verlag Press, Berlin.
- Harpending HC (1994) Signature of ancient populations growth in a low-resolution DNA mismatch distribution. *Human Biology* **66**, 591-600.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.
- 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 FST. *Nat Rev Genet* **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 *Gnarosiphia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6112-6117.
- Hull JM, Hull AC, Sacks BN, Smith JP, Ernest HB (2008) Landscape characteristics influence morphological and genetic differentiation in a widespread raptor (*Buteo jamaicensis*). *Molecular Ecology* **17**, 810-824.
- Isler ML, Isler PR, Brumfield RT, Zink RM (2005) Clinal Variation in vocalizations of an antbird (Thamnophilidae) and implications for defining species limits. *The Auk* **122**, 433-444.
- Jackson JA, Ouellet HR, Jackson BJS (2002) Hairy woodpecker. In: *The birds of North America* (eds. Poole A, Gill F), p. 32. The Academy of Natural Sciences. The American Ornithologists' Union, Philadelphia, Washington, DC
- Jackson ST, Webb RS, Anderson HA, Overpeck JT, Webb III T, Williams JW, Hansen BCS (2000) Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews* **19**, 489-508.
- Jarmillo-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.

- Johansen AD, Latta RG (2003) Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Molecular Ecology* **12**, 293-298.
- Keyghobadi N, Roland J, Strobeck C (1999) Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology* **8**, 1481-1495.
- 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. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **266**, 695-700.
- Klicka J, Zink RM, Jon CB, McGillivray WB, Doyle TJ (1999) Evidence supporting the recent origin and species status of the timberline sparrow. *The Condor* **101**, 577-588.
- Knowles LL (2009) Statistical Phylogeography. *Annual Review of Ecology, Evolution, and Systematics* **40**, 593-612.
- Krosby M, Rohwer S (2009) A 2000 km genetic wake yields evidence for northern glacial refugia and hybrid zone movement in a pair of songbirds. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **276**, 615-621.
- Latch E, Dharmarajan G, Glaubitz J, Rhodes O (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* **7**, 295-302.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.
- MacDougall-Shackleton EA, MacDougall-Shackleton SA (2001) Cultural and genetic evolution in mountain white-crowned sparrows: Song dialects are associated with population structure. *Evolution* **55**, 2568-2575.
- Mack RN, Rutter NW, Bryant VM, Valastro S (1978a) Reexamination of postglacial vegetation history in northern Idaho: Hager Pond, Bonner Co. *Quaternary Research* **10**, 241-255.
- Mack RN, Rutter NW, Valastro S, Bryant VM, Jr. (1978b) Late Quaternary vegetation history at Waits Lake, Colville River Valley, Washington. *Botanical Gazette* **139**, 499-506.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.
- Manthey JD, Klicka J, Spellman GM (2011) Cryptic diversity in a widespread North American songbird: Phylogeography of the brown creeper (*Certhia americana*). *Molecular Phylogenetics and Evolution* **58**, 502-512.
- McCormack J, Bowen B, Smith T (2008) Integrating paleoecology and genetics of bird populations in two sky island archipelagos. *BMC Biology* **6**, 28.
- McLeod TK, MacDonald GM (1997) Postglacial range expansion and population growth of *Picea mariana*, *Picea glauca* and *Pinus banksiana* in the western interior of Canada. *Journal of Biogeography* **24**, 865-881.

- Milá B, McCormack JE, Castañeda G, Wayne RK, Smith TB (2007a) Recent postglacial range expansion drive the rapid diversification of a songbird lineage in the genus *Junco*. *Proceedings of the National Academy of Sciences of the United States of America* **274**, 2653-2660.
- Milá B, Smith TB, Wayne RK (2007b) Speciation and rapid phenotypic differentiation in the yellow-rumped warbler *Dendroica coronata* complex. *Molecular Ecology* **16**, 159-173.
- Miller EH, Walters EL, Ouellet H (1999) Plumage, size, and sexual dimorphism in the Queen Charlotte Islands hairy woodpecker. *The Condor* **101**, 86-95.
- Mitton JB, Kreiser BR, Latta RG (2000) Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Molecular Ecology* **9**, 91-97.
- Moritz C, Faith DP (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular Ecology* **7**, 419-429.
- Morris-Pocock JA, Steeves TE, Estela FA, Anderson DJ, Friesen VL (2010) Comparative phylogeography of brown (*Sula leucogaster*) and red-footed boobies (*S. sula*): The influence of physical barriers and habitat preference on gene flow in pelagic seabirds. *Molecular Phylogenetics and Evolution* **54**, 883-896.
- Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics* **97**, 145-163.
- Nielson M, Lohman K, Sullivan J (2001) Phylogeography of the tailed frog (*Ascaphus truei*): implications for the biogeography of the Pacific Northwest. *Evolution* **55**, 147-160.
- O'Connell LM, Ritland K, Thompson SL (2008) Patterns of post-glacial colonization by western redcedar (*Thuja plicata*, Cupressaceae) as revealed by microsatellite markers. *Botany* **86**, 194-203.
- Orange DI, Riddle BR, Nickle DC (1999) Phylogeography of a wide-ranging desert lizard, *Gambelia wislizenii* (Crotaphytidae). *Copeia* **1999**, 267-273.
- 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 HR (1977) *Biosystematics and ecology of Picoides villosus (L.) and Picoides pubescens (L.) (Aves: Picidae)*. Ph.D. thesis, McGill University.
- Paetkau D, Strobeck C (1996) Mitochondrial DNA and the phylogeography of Newfoundland blackbears. *Canadian journal of Zoology* **74**, 192-196.
- Peakall ROD, Smouse PE (2006) GenAlEx 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.
- Peterson A, Martinez-Meyer E, Gonzalez-Salazar C (2004) Reconstructing the Pleistocene geography of the *Aphelocoma* jays (Corvidae). *Diversity and Distribution* **10**, 237-246.
- Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.

- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231-259.
- Pielou EC (1991) *After the ice age: the return of life to glaciated North America* University of Chicago Press, Ltd., Chicago.
- Pierson JC (2010) *Genetic population structure and dispersal of two North American woodpeckers in ephemeral habitats*. Ph. D. thesis, University of Montana.
- Pierson JC, Allendorf FW, Saab V, Drapeau P, Schwartz MK (2010) Do male and female black-backed woodpeckers respond differently to gaps in habitat? *Evolutionary Applications* **3**, 263-278.
- Piertney SB, Maccoll ADC, Bacon PJ, Dallas JF (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Molecular Ecology* **7**, 1645-1654.
- Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective allele size using allele frequency data. *The Journal of Heredity* **90**, 502-503.
- Porter SC, Pierce KL, Hamilton TD (1983) Late Wisconsin mountain glaciation in the western United States. In: *Late-Quaternary environments of the United States* (ed. Wright Jr. HE), pp. 71-111. University of Minnesota Press, Minneapolis.
- Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.
- Primmer CR, Ellegren H, Saino N, Moller AP (1996) Directional evolution in germline microsatellite mutations. *Nat Genet* **13**, 391-393.
- Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Rajora OP, DeVerno L, Mosseler A, Innes DJ (1998) Genetic diversity and population structure of disjunct Newfoundland and central Ontario populations of eastern white pine (*Pinus strobus*). *Canadian Journal of Botany* **76**, 500-508.
- Rambaut A, Drummond A (2006) FigTree: tree drawing tool, version 1.3.1. Available at <http://tree.bio.ed.ac.uk/software/figtree/>. Institute of Evolutionary Biology, University of Edinburgh.
- Ramos-Osnins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**, 2092-2100.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.
- Ridgely R, Alinutt T, Brooks T, McNicol D, Mehlman D, Young B, Zook J (2007) Digital distribution maps of the birds of the western hemisphere, version 3.0. In: *NatureServe*, Arlington, Virginia, USA.
- Ripper D, Bednarz JC, Varland DE (2007) Landscape use by hairy woodpeckers in managed forests of northwestern Washington. *Journal of Wildlife Management* **71**, 2612-2623.
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552-569.
- Saab VA, Dudley J, Thompson WL (2004) Factors influencing occupancy of nest cavities in recently burned forests. *The Condor* **106**, 20-36.

- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Slatkin M, Hudson RR (1991) Pairwise comparison of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555-562.
- Soltis DE, Gitzendanner MA, Strreng DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution* **206**, 353-373.
- Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15**, 4261-4293.
- Stutchbury BJM, Tarof SA, Done T, Gow E, Kramer PM, Tautin J, Fox JW, Afanasyev V (2009) Tracking long-distance songbird migration by using geolocators. *Science* **323**, 896.
- Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist* **166**, 581-591.
- 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.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetic analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* (Submitted).
- Topp CM, Winker K (2008) Genetic patterns of differentiation among five landbird species from the Queen Charlotte Islands, British Columbia. *The Auk* **125**, 461-472.
- Turmelle AS, Kunz TH, Sorenson MD (2011) A tale of two genomes: contrasting patterns of phylogeographic structure in a widely distributed bat. *Molecular Ecology* **20**, 357-375.
- 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 tridactylus*). *Molecular Ecology Resources* **8**, 1152-1154.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.
- Vila M, Robles H, Ciudad C, Olea PP, Baglione V (2008) Isolation and characterization of 12 microsatellite markers in the middle-spotted woodpecker (*Dendrocopos medius*). *Molecular Ecology Resources* **8**, 415-417.
- Virkkula R (2006) Why study woodpeckers? The significance of woodpeckers in forest ecosystems. *Annales Zoologici Fennici* **43**, 82-85.
- Waltari E, Hijmans RJ, Peterson AT, S. NÁ, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: Comparing phylogeographic and ecological niche model predictions. *PLoS ONE* **2**, e563.

- Warner BG, Mathews RW, Clague JJ (1982a) Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of late Wisconsin glaciation. *Science* **218**, 675-677.
- Warner BG, W. MR, Clague JJ (1982b) Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of late Wisconsin glaciation. *Science* **218**, 675-677.
- Webb DM, Moore WS (2005) A phylogenetic analysis of woodpeckers and their allies using 12S, Cyt b, and COI nucleotide sequences (class Aves; order Piciformes). *Molecular Phylogenetics and Evolution* **36**, 233-248.
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
- Weir J, Schluter D (2004a) Ice sheets promote speciation in boreal birds. *Proc R Soc Lond B Biol Sci* **271**, 1881-1887.
- Weir JT, Schluter D (2004b) Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 1881-1887.
- Whitlock C, Bartlein PJ (1997) Vegetation and climate change in northwest America during the past 125 kyr. *Nature* **388**, 57-61.
- Williams JW (2003) Variations in tree cover in North America since the last glacial maximum. *Global and Planetary Change* **35**, 1-23.
- Wilson GA, Rannala B (2003) Bayesian Inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.
- Wilson PJ, Grewal S, Lawford ID, Heal JNM, Granacki AG, Pennock D, Theberge JB, Theberge MT, Voigt DR, Waddell W, Chambers RE, Paquet PC, Goulet G, Cluff D, White BN (2000) DNA profiles of the eastern Canadian wolf and the red wolf provide evidence for a common evolutionary history independent of the gray wolf. *Canadian journal of Zoology* **78**, 2156-2166.
- Wright S (1946) Isolation by distance under diverse systems of mating. *Genetics* **31**, 39-59.
- Zink RM (2004) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 561-564.
- Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17**, 2107-2121.
- Zink RM, Drovetski SV, Rohwer S (2002a) Phylogeographic Patterns in the Great Spotted Woodpecker *Dendrocopos major* across Eurasia. *Journal of Avian Biology* **33**, 175-178.
- Zink RM, Rohwer S, Drovetski S, Blackwell-Rago RC, Farrell SL (2002b) Holarctic phylogeography and species limits of three-toed woodpeckers. *The Condor* **104**, 167-170.

Appendix 1: Locality information listed by population. Id refers to individual identification for each sample (sampling site and individual bird number). Haplotype refers to the mtDNA haplotype individuals belong to as shown in Figure 3.2 and Appendix 4. Museum samples from Burke Museum (UWBM), Smithsonian (SMITH), Royal Alberta (RABM), University of Northern British Columbia (UNBC), Louisiana State (LSUMZ), University of Michigan (UMI), North Carolina Museum (NCMNS), University of Columbia (UCOM), Queen’s University (QU), American Museum of Natural History (AMNH), Canadian Museum of Nature (CMN), Royal Saskatchewan (RSKM), Royal BC (RBCM) and Field Museum Chicago (FMC).

Location	band number	id	haplotype	Museum	Latitude (°N)	Longitude (°W)
AK-Alaska						
Eagle River Rd, AK	812-50886	AKA001	A		64.24	-149.29
Houston, AK	UWBM# 53974 DAB 726	AKA002	A	UWBM	61.28	-149.95
Houston, AK	UWBM# 53988 DAB 740	AKA003	A	UWBM	61.28	-149.95
Houston, AK	UWBM# 53978 DAB 730	AKA004	G	UWBM	61.28	-149.95
Elmendorf Air Force Base, Anchorage, AK	tissue#B13383 voucher#601784	AKA005	G	SMITH	61.16	-149.47
Elmendorf Air Force Base, Anchorage, AK	tissue#B13395 voucher#601818	AKA006	A	SMITH	61.18	-149.48
Eagle River nature center, AK	1342-17004	AKA007	A		64.88	-147.04
Eagle River nature center, AK	1342-17005	AKA008	A		61.23	-148.73
North Fork Eagle river, AK	1342-17006	AKA009	A		61.30	-148.47
Two Rivers Road, AK	2331-08701	AKA010	A		64.88	-146.96
AZ Arizona						
Dorman Dr. Forest	1272-31650	AZ001	AF		35.15	-110.34
Dorman Dr. Forest	1272-31651	AZ002	AF		35.15	-110.34
XYRanch RD 2	1272-31652	AZ003	AC		35.16	-110.35
XYRanch RD 2	1272-31653	AZ004	AZ4		35.16	-110.35
XYRanch RD 2	1272-31654	AZ005	AZ5		35.16	-110.35
Lake Mary Rd 2	1272-31655	AZ006	AZ6		35.15	-110.35
Lake Mary Rd 3	1272-31656	AZ007	AZ7		35.15	-110.36
BCR- British Columbia Revelstoke						

Arrow Lakes Reservoir	1731-05816	BCR001	Y		50.98	-118.19
Revelstoke Dump , Revelstoke BC	1731-05817	BCR002	Y		51.02	-118.23
CAB-central Alberta						
Edmonton, AB	1603-90501	CAB001	G		53.53	-113.55
Edmonton, AB	1731-05302	CAB002	CAB2		53.48	-113.56
Mt. Robson, BC	991-19824	CAB003	B		53.02	-119.22
west of Alder Flats, Improvement District 11, AB	Z95.15.23, cat#30265	CAB004	G	RABM	52.92	-115.07
west of Alder Flats, Improvement District 11, AB	Z95.15.24, cat#30251	CAB005	C	RABM	52.92	-115.07
Chain Lakes, AB	Z98.6.26, cat#32886	CAB006	B	RABM	54.98	-113.50
Chain Lakes, AB	Z98.6.27, cat#32887	CAB007	G	RABM	54.98	-113.50
CBC-central British Columbia						
Prince George, BC	1731-05303	CBC001	T		53.89	-122.82
Fort St. James, BC	1731-05312	CBC002	G		54.47	-124.47
Prince George, BC	1731-05304	CBC003	CBC3		53.89	-122.82
Prince George, BC	1731-05305	CBC004	G		53.92	-122.73
Prince George, BC	1731-05306	CBC005	G		53.92	-122.73
Fort St. James, BC	1731-05308	CBC006	CBC6		54.40	-124.29
Fort St. James, BC	1731-05309	CBC007	B		54.40	-124.29
Fort St. James, BC	1731-05310	CBC008	CBC8		54.40	-124.29
Fort St. James, BC	1731-95311	CBC009	CBC9		54.40	-124.29
Fort St. James, BC	1731-05313	CBC010			54.41	-124.27
Fort St. James, BC	1731-05325	CBC011	M		54.40	-124.29
Prince George, BC	UNBC specimen # 08-5	CBC012	R	UNBC	53.92	-122.77
Fort St. James, BC	1731-05307	CBC013	M		54.40	-124.29
Prince George, BC	UNBC specimen # 07-17	CBC014	G	UNBC	53.92	-122.77
Prince George, BC	UNBC specimen # B-337	CBC015		UNBC	53.92	-122.77
Prince George, BC	UNBC specimen # 07-16	CBC016	G	UNBC	53.92	-122.77
Prince George, BC	UNBC specimen # B-208	CBC017	F	UNBC	53.92	-122.77

Prince George, BC	UNBC specimen # B-231	CBC018	AS	UNBC	53.92	-122.77
Prince George, BC	Frozen specimen	CBC019	G	Frozen	53.92	-122.77
Prince George, BC	UNBC specimen # 8083	CBC020	G	UNBC	53.92	-122.77
Hazelton, BC	019091	CBC021		RABM	55.25	-127.67
CeOR- central Oregon						
Prineville, Crook Co. OR	UWBM# 69484 AWH 026	CeOR001	Y	UWBM	44.30	-120.83
Sunriver, Deschutes Co. OR	UWBM# 64480 WLK 110	CeOR002	AI	UWBM	43.88	-121.44
CO-Colorado						
Rist Canyon (Fire Station 3), CO	1272-31611	CO001	CO1		40.62	-105.35
Rist Canyon (Fire Station 3), CO	1272-31612	CO002	CO2		40.62	-105.35
Rist Canyon, CO	1272-31613	CO003	Q		40.63	-105.33
Poudre National Forest, CO	1272-31614	CO004	AF		40.66	-105.56
20 km W of Solder (Jans), CO	1272-31615	CO005	CO5		39.78	-105.36
20 km W of Solder (Jans), CO	1272-31616	CO006	AE		39.78	-105.36
Rollands Pass Road (FS 149), CO	1272-31618	CO007	AG		39.91	-105.61
10 km NE of BlackHawk (Bob Clemans), CO	1272-31619	CO008	Q		39.82	-105.39
Pickle Gulch, CO	1272-31620	CO009	AC		39.84	-105.52
Pickle Gulch, CO	1272-31621	CO010	AE		39.84	-105.52
N of Cottonwood (Stuart Wheeler), CO	1272-31622	CO011	AC		39.78	-105.39
N of Cottonwood (Trent and Cindy Miller), CO	1272-31623	CO012	AG		39.77	-105.39
N of Cottonwood (Trent and Cindy Miller), CO	1272-31624	CO013	AG		39.77	-105.39
N of Cottonwood (Trent and Cindy Miller), CO	1272-31625	CO014	AD		39.77	-105.39
N of Cottonwood (Larry Turner), CO	1272-31626	CO015	AG		39.77	-105.40
N of Cottonwood (Larry Turner), CO	1272-31627	CO016	AT		39.77	-105.40
N of Cottonwood (Chuck Serrat), CO	1272-31628	CO017	AA		39.77	-105.38
N of Cottonwood (Chuck Serrat), CO	1272-31629	CO018	AC		39.77	-105.38
Cottonwood (Molly and David Nevin), CO	1272-31630	CO019	AD		39.78	-105.40
Cottonwood (Molly and David Nevin), CO	1272-31631	CO020	CO20		39.78	-105.40

Cottonwood (Molly and David Nevin), CO	1272-31632	CO021			39.78	-105.40
Cottonwood (Molly and David Nevin), CO	1272-31633	CO022			39.78	-105.40
N of Central City (HWY 119), CO	1272-31634	CO023			40.05	-105.53
Grand Lake (Dave and Marilyn Bindey), CO	1272-31635	CO024			40.26	-105.86
Grand Lake (Dave and Marilyn Bindey), CO	1272-31636	CO025			40.26	-105.86
Grand Lake (Dave and Marilyn Bindey), CO	1272-31637	CO026			40.26	-105.86
W of Spencer Heights (Forestry Road), CO	1272-31639	CO027			40.26	-105.86
W of Spencer Heights (Forestry Road), CO	1272-31638	CO028			40.26	-105.86
NW of Spencer Heights (Forestry Road), CO	1272-31640	CO029			40.68	-105.86
CoOR-coastal Oregon						
Cape Meares, OR	2331-88101	CoOR001	AJ		45.48	-123.97
Cape Meares, OR	2331-88102	CoOR002	Y		45.48	-123.97
Drift Creek Wilderness, OR	2331-88103	CoOR003	CoOR3		45.48	-123.91
Drift Creek Wilderness, OR	2331-88104	CoOR004	AN		45.48	-123.91
CR-Costa Rica (Central America)						
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19818	CR001	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19843	CR002	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19856	CR003	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19900	CR004	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19908	CR005	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19917	CR006	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19932	CR007	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19936	CR008	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19943	CR009	AR	LSUMZ	9.56	-83.71
Villa Mills, Cartago Province, Costa Rica	LSUMZ #19946	CR010	AR	LSUMZ	9.56	-83.71
ID-Idaho						
1037 Showalter Rd, Moscow ID	1272-31657	ID001	AM		46.77	-116.84
1037 Showalter Rd, Moscow ID	1272-31658	ID002	G		46.77	-116.84

1358 4 Mile Rd, Moscow, ID	1272-31659	ID003	Y		46.84	-116.94
1358 4 Mile Rd, Moscow, ID	1272-31660	ID004	G		46.84	-116.94
1358 4 Mile Rd, Moscow, ID	1272-31661	ID005	G		46.84	-116.94
2136 Roop Rd Cocolalla, ID	1271-31663	ID006	D		48.13	-116.64
2136 Roop Rd Cocolalla, ID	1272-31662	ID007	D		48.13	-116.64
Republic, Ferry, WA	UWBM#53352, CSW 3891	ID008	G	UWBM	48.47	-118.50
LeClerc Creek, Ione, Pend Oreille, WA	UWBM#54045, DAB 483	ID009	G	UWBM	48.53	-117.28
LeClerc Creek, Ione, Pend Oreille, WA	UWBM#54052, DAB 490	ID010	R	UWBM	48.53	-117.28
LeClerc Creek, Ione, Pend Oreille, WA	UWBM#54053, DAB 491	ID011	Y	UWBM	48.53	-117.28
LA-Louisiana						
East Jetty Woods, Cameron Parish, LA	LSUMZ #804	LA001	C	LSUMZ	31.24	-92.14
East Jetty Woods, Cameron Parish, Louisiana	LSUMZ #3840	LA002	K	LSUMZ	31.24	-92.14
Garner Ridge, , Cameron Parish, LA	LSUMZ #8532	LA003	LA3	LSUMZ	29.76	-93.78
LAB-Labrador						
Gosling Bay , Happy Valley-Goose Bay, Lab	1731-05341	LAB001	H		53.41	-59.60
Gosling Bay , Happy Valley-Goose Bay, Lab	1731-05342	LAB002	G		53.41	-59.60
25 Palliser Cr., Happy Valley-Goose Bay, Lab	1731-05343	LAB003	P		53.30	-59.68
25 Palliser Cr., Happy Valley-Goose Bay, Lab	1731-05344	LAB004	H		53.30	-59.68
ME-Maine						
Maine	catalog#22969	ME001	ME001	NCMNS	40.71	-74.01
MI-Michigan						
Grass Lake Twp, Jackson Co., MI	228238	MI001		UMI	42.25	-84.21
T45N, R11W, Sec9, Luce Co., MI	239375	MI002	F	UMI	43.23	-84.03
Scio Twp, Sec 30, Pinecross Lane, Washtenaw Co., MI	239134	MI003	C	UMI	42.33	-83.84
Ann Arbor, Washtenaw Co., MI	236254	MI004	AS	UMI	42.27	-83.73
Mio, 6.5 Mi E, Au Sable River, Oscoda Co., MI	227391	MI005	F	UMI	44.65	-84.13
Mio, 6.5 Mi E, Au Sable River, Oscoda Co., MI	227392	MI006	S	UMI	44.65	-84.13
Markey Twp, Roscommon Co., MI	227393	MI007	G	UMI	44.50	-84.59

Markey Twp, Roscommon Co., MI	227394	MI008	MI8	UMI	44.50	-84.59
Markey Twp, Roscommon Co., MI	227395	MI009	F	UMI	44.50	-84.59
Sands Twp, Marquette Co., MI	238712	MI010	MI10	UMI	46.42	-87.41
Colfax Twp., T15N, R 9W, Sec. 12, Mecosta Co., MI	238419	MI011	C	UMI	43.74	-86.08
Sylvan Twp, Spring Lake Drive, Washtenaw Co., MI	235601	MI012	MI12	UMI	42.27	-83.63
Hartland, Livingston Co., MI	235106	MI013	MI13	UMI	42.65	-83.75
Rapid River, Delta Co., MI	240775	MI014	MI14	UMI	45.70	-86.94
Ann Arbor, Washtenaw Co., MI	238243	MI015	E	UMI	42.27	-83.73
MO-Missouri						
Rolla, MO	1342-17001	MO001	MO1		37.83	-90.25
Rolla, MO	1342-17002	MO002	W		37.83	-90.25
Ashland, state road Y, MO	1342-17003	MO003	K		38.76	-91.88
Moniteau Co, MO	catalog#1301	MO004		UCOM	38.58	-92.58
Moniteau Co, MO	catalog#1298	MO005		UCOM	38.58	-92.58
MT-Montana						
Twin Peaks Rd, Helena, MT	1272-31601	MT001	AM		46.54	-112.19
Montana City, near Helena, MT	1272-31602	MT002	G		46.53	-111.99
Montana City, near Helena, MT	1272-31603	MT003	G		46.53	-111.99
GL, Glacier NP MT	RB0705M	MT004	Y	J. Woolf	48.55	-114.01
GL, Glacier NP MT	RB0708F	MT005	N	J. Woolf	48.50	-114.04
MA, Missoula? MT	MA0701M	MT006	Y	J. Woolf	46.85	-114.21
MA, Missoula? MT	MA0706F	MT007	F	J. Woolf	46.93	-114.60
Montana City, near Helena, MT	1272-31604	MT008	G		46.53	-111.99
Bridger Woods Rd, Bozeman, MT	1272-31606	MT009	Y		45.69	-110.90
GL, Glacier NP MT	RB0709M	MT010	D	J. Woolf	48.50	-114.04
GL, Glacier NP MT	RB0710F	MT011	Y	J. Woolf	48.50	-114.04
GL, Glacier NP MT	RB0711M	MT012	Y	J. Woolf	48.51	-114.03
GL, Glacier NP MT	RB0712F	MT013	Y	J. Woolf	48.51	-114.03

GL, Glacier NP MT	RB0714M	MT014	X	J. Woolf	48.50	-114.04
MA, Missoula MT	MA0707M	MT015	Y	J. Woolf	46.84	-114.21
MA, Missoula MT	MA0708M	MT016	Y	J. Woolf	46.83	-114.15
MA, Missoula MT	MA0709M	MT017	G	J. Woolf	46.85	-114.20
MA, Missoula MT	MA0710M	MT018	MT18	J. Woolf	46.84	-114.20
MA, Missoula MT	MA0711F	MT019	MT19	J. Woolf	46.83	-114.15
MA, Missoula MT	MA0712M	MT020	Y	J. Woolf	46.94	-114.58
MA, Missoula MT	MA0714F	MT021		J. Woolf	46.92	-114.57
MA, Missoula MT	MA0517	MT022		J. Woolf	46.68	-113.91
MA, Missoula MT	WM0701F	MT023		J. Woolf	46.84	-114.20
MA, Missoula MT	WM0702F	MT024		J. Woolf	46.68	-113.92
MA, Missoula MT	WM0704M	MT025		J. Woolf	46.68	-113.93
Helena, MT	2331-08702	MT026			46.54	-112.11
Helena, MT	1272-31605	MT027			46.54	-112.11
Helena, MT	1272-31607	MT028			46.54	-112.11
Helena, MT	1272-31608	MT029			46.57	-112.22
Frontier Town (Forestry Road), MT	1272-31609	MT030			46.61	-112.26
Frontier Town (Forestry Road), MT	1272-31610	MT031			46.61	-112.26
NBC-northern British Columbia						
Prince Rupert Butze Trail, NBC	1681-65886	NBC001	B	QU	54.32	-130.33
Tyhee Lake, NBC	852-44046	NBC002	NBC2	QU	54.31	-130.31
Prince Rupert Wildlife Rehab Centre	1731-05803	NBC003	G		54.33	-130.29
3928 Mountainview Ave, Thornhill BC	1731-05809	NBC004	AP		54.51	-128.54
3928 Mountainview Ave, Thornhill BC	1731-05810	NBC005	R		54.51	-128.54
3928 Mountainview Ave, Thornhill BC	1731-05811	NBC006	B		54.51	-128.54
Ferry Island	1731-05812	NBC007	NBC7		54.51	-128.57
The Brady's Thornhill BC	1731-05814	NBC008	B		54.46	-128.48
Ferry Island 2	1731-05813	NBC009	B		54.50	-128.57

3928 Mountainview Ave, Thornhill BC	1731-05815	NBC010	B		54.51	-128.54
NC-North Carolina						
Dillingham, Buncombe Co. NC	UWBM# 86862 RBB 532	NC001	W	UWBM	35.75	-82.41
Murchison, Buncombe Co. NC	UWBM# 86870 RBB 540	NC002	E	UWBM	35.83	-82.32
Graham Co. NC	UWBM# 87014 RBB 716	NC004	K	UWBM	35.88	-90.17
Haywood Co, North Carolina	catalog#15094	NC005	F	NCMNS	35.76	-79.02
North Carolina	catalog#15182	NC006		NCMNS	35.76	-79.02
North Carolina	catalog#15190	NC007		NCMNS	35.76	-79.02
North Carolina	catalog#15191	NC008		NCMNS	35.76	-79.02
North Carolina	catalog#15208	NC009		NCMNS	35.76	-79.02
North Carolina	catalog#15213	NC010	NC10	NCMNS	35.76	-79.02
North Carolina	catalog#15228	NC011	NC11	NCMNS	35.76	-79.02
North Carolina	catalog#15236	NC012		NCMNS	38.92	-93.40
Buncombe CO, North Carolina	catalog#15240	NC013		NCMNS	35.76	-79.02
North Carolina	catalog#15305	NC014		NCMNS	35.76	-79.02
North Carolina	catalog#19518	NC015	NC15	NCMNS	35.76	-79.02
North Carolina	catalog#19823	NC016		NCMNS	35.76	-79.02
NEOR-northeast Oregon						
Wallowa, OR	DOT-15782	NEOR001	NEOR1	AMNH	45.57	-117.53
Wallowa, OR	DOT-15783	NEOR002	G	AMNH	45.57	-117.53
Wallowa, OR	DOT-15784	NEOR003	Y	AMNH	45.57	-117.53
Wallowa, OR	DOT-15787	NEOR004	NEOR4	AMNH	45.57	-117.53
Wallowa, OR	DOT-15790	NEOR005		AMNH	45.57	-117.53
Grande Ronde River, Asotin, Asotin, WA	UWBM#59057, SAR 6998	NEOR006	D	UWBM	46.08	-116.98
Grouse Creek, Asotin, Asotin, WA	UWBM#59090, SVD 983	NEOR007	D	UWBM	46.01	-117.42
NL-Newfoundland						
Barachois PP, NL	991-19814	NL001	P		48.45	-58.43
Barachois PP, NL	991-19815	NL002	P		48.45	-58.43
Passadena, NL	991-19816	NL003	AL		49.01	-57.60

Passadena, NL	991-19817	NL004	P		49.01	-57.60
Deer Lake, NL	991-19818	NL005	O		49.18	-57.42
Deer Lake, NL	991-19819	NL006	P		49.18	-57.42
Eagle Mountain, NL	991-19820	NL007	AK		49.84	-57.26
Humber Valley Distr., NL	77679	NL008		CMN	46.98	-67.08
Humber Valley Distr., NL	77680	NL009	NL9	CMN	49.26	-57.38
Campground, Sir Richard Squires PP, NL	1731-05336	NL010	NL10		49.35	-57.17
Lomond Campground, Gros Morne NP, NL	1731-05338	NL011	O		49.46	-57.76
Western Brook Pond Trail, Gros Morne NP, NL	1731-05338	NL012	U		49.79	-57.84
Shallow Bay Trail, Gros Morne NP, NL	1731-05339	NL013	P		49.95	-57.75
Shallow Bay Trail, Gros Morne NP, NL	1731-05340	NL014	G		49.95	-57.75
NS-Nova Scotia						
Antigonish, NS	991-19810	NS001	AS		45.62	-61.99
Cape North, NS	991-19809	NS002	I		46.89	-60.53
Mt Hanley, NS	991-19801	NS003	AL		44.75	-65.13
Mt Hanley, NS	991-19802	NS004	K		44.75	-65.13
Middleton, NS	922-97502	NS005	NS5		44.96	-65.07
Middleton, NS	991-19803	NS006	V		44.96	-65.07
Middleton, NS	991-19804	NS007	R		44.96	-65.07
Mt Hanley, NS	991-19805	NS008	NS8		44.75	-65.13
Middleton, NS	991-19806	NS009	X		44.96	-65.07
Economy Lake, NS	991-19807	NS010	X		45.39	-63.91
Fundy NP, NB	991-19808	NS011	B		45.62	-65.04
Antigonish, NS	991-19811	NS012	NS12		45.62	-61.99
Antigonish, NS	991-19812	NS013	T		45.62	-61.99
Antigonish, NS	991-19813	NS014	NS14		45.62	-61.99
Saint John Co., NB	005137	NS015		RNBM	-45.28	-66.17
Kings Co., NB	007292	NS016		RNBM	-45.68	-65.38
Saint John Co., NB	008874	NS017		RNBM	-45.20	-66.27

Kings Co., NB	008937	NS018		RNBM	-45.50	-65.97
Charlotte Co., NB	009561	NS019		RNBM	-44.73	-66.75
Two Neck Road, Quispamsis, NB	1731-05326	NS020			45.48	-65.92
Two Neck Road, Quispamsis, NB	1731-05327	NS021	U		45.48	-65.92
47 Silas Lewis Road, Second North River, NB	1731-05328	NS022	NS22		46.06	-65.05
47 Silas Lewis Road, Second North River, NB	1731-05329	NS023	G		46.06	-65.05
70 Browns Place Rd, Middle Musquodoboit, NS	1731-05330	NS024			45.03	-63.08
70 Browns Place Rd, Middle Musquodoboit, NS	1731-05331	NS025	U		45.03	-63.08
218 Fraser Road, East Hants, NS	1731-05332	NS026			45.11	-63.64
218 Fraser Road, East Hants, NS	1731-05333	NS027			45.11	-63.64
Near Lewis Lake, Musquodoboit Valley, NS	1731-05345	NS028			45.00	-63.01
1814 Fairmont Road, Antigonish, NS	1731-05334	NS029	U		45.72	-61.94
1814 Fairmont Road, Antigonish, NS	1731-05335	NS030			45.72	-61.94
ON-Ontario						
Mun. Rég. de Cté de Pontiac, QC	77992	ON001	G	CMN	45.63	-74.08
Mun. Rég. de Cté de Pontiac, QC	77993	ON002	G	CMN	46.47	-77.68
Mun. Rég. de Cté de Pontiac, QC	77994	ON003	G	CMN	46.47	-77.68
Mun. Rég. de Cté de Pontiac, QC	77995	ON004	ON4	CMN	46.33	-77.68
Mun. Rég. de Cté de Pontiac, QC	77996	ON005	ON5	CMN	46.47	-77.77
Mun. Rég. de Cté de Pontiac, QC	77997	ON006	G	CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	77998	ON007		CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	77999	ON008	X	CMN	46.47	-77.77
Mun. Rég. de Cté de Pontiac, QC	78000	ON009		CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	78001	ON010	I	CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	78002	ON011	ON11	CMN	46.43	-77.63
Mun. Rég. de Cté de Pontiac, QC	78003	ON012		CMN	46.43	-77.63
Mun. Rég. de Cté de Pontiac, QC	78004	ON013	F	CMN	46.43	-77.63
Mun. Rég. de Cté de Pontiac, QC	78005	ON014	ON14	CMN	46.43	-77.63
Mun. Rég. de Cté de Pontiac, QC	78006	ON015	G	CMN	46.58	-77.68

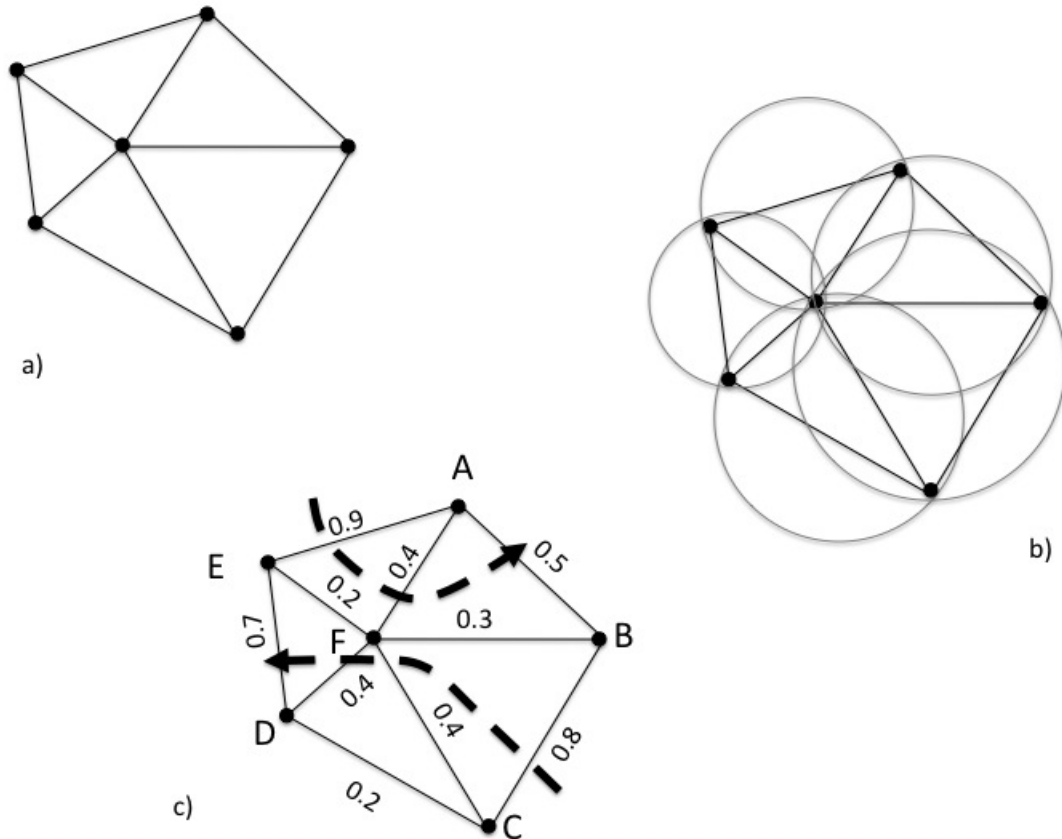
Mun. Rég. de Cté de Pontiac, QC	78007	ON016	B	CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	78008	ON017	ON17	CMN	46.23	-77.72
Mun. Rég. de Cté de Pontiac, QC	78009	ON018	G	CMN	46.50	-77.68
Mun. Rég. de Cté de Pontiac, QC	78010	ON019		CMN	46.58	-77.68
PAN-Panama (Central America)						
Dist. Boquete, Chiriquí Province, Panama	LSUMZ #28258	PA001	AR	LSUMZ	8.53	-80.78
Dist. Boquete, Chiriquí Province, Panama	LSUMZ #28323	PA002	PAN2	LSUMZ	8.53	-80.78
QCI-Queen Charlotte Islands/Haida Gwaii						
East Limestone, QCI	1581-43745	QCI001	AJ	QU	53.00	-132.00
Masset, BC	018076	QCI002	QCI2	RBCM	53.93	-132.15
Yakoun Lake, BC	018077	QCI003	AP	RBCM	53.32	-132.28
McClinton Bay, BC	018559	QCI004		RBCM	53.65	-132.58
Delkatla Parking Lot	1731-05804	QCI005	B		54.02	-132.05
Skate Board Park, QCI	1731-05805	QCI006	AP		53.26	-132.10
Skate Board Park, QCI	1731-05806	QCI007	QCI7		53.26	-132.10
Tlell Soccer Park	1731-05807	QCI008	AP		53.54	-131.95
Delkatla Parking Lot	1731-05808	QCI009	AJ		54.02	-132.05
Queen Charlotte Islands, Juskatla; BC	LSUMZ #19960	QCI010	AP	LSUMZ	53.37	-121.76
SAB-southern Alberta						
West Castle, AB	922-97503	SAB001	G		49.35	-114.42
Hwy 6, Waterton, S AB	1581-64945	SAB002	Y		49.11	-113.82
Castle River, Improvement District 5, AB	Z87.20.58, cat#23326	SAB003		RABM	49.30	-114.28
Lynx Creek, Improvement District 5, AB	Z87.20.96, cat#23356	SAB004	Y	RABM	49.47	-114.42
Lynx Creek, Improvement District 6, AB	Z87.20.100, cat#23560	SAB005	Y	RABM	49.47	-114.42
Beaver Mines, Improvement District 5, AB	Z91.17.16, cat#28973	SAB006		RABM	49.40	-114.33
Beaver Mines, AB	Z96.18.48, cat#32497	SAB007	SAB7	RABM	49.37	-114.38
Wardner, BC	019085	SAB008	SAB8	RABM	49.42	-115.42
Leihman trail, SAB	1142-49421	SAB009	Y		49.08	-113.97
SCA-southern California						
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #22959	SCA001		LSUMZ	34.12	-116.87

San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #24231	SCA002	SCA2	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #34542	SCA003	AJ	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #34543	SCA004	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #41112	SCA006	AI	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #41967	SCA007	AJ	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42038	SCA008	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42039	SCA009	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42040	SCA010		LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42066	SCA011	AJ	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42095	SCA012	AN	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42096	SCA013	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #42103	SCA014	AH	LSUMZ	34.12	-116.87
Frazier Mountain, Ventura Co., CA	LSUMZ #45537	SCA015	SCA15	LSUMZ	34.77	-119.77
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #51886	SCA016		LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #51940	SCA017		LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #53167	SCA018	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #53252	SCA019	AH	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #53277	SCA020	AN	LSUMZ	34.12	-116.87
San Bernardino Mtns., San Bernardino Co., CA	LSUMZ #53311	SCA021	AH	LSUMZ	34.12	-116.87
SCCA-south central California						
Sequoia National Forest, Tulare, CA	tissue#B21392	SCCA001	AI	SMITH	35.93	-118.39
Sequoia National Forest, Tulare, CA	voucher#636832	SCCA002	AI	SMITH	35.93	-118.39
3 mi. NE Sherman Pass, Tulare Co., California	tissue#B21398	SCCA004	AH	LSUMZ	36.78	-119.42
	voucher#636838					
	LSUMZ #24630					
SD-South Dakota						
South Dakota	SD0709F	SD001	AA	J. Woolf	44.22	-103.42
South Dakota	SD0710M	SD002	SD2	J. Woolf	44.22	-103.42
South Dakota	SD0711M	SD003	AA	J. Woolf	44.21	-103.40
South Dakota	SD0712F	SD004	AB	J. Woolf	44.21	-103.41
South Dakota	SD0713F	SD005	AT	J. Woolf	44.21	-103.40
SK-Saskatchewan						

Treebeard Trail, Prince Albert NP, SK	1731-05346	SK001	J		53.97	-106.29
Red Deer Trail Blue Loop, Prince Albert NP, SK	1731-05348	SK002	U		53.94	-106.06
Red Deer Trail Blue Loop, Prince Albert NP, SK	1731-05348	SK003	SK3		53.94	-106.06
SOR-southern Oregon						
near Paisley Oregon	OR0701M	SOR001	SOR1	J. Woolf	42.98	-121.15
near Paisley Oregon	OR0702F	SOR002	AN	J. Woolf	42.97	-121.17
near Paisley Oregon	OR0703M	SOR003	Y	J. Woolf	42.95	-121.18
near Paisley Oregon	OR0704M	SOR004	AK	J. Woolf	42.93	-121.14
near Paisley Oregon	OR0705F	SOR005	Z	J. Woolf	42.93	-121.13
near Paisley Oregon	OR0706M	SOR006	SOR6	J. Woolf	42.93	-121.14
near Paisley Oregon	OR0707F	SOR007	SOR7	J. Woolf	42.93	-121.14
near Paisley Oregon	OR0708F	SOR008	SOR8	J. Woolf	42.95	-121.18
near Paisley Oregon	OR0709M	SOR009	AJ	J. Woolf	42.92	-121.14
near Paisley Oregon	OR0710M	SOR010	AP	J. Woolf	42.92	-121.13
near Paisley Oregon	OR0711F	SOR011	AH	J. Woolf	42.91	-121.14
near Paisley Oregon	OR0712M	SOR012	Y	J. Woolf	42.93	-121.14
near Paisley Oregon	OR0713F	SOR013	D	J. Woolf	42.93	-121.14
near Paisley Oregon	OR0717MH	SOR014	Y	J. Woolf	42.89	-120.92
near Paisley Oregon	OR0718MH	SOR015	Z	J. Woolf	42.97	-121.00
near Paisley Oregon	OR0719FH	SOR016	AI	J. Woolf	42.97	-121.00
near Paisley Oregon	OR0720MH	SOR017	SOR17	J. Woolf	42.97	-121.00
near Paisley Oregon	OR0716F	SOR018	AI	J. Woolf	42.91	-121.14
McCloud, Shasta Nat. Forest ca., Siskiyou Co., CA	227376	SOR019	V	UMI	41.32	-121.92
McCloud, Shasta Nat. Forest ca., Siskiyou Co., CA	227377	SOR020	AH	UMI	41.32	-121.92
McCloud, Shasta Nat. Forest ca., Siskiyou Co., CA	227422	SOR021		UMI	41.32	-121.92
Warner Mtns., Modoc Co., California	LSUMZ #40966	SOR022	SOR022	LSUMZ	41.6	-121.6
Bear CA	tissue#B21463 voucher#634998	SOR023	SOR023	SMITH	41.21	-120.15
SSK-southern Saskatchewan						
Assiniboia, SK	catA6371/accn16356 exn 13	SSK001	SSK1	RSKM	49.63	-105.99
Caron, SK	catA6714/accn17119 exn 8	SSK002	SSK2	RSKM	50.46	-105.88
UT-Utah						

Cache National Forest (Logging Rd.), UT	1272-31641	UT001	AD		41.41	-111.50
Cache National Forest (Logging Rd.), UT	1272-31642	UT002	UT2		41.41	-111.50
SW Woodruff (Hwy 39), UT	1272-31643	UT003	G		41.46	-111.50
SW Woodruff (Hwy 39), UT	1272-31644	UT004	AB		41.46	-111.50
Cache National Forest (Forestry Rd.), UT	1272-31645	UT005	UT5		41.51	-111.51
W of Woodruff (Cache Forestry Rd.), UT	1272-31646	UT006	UT6		41.53	-111.46
Snow Basin Rd., UT	1272-31647	UT007	AB		41.23	-111.85
3 Mile Creek, UT	2331-08706	UT008	AB		41.45	-111.85
Wasatche-Cache, FR 73, UT	1272-31648	UT009	AB		40.94	-111.43
Uinta NF, FR 83, UT	1272-31649	UT010	F		40.34	-111.05
Vernal, Daggett Co. UT	UWBM# 70871 JMB 1778	UT011	AC	UWBM	40.46	-109.53
Vernal, Daggett Co. UT	UWBM# 70870 JMB 1777	UT012	AC	UWBM	40.46	-109.53
Vernal, Daggett Co. UT	UWBM# 70869 JMB 1776	UT013	AE	UWBM	40.46	-109.53
Vernal, Daggett Co. UT	UWBM# 70868 JMB 1775	UT014	AC	UWBM	40.46	-109.53
Vernal, Daggett Co. UT	UWBM# 70867 JMB 1774	UT015	AC	UWBM	40.46	-109.53
VI-Vancouver Island						
Langford, BC	017898	VI001		RBCM	48.45	-123.50
Victoria, BC	019080	VI002	VI2	RBCM	48.43	-123.37
Victoria, BC	020023	VI003	VI3	RBCM	48.43	-123.37
Seal Bay Nature Park	1731-05801	VI004	VI4		49.74	-124.97
Westwood Lake	1731-05802	VI005	VI5		49.16	-123.99
Lazo-Comox Marsh	HAWO1	VI006	B		49.69	-124.89
WA-Washington						
Mt. Rainier Paradise, WA	2331-88105	WA001	N		46.78	-121.74
Lake Tapps 16318 37St. Cr. E., WA	2331-88106	WA002	AQ		47.22	-122.21
Lake Tapps 16318 37St. Cr. E., WA	2331-88107	WA003	AQ		47.22	-122.21
Newhalem, WA	1581-43887	WA004	WA4	QU	48.67	-121.25
Ellensburg, Kittitas Co. WA	UWBM# 69701 BKS 1420	WA005	Y	UWBM	47.05	-120.75
Ellensburg, Kittitas Co. WA	UWBM# 61805 KJP 50	WA006	Y	UWBM	47.05	-120.75
Ellensburg, Kittitas Co. WA	UWBM# 68227 MMH 001	WA007	L	UWBM	47.05	-120.75
Cle Elum, Kittitas Co. WA	UWBM# 74111 SVD 1296	WA008	WA8	UWBM	47.25	-120.07
Ellensburg, Kittitas Co. WA	UWBM# 61788 CDS 5073	WA009	WA9	UWBM	47.05	-120.75

Easton, Kittitas Co. WA	UWBM# 49986 CSW 3957	WA010	Y	UWBM	47.24	-121.18
Ellensburg, Kittitas Co. WA	UWBM# 72663 SVE 1288	WA011	Y	UWBM	47.05	-120.75
Ellensburg, Yakima Co. WA	UWBM# 57151 GAV 304	WA012	AJ	UWBM	47.00	-120.55
Ellensburg, Yakima Co. WA	UWBM# 57163 JMB 1573	WA013	D	UWBM	46.73	-120.70
Naches, Yakima Co. WA	UWBM# 72668 SAR 6306	WA014	Y	UWBM	47.00	-120.55
Ellensburg, Yakima Co. WA	UWBM# 62629 RAP 24	WA015	Y	UWBM	47.00	-120.55
Ellensburg, Yakima Co. WA	UWBM# 57147 GAV 300	WA016	Y	UWBM	47.00	-120.55
Ellensburg, Yakima Co. WA	UWBM# 57164 JMB 1577	WA017	Y	UWBM	47.00	-120.55
Ellensburg, Yakima Co. WA	UWBM# 57156 JMB 1576	WA018	Y	UWBM	47.00	-120.55
Ellensburg, Kittitas Co. WA	UWBM# 61817 MLB 25	WA019	WA19	UWBM	47.05	-120.75
Cle Elum, Kittitas, Washington	tissue#B07653 voucher#586093	WA020		SMITH	47.25	-120.07
Naches, Yakima, Washington	tissue#B07747 voucher#586125	WA021	Y	SMITH	46.85	-120.82
Naches, Yakima, Washington	tissue#B07748 voucher#621179	WA022	L	SMITH	46.85	-120.83
WI-Wisconsin						
New Franken, County Rd P, Brown Co, WI	363492	WI001	J	FMC	44.53	-87.83
Green Bay, Brown Co, WI	428786	WI002	WI2	FMC	44.52	-88.02
WI	428812	WI003	WI3	FMC	44.52	-88.02
Green Bay, Brown Co, WI	438443	WI004	X	FMC	44.52	-88.02
WI	452082	WI005	G	FMC	44.52	-88.02
WI	454862	WI006	S	FMC	44.52	-88.02



Appendix 2: Diagram demonstrating the principles of Delaunay triangulation and Monmonier's algorithm as performed in the program BARRIER. a) The program connects geographic points entered into the program and connects a series of three points with triangles. Under the principles of Delaunay triangulation points can only be connected via triangles as long as no more than three points are encircled by circumcircles. b) Circumcircles can intersect but if they encompass more than three points, the triangle will not be completed. Once triangulation is completed the program uses Monmonier's algorithm. c) To determine potential barriers, Monmonier's algorithm starts at the edge of the triangulation matrix, beginning with the highest value (F_{ST}) written on the line between two points. The program then moves towards the centre of the matrix towards the next highest value. The program will continue to move towards the highest statistical value until it reaches another edge. The program then starts again at the edge at the next highest statistical value and undergoes the same process again. Under the hypothetical dataset, the program discovered two potential barriers. The first separates point A from all other points, while the second separates points C and D from all others. Based on Delaunay triangulation and Monmonier's algorithm, points B, E and F would share similar genetic signatures, with gene flow occurring between the three points.

Appendix 3: Summary of expected (He) and observed (Ho) heterozygosity for individual locus and combined loci for each population. Total represents the overall averaged He and Ho per individual locus and combined loci for 23 populations. Asterisk (*) denotes missing data.

Pop	Dlu1		Dlu5		DMC111		DMC115		DMD118		Ptri3		Average	
	Ho	He	Ho	He	Ho	He	Ho	He	Ho	He	Ho	He	Ho	He
AK	0.11	0.10	0.57	0.79	0.60	0.57	0.90	0.63	0.70	0.51	0.70	0.79	0.60	0.56
AZ	0.14	0.34	0.67	0.81	0.57	0.46	1.00	0.74	1.00	0.80	1.00	0.60	0.73	0.62
CAB	0.50	0.40	0.60	0.58	0.86	0.61	0.86	0.77	0.86	0.83	0.67	0.81	0.72	0.67
CBC	0.62	0.71	0.65	0.75	0.86	0.73	0.90	0.78	1.00	0.87	0.95	0.88	0.83	0.79
CO	0.34	0.69	0.10	0.82	0.71	0.71	0.52	0.71	0.97	0.87	0.83	0.80	0.58	0.77
ID	0.43	0.62	0.29	0.69	0.86	0.78	1.00	0.72	0.86	0.84	0.86	0.79	0.71	0.74
MI	0.47	0.76	0.31	0.77	0.57	0.73	0.71	0.60	0.77	0.86	0.73	0.84	0.59	0.76
MO	0.25	0.53	0.75	0.78	1.00	0.80	0.75	0.66	0.67	0.72	1.00	0.81	0.74	0.72
MT	0.42	0.69	0.57	0.77	0.80	0.78	0.63	0.80	0.84	0.89	0.87	0.88	0.69	0.80
NBC	0.30	0.68	0.22	0.78	0.70	0.75	0.90	0.81	0.90	0.85	0.70	0.71	0.62	0.76
NC	0.50	0.76	0.69	0.76	0.67	0.67	0.83	0.74	0.64	0.86	1.00	0.76	0.72	0.76
NEOR	0.60	0.74	0.40	0.66	0.60	0.74	0.75	0.66	1.00	0.84	0.60	0.76	0.66	0.73
NL	0.31	0.71	0.82	0.83	0.79	0.79	0.92	0.80	0.93	0.79	0.77	0.84	0.76	0.79
NS	0.52	0.77	0.36	0.78	0.67	0.77	0.71	0.77	0.89	0.89	0.82	0.87	0.66	0.81
ON	0.58	0.61	0.47	0.71	0.69	0.79	*	*	0.75	0.86	0.67	0.75	0.53	0.62
QCI	0.57	0.41	0.22	0.85	0.50	0.41	0.88	0.63	0.56	0.49	0.75	0.76	0.58	0.59
SAB	0.38	0.61	0.57	0.73	0.63	0.63	1.00	0.79	0.63	0.81	0.75	0.80	0.66	0.73
SD	0.25	0.53	0.60	0.70	0.25	0.22	0.60	0.58	1.00	0.86	0.80	0.76	0.58	0.61
SOR	0.58	0.74	0.41	0.79	0.57	0.69	0.81	0.81	0.79	0.86	0.95	0.83	0.69	0.79
UT	0.33	0.52	0.40	0.82	0.64	0.76	0.87	0.80	0.93	0.82	0.93	0.79	0.68	0.75
VI	0.60	0.42	0.20	0.70	0.75	0.56	0.75	0.56	0.60	0.80	1.00	0.82	0.65	0.64
WA	0.47	0.66	0.38	0.76	0.58	0.81	0.75	0.81	0.90	0.90	0.95	0.86	0.67	0.80
WI	0.20	0.46	0.17	0.76	0.00	0.00	0.80	0.76	0.83	0.89	1.00	0.82	0.50	0.62
Average	0.41	0.59	0.45	0.76	0.65	0.64	0.78	0.69	0.83	0.81	0.84	0.80	0.66	0.71

Appendix 4: Variable sites table for hairy woodpecker mtDNA control region haplotypes (hap). Table shows the 50 variable sites for an 825 bp sequence. Sequences begin at site 14940 of the L strand (referenced from Pileated woodpecker (*Dryocopus pileatus*) genome; accession no. DQ780879).

Hap	Variable sites starting at 14940 L strand				
	1111111	1111122222	2222333344	4566667788	
	22447778	8899011125	5677700022	4588236868	8066798901
	1267370790	3503714924	6105603404	9058628260	1305721016
A	GTTGAGACCT	CACGATCTCT	CTCCCCATTC	GTCCCATTTA	CAGGAGTAGC
BT.
CT.	A.....
DT.	A.....	...A.....
ET.	..T.....	A.....
FT.	.C.....	A.....
GT.	.C.....	A.....	...A.....
HT.	.C.....	A.....	..AA.....
IT.	.C.....CT	A.....	...A.....
JT.	.C.....	A.T.....	...A.....
KT.	.C.....T	A.....	...A.....
LT.	.C.....	A..T.....	...A.....
MT.	.C.T.....	A.....	..AA.....
NTC	.CT.....	A.....	...A.....
OG.....T.
PG.....T.	A.....
QT..C	..TT.....T	ACTTT.....	TT.....
RT..T.
ST..T.	A.....
TT..T.	A.....G..
UT..T.	A.....	...A.....
VT..T.	A.....	..A.....
WT..T.	..T.....	A.....
XT..T.	.C.....	A.....
YT..T.	.C.....	A.....	...A.....
ZT..T.	.C..T.....	A.....	...A.....
AAT..TC	..T.....T	ACT-.....	TT.....
ABT..TC	..T.....T	ACT.T.....	TT.....
ACT..TC	..TT.....T	ACT.T.....	TT.....
ADT..TC	..TT....CT	A.TTT.....	TTA.....
AET..TC	..TT....CT	ACTTT.....	TT.....A.
AFT..TC	..TT.T...T	AC.TT.....	TT.....
AGT..CTC	..TT.....T	ACT.T.....	TT.....
AHT..CTC	..TT..G.CT	AC.....	TT.....
AIT..	T.....CTC	..TT..G..T	ACT.....	TT.....
AJT..CTC	..TT..G.CT	ACT.....	TT.....
AKT..	.G.....T.	.C..T.....	A.....	...A.....
ALT..	.G.....T.	A.....

AMT..	.G.....T.	.C.....	A.....	...A.....
ANT..	T.....CTC	..TT..G..T	AC.....	TT.....
AOT..TC	..TT.....T	ACTTT.....	TT.....
APT..	T.....CTC	..TT..G..CT	AC.....	TT.....
AQ	...G..TT.TC	..TT..G..CT	AC.....	TT.....
ART..	TG.....T.	..TT.....TT	.TA.....NN
AST..T.	.C.....A.....
AZ4A.T..TC	..TT.....T	ACT.T.....	TT.....
AZ5A.T..TC	..TT.....T	ACTTT.....	TT..N.NNNN
AZ6A.TT.TC	..TT....CT	AC.TT.....	TT....NNNN
AZ7T..	.G.....CTC	..TT....CT	A.T.T.....	TT.....
CAB2T.	A.....G..
CBC3T..T.	.C.....T	A.....
CBC6	NN.....T.	.C.....
CBC8	A.....T.	A.....G..
CBC9	A.....T..T.	.C.....	A.....	...A.....
CO1T..TC	..TT.....T	A.TTT.....	TT..G.....
CO2T.ATC	..T.....T	ACT.T.....	TT..G.....
CO5	.N.....T..TC	..TT.....T	AC..T.....	TT.....
CO20T..TC	..T.....T	AC..T.....	TT.....
CoOR3TT.CTC	..TT..G..CT	AC.....	TT.....A.
LA3T..T.	.C.....CT	A.....	...A.....N
ME1T.	.CT.....	A.....
MI4T..T.	.C.....C...	...A.....
MI8T..T.	.C.....C.	A.....	...A.....
MI10C.....A.....
MI12T..T.	.C.....C.C.
MI13T.	.C.....	A.....C.
MI14TC	A.....
MO1T..T.	.C.....	A.....	..AA..NNNN
MT18	.N.....T..T.	A.....	A.....	...A.....
MT19T..T.	.C.....	A...G.....	...A.....
NBC2	A.....T.
NBC7	T.....T.	.C.....	A.....	...A.....
NC10	NN.....T.	.C.....	A.....	.T.....
NC11T.	.CT.....	A.....	..AA.....
NC15T..	T.....T.	..T...C..	A.T.....
NEOR1	NNNNNNNNNN	NN...C..T.	.C.....	A.....	...A...NN
NEOR4T..CT.	.C.....	A.....	...A...NN
NL9G.....T.	A.....A.
NL10GT..CT.	A.....	...A.....
NS5T..T.	A.....	..AA..NNNN
NS8T..T.	A.T.....
NS12T.	..T.....	A.....C...
NS14G.....T.	.C.....	A.....	..AA.....
NS22T..T.	.C.....NNNN
PAN2T..	TG.....T.	..TT.....TT	TTA..NNNNN
QC4T.	..T.....T.	.C.....	A.....	...A.....

QC5	NNNNNNNNNN	NN.....CT.	A.....	...A.....
QC11T..T.	.C....G...	A.....
QC14T.T.	A.....	...A.....
QC17	..CA.....T.	A.T.....
QCI2	NNNNN..T..	T.....CT.	.CTT..G.CT	AC.....	TT.....
QCI7	.N.....T..	T.....CTC	..TT..G.CT	AC.....	TT....NNNN
SAB7T..T.	.C....G...	A.....	...A.....
SAB8T..TC	.CT.....CT	ACTTT.....	TTA.....
SCA2T..CTC	.CTT...CT	ACT.....	TT....NNNN
SCA15T..TC	..TT.....T	AC.TT.....	TT.....
SD2TT.TC	..T.....CT	AC..T.....	TT.....
SK3T..T.G.N
SOR1T..CTC	..TT..G.CT	AC.....	TT.....A.
SOR6T..TC	..TT..G..T	ACTTT.....	TT.....
SOR7T..CTC	..TT..G.CT	ACT.....	TT.....A.
SOR8TT.CTC	..TT..G.CT	AC.....	TT.....
SOR17T..	TG.....CTC	..TT..G.CT	ACT.....	TT.....
SOR22T..	T.....TC	..TT..G..T	ACT.....	TT...NNNNN
SOR23T..	TG.....CTC	..TT..G.CT	AC.....	TT....NNN
SSK1T..T.T.....G..
SSK2T..T.	.C.....	AC.....	...A.....
UT2T..	T.....TC	..TT.....T	ACT.T.....	TT.....
UT5T..	T.....TC	..TT.....T	ACT.T.....	TT...T....
UT6	.C.....T..TC	..T.....T	AC..T.....	TT.....
VI2T..	T.....CTC	..TT..G..T	ACT.....N	N.....
VI3T..	.N.....T.	AC.....NNNN
VI4	T.....CTC	..TT..G.CT	ACT.....	TT.....
VI5	.C.....T..	T.....CTC	..TT..G.CT	ACT.....	TT.....
WA4TT.CTC	..TT..G.CT	ACT.....	TT.....
WA8T..	T.....CTC	..TT..G..T	ACT.....	TT.....A.
WA9T..	T.....CT.	..TT..G..T	AC.....	TT.....
WA19T..CTC	..TT...CT	ACT.T.....	TT.....
WI2T.	.C.....A.....
WI3T..T.A.....T

Appendix 5: Chart showing distribution of haplotypes (hap, as found in Figure 3.2 and Appendix 4) in sampling sites (SCA and SCCA combined (SCA) and PAN and CR combined (CAm)). Singles represent haplotypes found in only one individual, Nh is number of haplotypes and #ind the number of individuals.

hap	AK	BCR	CAB	CBC	ID	LA	LAB	ME	MI	MO	MT	NBC	NEOR	NC	NL	NS	ON	SAB	SK	SSK	WI	WA	SOR	CeOR	CoOR	VI	QCI	SCA	AZ	CO	SD	UT	CAm	Total
A	8																																	8
B			2	1								5				1	1									1	1							12
C			1			1			2																									4
D					2						1		2										1	1										7
E									1					1																				2
F				1					3	1				1			1														1			8
G	2		3	7	5		1		1		4	1	1		1	1	6	1				1										1		36
H							2																											2
I																1	1																	2
J																			1			1												2
K						1				1				1		1																		4
L																							2											2
M				2																														2
N												1											1											2
O																2																		2
P							1									5																		6
Q																															2			2
R				1	1							1					1																	4
S									1													1												2
T				1													1																	2
U															1	3			1															5
V																	1							1										2
W										1				1																				2
X											1					2	1					1												5
Y		2			2							9		1					4				10	3	1	1								33
Z																								2										2

Appendix 6: Table of bioclimate variables used for ecological niche modeling (Figure 3.4). Rank indicates the importance of the bioclimate variable as determined using MAXENT.

Bioclimate variable	Description	Rank
BI01	Annual mean temperature	2
BI02	Mean diurnal temperature range	4
BI03	Isothermality (mean diurnal range/temperature annual range)	8
BI04	Temperature seasonality	14
BI05	Maximum temperature of warmest month	6
BI06	Minimum temperature of coldest month	12
BI07	Temperature annual range	9
BI08	Mean temperature of wettest quarter	17
BI09	Mean temperature of driest quarter	15
BI010	Mean temperature of warmest quarter	3
BI011	Mean temperature of coldest quarter	7
BI012	Annual precipitation	5
BI013	Precipitation of wettest month	18
BI014	Precipitation of driest month	16
BI015	Precipitation seasonality	13
BI016	Precipitation of wettest quarter	11
BI017	Precipitation of driest quarter	19
BI018	Precipitation of warmest quarter	10
BI019	Precipitation of coldest quarter	1