

**POPULATION GENETICS OF TWO NORTH AMERICAN PASSERINES WITH
MIXED MIGRATION STRATEGIES**

DANIKA JANELLE SCHRAMM
Bachelor of Science, St. Mary's University, 2018

A thesis submitted
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOLOGY

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

POPULATION GENETICS OF TWO NORTH AMERICAN PASSERINES WITH
MIXED MIGRATION STRATEGIES

DANIKA JANELLE SCHRAMM

Date of Defense: November 16, 2022

Dr. T. Burg
Thesis Supervisor

Professor

Ph.D.

Dr. A. Iwaniuk
Thesis Examination Committee Member

Associate Professor

Ph.D.

Dr. A. Zovoilis
Thesis Examination Committee Member

Associate Professor

Ph.D.

Dr. C. Goater
Chair, Thesis Examination Committee

Professor

Ph.D.

Abstract

This thesis used molecular markers to examine population genetics of two North American passerines. My study focused on golden-crowned kinglets and red-breasted nuthatches, forest dwelling birds with migratory and resident populations. I used restriction site-associated DNA sequencing to analyse single nucleotide polymorphisms in both species. Additionally, I used microsatellite markers to study red-breasted nuthatch populations. Golden-crowned kinglets show an east-west split caused by their time in multiple glacial refugia during the Pleistocene. I also examined differences between their resident and migratory populations and identified potential candidate genes for migration. In red-breasted nuthatches, I found little population structure despite their large range. However, some differentiation is present in peripheral populations. They likely experienced rapid expansion from one glacial refugium following the Pleistocene and gene flow between populations is maintained by irruptive migrations. My thesis provides an example of how species develop different genetic patterns in response to historical and contemporary processes.

Acknowledgments

I would like to begin by thanking my supervisor, Dr. Theresa Burg. The first time we spoke I told her I had very little genetic background, all I knew is that I liked birds. She responded with “you’ll learn.” At the time that was an intimidating prospect, but her guidance and support has helped me through to the end of this journey. Thank you for your patience and all your contributions into making my project a success.

To my committee members, Dr. Andrew Iwaniuk and Dr. Athanasios Zovoilis, thank you for challenging me and supporting me through this program. This has definitely not been a normal two years. I appreciate your flexibility and input while we all navigated these waters.

Thank you to all those who have provided samples for my thesis and the work beyond: Dr. Vicki Friesen at Queen’s University, Dr. Garth Spellman at the Denver Museum of Nature and Science, Dr. Ben Winger at the Museum of Zoology at the University of Michigan, and Dr. Andy Jones at the Cleveland Museum of Natural History.

None of this research would be possible without the contributions of the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant. Thank you to the Northern Scientific Training Program put together by the Government of Canada for funding my field work.

I could write a second thesis expressing how grateful I am for the past and current members of the Burg Lab. I started graduate school amid a pandemic, a recipe that had all of the ingredients for disaster. Without their support I would have packed up and run away into the woods approximately two years ago.

I would like to thank (soon to be Dr.) Dilini Abeyrama for her endless patience, assistance, and coffee runs. Your help and friendship have been a blessing. Bukola Oguntuase-Osagie, Thilini Thakshila, Kanishka Senevirathna and Aaron Veale, thank you for all your ideas and your dedication to our Friday lunches. Thank you to Dr. Brendan Graham for working with me in the field and providing your insights on my project.

Last, but certainly not least, I would like to express my gratitude to my family and friends for supporting me these past two years. Thank you to my lovely mother Debbie (and her dog Sadie, they're a package deal) for always being excited to see me and encouraging me to pursue my education. My close friends, Dr. Danielle Hilborn and Jenna LeBlanc, I appreciate every moment you have listened to my worries and replaced them with laughter.

And who would I be if I didn't extend my gratitude to the man who has been by my side through two degrees, Mr. Sullivan, my cat.

Table of Contents

Abstract	iii
Acknowledgements	iv
List of Tables	ix
List of Figures	x
List of Abbreviations	xi
Chapter 1: General Introduction	1
1.1 Background	1
1.1.1 Population genetics	1
1.1.2 Migration	1
1.1.3 Population differentiation	5
1.2 Study species	7
1.2.1 Rationale	7
1.2.2 <i>Regulus satrapa</i>	7
1.2.3 <i>Sitta canadensis</i>	8
1.3 Molecular markers	9
1.4 Thesis goals	10
1.5 Literature cited	13
Chapter 2: Population genomics of golden crowned kinglets (<i>Regulus satrapa</i>) using RADseq	19
2.1 Introduction	19
2.2 Methods	23
2.2.1 Sample collection and DNA extractions	23
2.2.2 RADseq	24
2.2.3 Data processing and filtering	24
2.2.4 Population analyses	25
2.2.5 Isolation-by-distance	26
2.2.6 Divergence time estimation	27
2.2.7 Aligning to a reference genome	27
2.2.8 Genome scans	28
2.2.9 Comparison of golden-crowned kinglets to previously identified candidate genes in birds	29
2.3 Results	29
2.3.1 Data processing and filtering	29
2.3.2 Population analyses	30
2.3.3 Isolation-by-distance	31
2.3.4 Divergence time estimation	32
2.3.5 Reference genome alignment	32
2.3.6 Genome scans	32
2.3.7 Gene ontology	34
2.3.8 Comparison of golden-crowned kinglets to previously identified candidate genes in birds	34
2.4 Discussion	35
2.4.1 Population structure and barriers to gene flow	35

2.4.2 Genome scans in east and west population	39
2.4.3 Differences in migratory and resident populations	39
2.5 Conclusions	43
2.6 Literature cited	55
Chapter 3: Population genetics of an irruptive migrant (<i>Sitta canadensis</i>) using microsatellites and RADseq	65
3.1 Introduction	65
3.2 Methods	69
3.2.1 Sample collection and DNA extraction	69
3.2.2 Microsatellite markers	69
3.2.2.1 Genotyping	69
3.2.2.2 Population structure	70
3.2.2.3 F_{ST} and heterozygosity	71
3.2.3 RADseq and SNP analyses	72
3.2.3.1 RADseq	72
3.2.3.2 Data processing and filtering	72
3.2.3.3 Population analyses	73
3.2.3.4 Isolation-by-distance	74
3.2.3.5 Aligning to a reference genome	75
3.2.3.6 Genome scans	75
3.3 Results	77
3.3.1 Microsatellite analyses	77
3.3.1.1 Population structure	77
3.3.1.2 F_{ST} and heterozygosity	78
3.3.2 RADseq and SNP analyses	78
3.3.2.1 Data processing and filtering	78
3.3.2.2 Population analyses	78
3.3.2.3 Isolation-by-distance	80
3.3.2.4 Aligning to a reference genome	80
3.3.2.5 Genome scans and outlier PCoA	80
3.4 Discussion	82
3.4.1 Population structure and trends	82
3.4.2 Molecular markers	83
3.4.3 Irruptive migration and the LGM	84
3.4.4 Peripheral populations	87
3.4.5 Sky islands	88
3.5 Conclusions	90
3.6 Literature cited	104
Chapter 4: General Discussion	114
4.1 Introduction and main findings	114
4.2 Limitations and future directions	117
4.3 Closing statement	120
4.4 Literature cited	121

Appendices

Appendix 1: Supplementary information for Chapter 2	125
1.1 Individual details of golden-crowned kinglet samples	126
Appendix 2: Supplementary information for Chapter 3	128
2.1 Pairwise F_{ST} and P -values from microsatellite analysis	129
2.2 Individual details of red-breasted nuthatch samples	131

List of Tables

Table 2.1 Sample sites and sizes of golden-crowned kinglets	45
Table 2.2 Observed and expected heterozygosities	46
Table 2.3 Pairwise F_{ST} heatmap for golden-crowned kinglets using RADseq	47
Table 2.4 Outlier scaffolds and the genes associated with them in east-west and migratory-resident comparisons of golden-crowned kinglets	48
Table 3.1 Sample sites and sizes of red-breasted nuthatches	91
Table 3.2 Microsatellite markers	92
Table 3.3 F_{ST} values from microsatellite analysis	93
Table 3.4 F_{ST} values from RADseq data in red-breasted nuthatches	94
Table 3.5 Observed and expected heterozygosities for both microsatellite and RADseq data in red-breasted nuthatches	95
Table 3.6 Outlier scaffolds and the genes associated with them in the sky island comparisons for red-breasted nuthatches	96

List of Figures

Figure 2.1 Range map and sample sites of golden-crowned kinglets	49
Figure 2.2 PCoA analyses of golden-crowned kinglets using RADseq	50
Figure 2.3 DAPC analyses of golden-crowned kinglets using RADseq	51
Figure 2.4 Ancestry matrices from RADseq data in golden-crowned kinglets	52
Figure 2.5 East vs. west groups genome scans in golden-crowned kinglets	53
Figure 2.6 Resident vs. migratory populations genome scans in golden-crowned kinglets	54
Figure 3.1 Range map and samples sites of red-breasted nuthatches	97
Figure 3.2 STRUCTURE plots using microsatellite data	98
Figure 3.3 PCoA of 27 red-breasted nuthatch populations using microsatellites	99
Figure 3.4 DAPC analysis using RADseq data in red-breasted nuthatches	100
Figure 3.5 UPGMA tree of red-breasted nuthatches using RADseq	101
Figure 3.6 MSN for red-breasted nuthatch populations using RADseq	102
Figure 3.7 PCoAs of sky island populations using only outlier SNPs	103

List of abbreviations, acronyms, and symbols

°C	degrees Celsius
π	pi/nucleotide diversity
bp	base pairs
bwa	Burrows-Wheeler Alignment Tool
DAPC	discriminant analysis of principal components
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
F_{ST}	fixation index
GCKI	golden-crowned kinglet
Gb	giga base pairs
H_e	expected heterozygosity
H_o	observed heterozygosity
HWE	Hardy-Weinberg Equilibrium
K	genetic cluster
Kb	kilo base pairs
LEA	LEA: An R Package for Landscape and Ecological Association Studies
LGM	last glacial maximum
McMC	Markov chain Monte Carlo
MgCl ₂	magnesium chloride
mM	millimolar
MSN	minimum spanning network
mtDNA	mitochondrial DNA
n	sample size
NF	National Forest
PC	principal component
PCA	principal component analysis
PCoA	principal coordinate analysis
PCR	polymerase chain reaction
Q	ancestry coefficient
RADseq	restriction site-associated DNA sequencing
RBNU	red-breasted nuthatch
s	seconds
SNAPP	SNP and AFLP Package for Phylogenetic analysis
SNP	single-nucleotide polymorphism
T_A	annealing temperature
μ M	micromolar
UPGMA	unweighted pair group method with arithmetic mean
vcf	variant call format

Populations

AK	Alaska
BC	British Columbia
CA	California
CAB	Central Alberta
CBC	Central British Columbia
CO	Colorado
HG	Haida Gwaii
ID	Idaho
LETH	Lethbridge
MB	Manitoba
MI	Michigan
MT	Montana
NB	New Brunswick
NBC	North British Columbia
NC	North Carolina
NEAB	North-east Alberta
NEWA	North-east Washington
NL	Newfoundland
NS/NB	Nova Scotia/New Brunswick
NV	Nevada
NWWA	North-west Washington
ON	Ontario
OR	Oregon
OR	Oregon
SAB	South Alberta
SD	South Dakota, Black Hills
SEAB	South-east Alberta, Cypress Hills
SEBC	South-east British Columbia
SK	Saskatchewan
SWBC	South-west British Columbia
UT	Utah
VI	Vancouver Island
WA	Washington
WY	Wyoming

Chapter 1: General Introduction

1.1 Background

1.1.1 Population genetics

Populations of the same species may evolve to be genetically distinct from one another due to limited gene flow between populations. These distinct populations can be of special conservation concern, as barriers to gene flow can increase inbreeding and genetic drift within a population (Daniels and Walters, 2000; Haas et al., 2010; Tallmon et al., 2004). This can negatively impact the genetic diversity of a population and put isolated populations at increased risk of population decline. Population differentiation can arise for a variety of reasons such as physical, environmental, and behavioural barriers. Breaks in habitat can be due to bodies of water, habitat fragmentation, mountain ranges, large areas of unsuitable habitat such as desert, and past periods of glaciation. Even small breaks in continuous habitat can result in populations becoming isolated, including in taxa with high dispersal capabilities (Adams and Burg, 2015; MacDonald et al., 2010; Perrin et al., 2020). Barriers to gene flow are not always physical and can be caused by behavioural differences in populations due to selective pressures, sexual selection, feeding habits and migratory behaviour (Delmore et al., 2020; Ruegg et al., 2014).

1.1.2. Migration

Migration is utilized by a variety of taxa to cope with the fluctuation of resources throughout the year, such as food and breeding locations (Pulido, 2007). Migratory behaviour can differ both between and within species where populations of the same species exhibit different migration patterns (Delmore et al., 2020; O'Malley et al., 2010;

Thompson et al., 2020). Resident populations do not migrate and remain in the same location year-round, while migratory populations contain individuals that breed in one area and overwinter in another. Migratory behaviours are generally described in terms of the direction an individual migrates (the orientation), distance traveled (i.e. altitudinal, short-, medium- and long-distance migrants), and the inclination to migrate (Delmore et al., 2020). Life-history and geographic factors can help explain some of the behavioural differences (Somveille et al., 2015). Approximately one fifth of the world's bird species migrate in some way, including those that migrate short distances to different altitudes (Pulido, 2007).

Several theories try to explain how migration in avian species could have originated. While there is a consensus that migration developed through small, gradual movements over time, there are a few possibilities on how this occurred (Gauthreaux, 1982; Levey and Stiles, 1992; Volker and Bruno, 2007; Winger et al., 2018). The “southern-home-theory” suggests that bird migration originated in the tropics to allow birds to breed at higher latitudes (Levey and Stiles, 1992). The “northern-home-theory” suggests that bird migration originated due to birds shifting their non-breeding activities away from the higher latitude breeding grounds (Gauthreaux, 1982). A separate theory argues that migration is more like hibernation and is used to survive seasonal changes (Winger et al., 2018). These theories are not mutually exclusive, as migration in birds has arisen independently in multiple avian lineages (Zink, 2002). The origin of migratory behavior is important to help understand the diversity of migratory behaviours both within and between species.

Birds undergo various metabolic and hormonal shifts both in preparation for and during their migratory expeditions (Dingle and Drake, 2007). These include changes in circadian rhythm and feeding habits, increases in fat stores, changes in muscle and organ sizes and increases in aerobic capacity (McCabe and Guglielmo, 2019). Changes in day length combined with endogenous cues trigger birds to begin preparing for and undergoing migration (Marra et al., 2005). During migration, birds use internal and environmental cues such as the stars and the earth's magnetic field to orient and navigate themselves throughout their journey (Wiltschko and Wiltschko, 1978). There are, however, strong genetic components to migration, some of which influence aspects of migration such as an individual's departure and arrival times as well as the orientation of their migration (Lundberg et al., 2017; Pearse et al., 2014).

While migration has been studied for decades, less is known about the relationship between the genome and migratory behaviour (Delmore et al., 2020; Liedvogel et al., 2011; Lundberg et al., 2017; O'Malley et al., 2010; Pearse et al., 2014; Peterson et al., 2013; Toews et al., 2019). It is critical to understand the genetics of migration because climate change, habitat loss and infrastructure development may cause birds to alter their migratory behaviours (Wilcove and Wikelski, 2008). Birds have small genomes (1.0-8.2 Gb) which makes sequencing cost effective and time efficient (Toews et al., 2015). Additionally, highly annotated reference genomes are available making it easier to understand gene function (Toews et al., 2015).

The genetic basis for migration was first seen in cross breeding studies of Eurasian blackcaps (*Sylvia atricapilla*) (Helbig, 1991). When blackcaps with different migratory pathways were bred, their offspring exhibited an intermediate pathway to those

of their parents (Helbig, 1991). This sparked more widespread studies of migration genomics. The results from these studies, however, are not consistent and do not demonstrate a common genetic mechanism of migration between species. For example, the orientation of migration in Swainson's thrush (*Catharus ustulatus*) is correlated with a gene region on chromosome 4, whereas a study on willow warblers (*Phylloscopus trochilus*) identified different gene regions on chromosomes 1, 3 and 5 (Delmore et al., 2016; Lundberg et al., 2017). The genes within these correlated regions also vary between species (Delmore et al., 2020; Delmore et al., 2016; Lundberg et al., 2017; Parody-Merino et al., 2019; Peterson et al., 2013). While the location of gene regions correlated to migration differ from species to species, migration in passerines appears to be controlled by only a few gene regions (Delmore et al., 2020; Lundberg et al., 2017; Toews et al., 2019). It is important to compare intraspecific migratory behaviour because the differences seen in the genome are more likely to be related to migratory behaviour, rather than the more numerous differences that would be seen between species.

Past studies have discovered common patterns of gene flow in migratory and resident populations. Compared to migrants, resident populations generally experience less gene flow and increased genetic drift, causing higher levels of population structure to arise (Delmore et al., 2020; Ruegg et al., 2014). As a result, resident populations are typically more genetically distinct from one another, while migratory populations tend to be more similar. Differing levels of genetic diversity between migratory and resident populations could potentially lend itself to determining a species' or populations' adaptability and vulnerability to habitat disturbance (Lundberg et al., 2017).

It is critical to study migration because if migratory and resident populations are genetically distinct, special care needs to be taken to preserve each group. In comparison to resident birds, migratory species typically utilize more habitat in both area and diversity. Many migratory birds require the use of breeding grounds, wintering grounds, migratory stop-over sites as well as migratory flyways (Kirby et al., 2008). Warmer temperatures have been linked to earlier arrival at the breeding sites for migratory birds, causing some of their population trends to be driven by climate change (Koleček et al., 2020). Climate change has influenced North American bird species to alter their ranges, with resident populations shifting northward (Rushing et al., 2020).

1.1.3. Population differentiation

In examining population structure related to migratory behaviour, it is impossible to ignore underlying population structure within a species caused by other factors. Physical, environmental, and behavioural barriers may cause populations to become genetically distinct. Genetically isolated populations can be of special concern due to increased occurrences of inbreeding and genetic drift resulting in loss of genetic diversity (Daniels and Walters, 2000; Haas et al., 2010; Tallmon et al., 2004). Gaps in continuous habitat can result in habitat fragmentation, reductions in gene flow and loss of genetic diversity. Gene flow can also be affected by geographical barriers such as rivers, lakes, and areas of unsuitable habitat (MacDonald et al., 2010; Perrin et al., 2020). Reduction in gene flow can cause populations to become isolated and evolve independently of one another through drift, mutation, and selection, resulting in population structure. For example, a previous study of black-capped chickadees (*Poecile atricapillus*) found that small breaks in southern Alberta's riparian areas can significantly influence genetic

variation (Adams and Burg, 2015). This indicates that even small gaps in habitat of high-dispersal species can result in population structure and limit gene flow between populations.

An excellent example of habitat that could potentially contain isolated populations is Cypress Hills Interprovincial Park, a “sky-island” straddling the Alberta and Saskatchewan border. Sky islands refer to areas of significantly increased elevation in comparison to their lowland surroundings. The Cypress Hills contains genetically isolated populations of snails and warbling vireos (*Vireo gilvus*); however, it is not known if the Cypress Hills consistently provides isolated habitat between bird species with high dispersal capabilities (Carpenter et al., 2022; Dempsey et al., 2020). White-breasted nuthatches (*Sitta carolinensis*) in Arizona sky islands have shown population differentiation (Manthey and Moyle, 2015). These previous studies suggest that other bird species may also develop genetically isolated populations in response to these unique habitats.

While contemporary barriers can often explain some of the population structure, historical barriers can also contribute to the current population structure. One significant historical barrier is widespread glaciation across North America, as during the Pleistocene ice sheets covered most of North America (Pielou, 1991). During the last glacial maximum (LGM), approximately 21,000 years ago, the ice sheets were at their greatest extent (Peltier, 1994). This was particularly significant for high latitude species, as these large ice sheets caused them to remain solely on the peripheries of the ice or in ice-free refugia (Pielou, 1991). When populations are physically isolated, they can diverge genetically from one another through genetic drift and selection pressure from different

habitats (Hickerson et al., 2010). In North America, many cases of genetic differentiation are attributed to this period of isolation in refugia, followed by re-colonization (Burg et al., 2006; Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004).

1.2 Study species

1.2.1 Rationale

Two common and widely distributed birds were chosen for this study. The golden-crowned kinglet (*Regulus satrapa*) and red-breasted nuthatch (*Sitta canadensis*) were chosen due to their similar distribution and contrasting migratory habits. Both have migratory and resident populations, making them good examples for studying how migratory behaviour may affect gene flow. Neither study species are currently of conservation concern (IUCN, 2022). Additionally, they are both primarily forest birds, which may be more sensitive to anthropogenic disturbances such as deforestation. Together they provide a system to demonstrate how birds with similar ranges and overlapping habitats may differ in their population genetics. Studying abundant and widespread birds may help provide insights into harder to capture and record species that are at conservation risk.

1.2.2. Regulus satrapa

Golden-crowned kinglets are a small (~5-7 g) passerine species widespread throughout North America. They primarily breed in mature conifer forests, however, those that migrate overwinter in a variety of forest types (Swanson et al., 2020). Populations in the east (Newfoundland to West Virginia) and west (Kodiak Island to

California) are year-round residents while central populations (southern British Columbia to central Quebec) migrate south into the United States every winter. These migratory individuals are typically short-distance migrants, however, the distance of their migration is flexible and can change based on food availability and climate (Swanson et al., 2020).

Currently, at least five subspecies are recognized, three of which breed in Canada and the United States (Swanson et al., 2020). All of these subspecies are based on morphological characteristics and there is no evidence to suggest that most of these subspecies are genetically distinct. Using microsatellites and mitochondrial DNA, at least two main genetic groupings have been identified (Burg et al., 2014; Graham et al., 2020). Individuals collected in the east (Ontario to Newfoundland) consistently separate from individuals collected in the west (Alberta to western Alaska). These genetic groups are likely the result of the most recent Pleistocene glaciation forcing these populations to be isolated in different glacial refugia. Studies also suggest that populations on Haida Gwaii are genetically distinct from the rest of the western populations (Burg et al., 2014; Graham et al., 2020). In this case, it is likely Hecate Strait which separates Haida Gwaii from mainland British Columbia acts as a barrier to gene flow. It is currently unknown where central individuals (Manitoba to Saskatchewan) fall in the genetic patterns of golden-crowned kinglets.

1.2.3. Sitta canadensis

Red-breasted nuthatches are another small, widespread passerine species found throughout North America. Unlike other nuthatches native to North America, red-breasted nuthatches prefer forests that have more spruce and fir trees, although they are frequently found in suburban areas and mixed wood forests (Ghalambor and Martin,

2020). Their breeding range stretches from north-west Alaska south to New Mexico, and spreads across Canada to northern Labrador (Ghalambor and Martin, 2020). Most individuals are year-round residents; however, the northernmost populations move south on a yearly basis (Harrap and Quinn, 1995). Red-breasted nuthatches are also known for their irruptive movements every two to five years when conifer cone production is relatively poor (Cannings et al., 1987; Dorn and Dorn, 1990). Even individuals that are resident in other years will irrupt southward to utilize better food sources.

Despite their large geographic range, there are currently no subspecies or any regional variation recognized. It has been suggested that they may have previously been distributed throughout the Holarctic, forming a homogenous species complex with several other closely related species (Ghalambor and Martin, 2020). Their dispersal and migratory capabilities may play a role in the lack of visible population variation in the species. Currently, there are no thorough genetic studies on red-breasted nuthatches despite their abundance and distinctive characteristics.

1.3 Molecular markers

To evaluate any potential population structure in red-breasted nuthatches, I utilized both microsatellite markers and single-nucleotide polymorphisms (SNPs) from restriction site-associated DNA sequencing (RADseq). As previous studies have performed in-depth analyses using microsatellites on golden-crowned kinglets (Burg et al., 2014; Graham et al., 2020), I have utilized SNPs with RADseq for this species. Microsatellites are mostly neutral, nuclear genetic markers that consist of tandem repeats of 1-6 nucleotides (Selkoe and Toonen, 2006). The flanking regions surrounding these areas are highly conserved, making them easy and practical to use within the same family.

They are highly variable, often being comprised of five to 40 repeats (Selkoe and Toonen, 2006). Additionally, microsatellites are biparental markers and have a high mutation rate which results in high levels of allelic diversity (Selkoe and Toonen, 2006; Vieira et al., 2016). Microsatellites can therefore lend themselves to studies looking at questions related to gene flow between populations, parental history and relatedness, genetic distance, and other measures of population differentiation (Selkoe and Toonen, 2006; Vieira et al., 2016).

Whereas the microsatellites used in my thesis represent neutral genetic markers (i.e. non-coding), using SNPs from RADseq allows me to look at thousands of markers across the genome, including those that may be associated with particular genes or areas of the genome under selection (Peterson et al., 2012). A SNP is a single nucleotide change in the DNA sequence. Although RADseq is more costly than using microsatellite markers, being able to use thousands of SNP markers allows for population sample sizes to smaller than with fewer microsatellite markers. SNPs are very effective in evaluating both broad and fine-scale population structure in many taxa (Garroway et al., 2013; Wang et al., 2022). Combining SNP and microsatellite data can increase the resolution of the study and allow me to look at both neutral and non-neutral patterns of genetic diversity.

1.4 Thesis goals

The overall objective of my research is to evaluate the population structure of both red-breasted nuthatches and golden-crowned kinglets to examine if and how migratory behaviour impacts their population genetics. My thesis strives to explore two main questions: 1) Are there genomic differences between resident and migratory populations of golden-crowned kinglets and red-breasted nuthatches? 2) Are there genetically distinct

populations of each species? And if so, what are the possible reasons for population structure in these species?

The first data chapter of my thesis is on golden-crowned kinglets. Previous research on this species has uncovered significant population structure, showing an eastern and a western genetic cluster. These studies have used both microsatellites and mitochondrial DNA (Burg et al., 2014; Graham et al., 2020). No published studies to date have used genome-wide SNP sampling to evaluate their populations. This allows me to look for genes or genomic regions correlated to the east-west split. The previous study on migratory behaviour of golden-crowned kinglets only used microsatellite markers (Graham et al., 2020) which are generally neutral markers and cannot be used to look for specific loci involved in migration. Genetic markers related to migratory behaviour are thought to be restricted to a few small gene regions (Delmore et al., 2020; Lundberg et al., 2017; Toews et al., 2019). In my research I use a principal coordinate analysis (PCoA), discriminant analysis of principal components (DAPC), and an ancestry matrix to test for population structure. Additionally, I use F_{ST} values to determine levels of population differentiation. I then use outlier analyses to select areas of the genome that are differentiated between both east and west populations, and then migratory and resident populations, to see what genes are located in these areas and may be under selection.

The second data chapter investigates red-breasted nuthatches, for whom no comprehensive genetic study of their populations has been published. Using eight variable microsatellite markers and SNPs from RADseq, I will test for population structure. Again, I use PCoA, DAPC, and an ancestry matrix to test for any population structure in the sampled populations. F_{ST} values are used to examine levels of population

differentiation between populations. Other population statistics such as expected heterozygosity (H_e) and observed heterozygosity (H_o) are evaluated to look at the population makeup of the species. Outlier analyses are performed with sky island populations to determine if differing selective pressures may be causing these unique habitats to become genetically isolated.

1.5 Literature Cited

- Adams, R. V., and Burg, T. M. (2015). Gene flow of a forest-dependent bird across a fragmented landscape. *PLoS One* 10, 1-22.
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664.
- Burg, T. M., Gaston, A. J., Winker, K., and Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, 15(9), 2409–2419.
- Burg, T., Taylor, S., Lemmen, K., Gaston, T., and Friesen, V. (2014). Postglacial population genetic differentiation potentially facilitated by a flexible migratory strategy in golden-crowned kinglets (*Regulus satrapa*). *Canadian Journal of Zoology*, 92(2), 163-172.
- Cannings, R. A., Cannings, R. J., and Cannings S. G. (1987). *Birds of the Okanagan Valley, British Columbia*. Royal British Columbia Museum, Canada.
- Carpenter, A. M., Graham, B. A., Spellman, G. M., Klicka, J., and Burg, T. M. (2022). Genetic, bioacoustics and morphological analyses reveal cryptic speciation in the warbling vireo complex (*Vireo gilvus*: Vireonidae: Passeriformes). *Zoological Journal of the Linnean Society*, 195(1), 45-64.
- Daniels, S. J., and Walters, J. R. (2000). Inbreeding depression and its effects on natal dispersal in red-cockaded woodpeckers. *The Condor*, 102, 482-491.
- Delmore, K. E., Toews, D. P. L., Germain, R. R., Owens, G. L., and Irwin, D. E. (2016). The genetics of seasonal migration and plumage color. *Current Biology*, 26(16), 2167–2173.
- Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., and Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *Elife*, 9, e54462.
- Dempsey, Z. W., Goater, C. P., and Burg, T. M. (2020). Living on the edge: Comparative phylogeography and phylogenetics of *Oreohelix* land snails at their range edge in Western Canada. *BMC Evolutionary Biology*, 20(1), 3.
- Dingle, H., and Drake, V. A. (2007). What is migration? *BioScience*, 57(2), 113–121.

- Dorn, J. L., and Dorn, R. D. (1999) *Wyoming Birds Second Edition*. Mountain West Publishing, Wyoming.
- Garroway, C. J., Radersma, R., Sepil, I., Santure, A. W., De Cauwer, I., Slate, J., and Sheldon, B. C. (2013). Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. *Evolution*, 67(12), 3488–3500.
- Gauthreaux, S. A. (1982). The ecology and evolution of avian migration systems. *Avian Biology*, 6, 93-168.
- Ghalambor, C. K., and Martin, T. E. (2020). Red-breasted Nuthatch (*Sitta canadensis*), version 1.0. *Birds of the World*.
- Graham, B. A., Carpenter, A. M., Friesen, V. L., and Burg, T. M. (2020). A comparison of neutral genetic differentiation and genetic diversity among migratory and resident populations of golden-crowned kinglets (*Regulus satrapa*). *Journal of Ornithology*, 161(2), 509–519.
- Haas, Sarah. E., Cox, J. A., Smith, J. V., and Kimball, R. T. (2010). Fine-scale spatial genetic structure in the cooperatively breeding brown-headed nuthatch (*Sitta pusilla*). *Southeastern Naturalist*, 9(4), 743–756.
- Harrap, S., and D. Quinn. (1995). *Chickadees, tits, nuthatches and treecreepers*. Princeton University Press, Princeton.
- Helbig, A. J. (1991). Inheritance of migratory direction in a bird species: A cross-breeding experiment with SE- and SW-migrating blackcaps (*Sylvia atricapilla*). *Behavioral Ecology and Sociobiology*, 28(1), 9–12.
- Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F., and Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54(1), 291–301.
- Ingold, J., Galati, R., Poole, A., and Gill, F. (1997). Golden-crowned Kinglet (*Regulus satrapa*). *The Birds of North America Online*.
- Kirby, J. S., Stattersfield, A. J., Butchart, S. H. M., Evans, M. I., Grimmett, R. F. A., Jones, V. R., O'Sullivan, J., Tucker, G. M., and Newton, I. (2008). Key

- conservation issues for migratory land- and waterbird species on the world's major flyways. *Bird Conservation International*, 18(S1), S49–S73.
- Knowles, L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, 10, 691–701.
- Koleček, J., Adamík, P., and Reif, J. (2020). Shifts in migration phenology under climate change: Temperature vs. abundance effects in birds. *Climatic Change*, 159, 177–194.
- Levey, D., and Stiles, F. (1992). Evolutionary precursors of long-distance migration: resource availability and movement patterns in Neotropical landbirds. *The American Naturalist*, 140(3), 447–476.
- Liedvogel, M., Akesson, S., and Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology and Evolution*, 26(11), 561–569.
- Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., and Bensch, S. (2017). Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. *Evolution Letters*, 1(3), 155–168.
- MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, S. E., and Nielsen, S. E. (2020) Gene flow and climate-associated genetic variation in a vagile habitat specialist. *Molecular Ecology*, 29(20), 3889–3906.
- Macfarlane, C. B. A., Natola, L., Brown, M. W., and Burg, T. M. (2016). Population genetic isolation and limited connectivity in the purple finch (*Haemorhous purpureus*). *Ecology and Evolution*, 6(22), 8304–8317.
- Manthey, J. D., and Moyle, R. G. (2015). Isolation by environment in white-breasted nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology*, 24(14), 3628–3638.
- Marra, P. P., Francis, C. M., Mulvihill, R. S., and Moore, F. R. (2005). The influence of climate on the timing and rate of spring bird migration. *Oecologia*, 142(2), 307–315.
- McCabe, B. J., and Guglielmo, C. G. (2019). Migration takes extra guts for juvenile songbirds: energetics and digestive physiology during the first journey. *Frontiers in Ecology and Evolution*, 7(381), 1–18.

- O'Malley, K. G., Ford, M. J., and Hard, J. J. (2010). Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences*, 277(1701), 3703-3714.
- Parody-Merino, A. M., Battley, P. F., Conklin, J. R., and Fidler, J. R. (2019). No evidence for an association between Clock gene allelic variation and migration timing in a long-distance migratory shorebird (*Limosa lapponica baueri*). *Oecologia*, 191, 843-859.
- Pearse, D. E., Miller, M. R., Abadia-Cardoso, A., and Garza, J. C. (2014). Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, 281(1783), 20140012.
- Peltier, W. (1994). Ice age paleotopography. *Science*, 265, 195–201.
- Perrin, A., Khimoun, A., Faivre, B., Ollivier, A., Pracontal, N., Théron, F., Loubon, M., Leblond, G., Duron, O., and Garnier, S. (2020). Habitat fragmentation differentially shapes neutral and immune gene variation in a tropical bird species. *Heredity*, 126, 148-162.
- Peterson, M. P., Abolins-Abols, M., Atwell, J. W., Rice, R. J., Milá, B., and Ketterson, E. D. (2013). Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *F1000 Research*, 2(115), 1-16.
- Peterson, B. K., Weber J. N., Kay E. H., Fisher H. S., and Hoekstra H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135.
- Pielou, E. C. (1992). *After the Ice Age: The Return of Life to Glaciated North America* (1st edition). University of Chicago Press, Chicago.
- Pulido, F. (2007). The genetics and evolution of avian migration. *BioScience*, 57(2), 165–174.
- Ralston, J., and Kirchman, J. J. (2013). Predicted range shifts in North American boreal forest birds and the effect of climate change on genetic diversity in blackpoll warblers (*Setophaga striata*). *Conservation Genetics*, 14(2), 543–555.

- Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., DeSante, D. F., Moore, F., and Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23(23), 5726–5739.
- Rushing, C. S., Royle, J. A., Ziolkowski, D. J., and Pardieck, K. L. (2020). Migratory behavior and winter geography drive differential range shifts of eastern birds in response to recent climate change. *Proceedings of the National Academy of Sciences*, 117(23), 12897–12903.
- Selkoe, K. A., and Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), 615–629.
- Somveille, M., Rodrigues, A. S. L., and Manica, A. (2015). Why do birds migrate? A macroecological perspective. *Global Ecology and Biogeography*, 24(6), 664–674.
- Swanson, D. L., Ingold, J. L., and Galati, R. (2020). Golden-crowned Kinglet (*Regulus satrapa*), version 1.0. *Birds of the World*.
- The IUCN Red List of Threatened Species*. (2022). IUCN Red List of Threatened Species. Version 2021-3.
- Tallmon, D. A., Luikart, G., and Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, 19(9), 489–496.
- Thompson, N. F., Anderson, E. C., Clemento, A. J., Campbell, M. A., Pearse, D. E., Hearsey, J. W., Kinziger, A. P., and Garza, J. C. (2020). A complex phenotype in salmon controlled by a simple change in migratory timing. *Science*, 370(6516), 609–613.
- Toews, D. P. L., Campagna, L., Taylor, S. A., Balakrishnan, C. N., Baldassarre, D. T., Deane-Coe, P. E., Harvey, M. G., Hooper, D. M., Irwin, D. E., Judy, C. D., Mason, N. A., McCormack, J. E., McCracken, K. G., Oliveros, C. H., Safran, R. J., Scordato, E. S. C., Stryjewski, K. F., Tigano, A., Uy, J. A. C., and Winger, B. M. (2015). Genomic approaches to understanding population divergence and speciation in birds. *The Auk*, 133(1), 13–30.

- Toews, D. P. L., Taylor, S. A., Streby, H. M., Kramer, G. R., and Lovette, I. J. (2019). Selection on *VPS13A* linked to migration in a songbird. *Proceedings of the National Academy of Sciences*, 116(37), 18272–18274.
- Vieira, M. L. C., Santini, L., Diniz, A. L., and Munhoz, C. de F. (2016). Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology*, 39(3), 312–328.
- Wang, X., Lu, B., Shao, L., Li, Z., Ali, A., Yu, F., Fu, Z., and Sun, F. (2022). Genome-wide SNPs reveal the fine-scale population structure of *Laodelphax striatellus* in China using double-digest restriction site-associated DNA sequencing. *Genomics*, 114(3), 110329.
- Weir, J., and Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings. Biological Sciences/The Royal Society*, 271, 1881–1887.
- Wilcove, D. S., and Wikelski, M. (2008). Going, going, gone: is animal migration disappearing. *PLoS Biology*, 6(7), e188.
- Wiltschko, R., and Wiltschko, W. (1978). Relative importance of stars and the magnetic field for the accuracy of orientation in night-migrating birds. *Oikos*, 30(2), 195–206.
- Winger, B. M., Auteri, G. G., Pegan, T. M., and Weeks, B. C. (2019). A long winter for the Red Queen: Rethinking the evolution of seasonal migration. *Biological Reviews*, 94(3), 737–752.
- Zink, R. M. (2002). Towards a framework for understanding the evolution of avian migration. *Journal of Avian Biology*, 33(4), 433–436.

Chapter 2: Population genomics of golden-crowned kinglets (*Regulus satrapa*) using RADseq

2.1 Introduction

Genetic differentiation between populations of the same species can arise for a variety of reasons. Physical, environmental, and behavioural barriers can all lead to populations becoming genetically distinct. These distinct populations may be of special concern, as barriers to gene flow can increase inbreeding and genetic drift within a population, negatively impacting genetic diversity and putting isolated populations at increased conservation risk (Daniels and Walters, 2000; Haas et al., 2010; Tallmon et al., 2004). Relatively small breaks in continuous habitat can result in populations becoming isolated, even in taxa with high dispersal capabilities (Adams and Burg, 2015; MacDonald et al., 2010; Perrin et al., 2020).

Contemporary barriers such as mountains, lakes, rivers, and areas of unsuitable habitat can explain some of the population structure in species we see today, however, historical barriers are often responsible for the development of genetically distinct populations (Pielou, 1991; Weir and Schluter, 2004). The most recent glacial maxima covered most of North America in ice sheets until approximately 21,000 years ago (Peltier, 1994; Pielou, 1991). High-latitude species were predominantly affected by the ice sheets, as they were forced to remain solely on the peripheries of the ice in ice-free refugia (Pielou, 1991). Many cases of population differentiation in different taxa, can be attributed to this period of isolation in refugia (Burg et al., 2006; Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004). This is primarily due to genetic drift over time and populations potentially residing in refugia with differing selective pressures or habitats (Hickerson et al., 2010). Two of the main ice sheets

covering North America were the Cordilleran ice sheet which covered southern Alaska and most of British Columbia, and the Laurentide ice sheet that stretched thousands of kilometers east of the Rocky Mountains to cover the rest of Canada and the northern United States (Peltier, 1994). Once the ice began to melt, species started to recolonize now ice-free North America, however, many maintained their genetic differences from their time in isolation (Pielou, 1991). Behavioural differences may also arise during this time, which may be reflected in population's genetic makeup.

Behavioural traits such as differing migratory strategies within a species can influence population structure (Delmore et al., 2020; Ralston and Kirchman, 2013; Ruegg et al., 2014). It has been suggested that on average resident populations experience less gene flow and increased genetic drift compared to migratory populations (Delmore et al., 2020; Ruegg et al., 2014). This causes higher levels of population structure to arise in resident populations, whereas migratory populations tend to be difficult to differentiate from one another. It is also hypothesized that the genomic differences in populations with differing migratory phenotypes are contained within only a few gene regions, however these regions do not appear to be consistent between species (Delmore et al., 2020; Delmore et al., 2016; Lundberg et al., 2017; Parody-Merino et al., 2019; Peterson et al., 2013). Migratory birds use breeding grounds, wintering grounds, migratory stop-over sites as well as migratory flyways to survive and reproduce (Kirby et al., 2008). Birds with migratory populations are of special interest to study as they utilize more habitat in both area and diversity than birds that do not migrate. If migratory and resident populations are genetically distinct in a species, special care must be taken to preserve their habitats. Factors such as climate change influence the behaviour and timing of

migration in migratory populations, indicating the need to protect migratory species. For example, warmer temperatures have been linked to birds arriving earlier at their breeding grounds, particularly in North America (Kolecek et al., 2020).

An example of one of North America's migratory birds is one of Canada's smallest passerines, the golden-crowned kinglet (*Regulus satrapa*) (Swanson et al., 2020). They are widespread throughout North America, and they primarily breed in mature conifer forests (Swanson et al., 2020). Populations in the east (Newfoundland to West Virginia) and west (Kodiak Island to California) are year-round residents, while central populations (southern British Columbia to central Quebec) migrate south into the United States every winter (Figure 2.1). The majority of these central migratory populations are short-distance migrants, although their migration strategies are flexible and may change based on availability of food and environmental conditions (Swanson et al., 2020).

There are currently five recognized subspecies of golden-crowned kinglet (Phillips, 1991; Swanson et al., 2020). *R. s. olivaceous* breeds from southeast Alaska and western British Columbia south to southwest Oregon. *R. s. apache* breeds from southern Alaska and central Yukon south to the Rocky Mountains. *R. s. satrapa* breeds from northeast Alberta east to Newfoundland and Labrador, encompassing the largest range of any of the subspecies. The remaining two subspecies (*R. s. aztecus* and *R. s. clarus*) are resident in central America and were not included in this study. A sixth subspecies has been suggested by Watson (1986), however, there is no genetic or morphological evidence to suggest the presence of a sixth subspecies. Previous genetic analyses using microsatellites and mitochondrial DNA (mtDNA) support the splitting of eastern and

western subspecies, however, there is currently no evidence to suggest that the two western subspecies are genetically different (Burg et al., 2014; Graham et al., 2020).

Apart from the east and west groupings, all recognized subspecies are based on morphological characteristics (Phillips, 1991). Size may vary across their range and individuals in the east are darker and greyer than their green-tinted western counterparts (Phillips, 1991; Sibley, 2014). The genetic groupings are likely the result of the Pleistocene glaciation causing population to be isolated in different glacial refugia. Studies also suggest that populations on Haida Gwaii are genetically distinct from other western populations, likely due to Hecate Strait which separates Haida Gwaii from mainland British Columbia acting as a barrier to gene flow (Burg et al., 2014; Graham et al., 2020).

Previous studies on golden-crowned kinglets have used mtDNA and microsatellites to evaluate their population structure. Microsatellites are typically neutral genetic markers (i.e. non-coding) and mtDNA is uniparentally inherited, and while both are widely used, they provide information on a small number of loci. Utilizing single-nucleotide polymorphisms (SNPs) from genomic DNA allows examination of thousands of markers across the genome, including those that may be associated with genes or areas of the genome under selection (Lundberg et al., 2017; Ravinet et al., 2018). A common method of finding these SNPs is using restriction site-associated DNA sequencing (RADseq) where DNA is fragmented prior to sequencing using restriction enzymes. While more expensive, RADseq requires fewer individuals per sample site than microsatellites, as RADseq provides data on thousands of markers (Nazareno et al., 2017). Both fine and broad scale population structure can be detected using SNPs from

RADseq (Garroway et al., 2013; Wang et al., 2022). This study strives to provide additional information about the population genetics of golden-crowned kinglets using SNPs from RADseq. Additionally, I attempt to identify genes or gene regions associated with the population structure and behavioural differences in this species. As past studies did not use genome-wide data, it is possible that differences in migratory phenotypes were missed as they are often restricted to a small number of genes or gene regions (Delmore et al., 2020; Delmore et al., 2016; Lundberg et al., 2017; Parody-Merino et al., 2019; Peterson et al., 2013).

I hypothesize that the most significant population structure will remain the same as previous studies, with the east and west populations being genetically distinct. There may also be many gene regions that are significantly different between the east and west groups, as they would have diverged thousands of years ago during the last glacial maximum (LGM). Between migratory and resident populations, I hypothesize that there are likely a few gene regions that differ significantly between the migratory phenotypes, even if they do not show distinct population differentiation overall.

2.2 Methods

2.2.1. Sample collection and DNA extractions

Between 2002 to 2021 blood samples were collected from adult golden-crowned kinglets. Birds were caught using mist nets and song playback and sampled during the summer breeding season (between April and July). Each bird was fitted with a numbered aluminum band to ensure no individual was sampled more than once. Samples were stored in ethanol at ambient temperature until they could be returned to the lab where they

were stored at -20°C. In total, 40 samples from 12 sample locations were used and all sites were classified as resident or migratory depending on their location (Figure 2.1; Table 2.1). To extract DNA from blood, I used a modified salting-out DNA extraction protocol (Miller et al., 1988).

2.2.2. RADseq

This project used a type of RADseq where three restriction enzymes were used (Peterson et al., 2012). The extracted DNA samples were digested using the restriction enzymes *PstI* (6-base pair cutter), *NsiI* (5 base-pair cutter) and *MspI* (4 base-pair cutter). Library preparation was performed at Université Laval and samples were sequenced on an Illumina NovaSeq 6000 with paired-end reads at Génome Québec.

2.2.3. Data processing and filtering

Raw sequence reads were analyzed using Fastqc 0.11.5 to generate quality reports to ensure sequencing was successful and no major errors occurred during sequencing (Andrews, 2011). Raw reads obtained from Génome Québec were demultiplexed using Sabre 1.00 to assign reads to individuals. Adapter and barcode sequences were removed from the reads using Cutadapt (Martin, 2011). Cutadapt was also used to trim reads to 80 base-pairs in length as the variant calling program requires all reads to be the same length.

Genotyping and variant calling was performed using Stacks 2.3e, a program specifically designed for restriction site data (Catchen et al., 2013). I used the Stacks *de novo* pipeline with default parameters to analyze my data without the use of a reference genome. The *population* function within the *de novo* pipeline indicates number of loci that are out of Hardy-Weinberg Equilibrium (HWE). The resulting variant call format

(vcf) file was filtered using VCFtools 0.1.16 (Danecek et al., 2011). SNPs with more than 90% missing data were removed before removing individuals with more than 70% missing data prior to downstream analyses. This filtered dataset contained 217,093 SNPs and 40 individuals.

2.2.4. Population analyses

Mean observed and expected heterozygosities for each population were calculated in Arlequin 3.5.2.2 (Excoffier and Lisher, 2010). Arlequin was also used to calculate pairwise F_{ST} and the corresponding p -values. The p -values were corrected using the Benjamini-Hochberg method to account for multiple comparisons (Benjamini and Hochberg, 1995). The OR and NB samples were excluded from heterozygosity measures and pairwise F_{ST} calculations as these two sample sites only contained a single sample after filtering.

A principal coordinate analysis (PCoA) was performed to visualize potential population structure in the dataset. First, pairwise genetic distances were calculated between each individual and output as a matrix using the Adegenet package in R-Studio 2021.09.0 (Jombart, 2008). The PCoA was run and visualized in the excel add-in GenAlEx 6.5 using the genetic distance matrix calculated in Adegenet (Peakall and Smouse, 2006, 2012).

A discriminant analysis of principal components (DAPC) was performed using Adegenet 2.0.0 to again visualize potential population structure in the dataset (Jombart, 2008). While this analysis typically provides similar results to a PCoA, it includes an additional step where the distance between clusters is maximized. This can assist in

detecting very fine-scale population structure in a large SNP dataset. For the DAPC, first the data were transformed using a principal component analysis (PCA) to analyze differences between individuals. Then a discriminant analysis was performed to maximize the distance between clusters, where the number of clusters is defined by the user as a value of K. The DAPC was run for K=1 to K=4 even though the Bayesian Information Criterion indicated that the best value of K was 2.

An ancestry matrix was constructed using LEA: An R Package for Landscape and Ecological Association Studies, with the cross-entropy function enabled (Frichot and Francois, 2015). The program ran for 100 iterations with the number of assumed populations between 1 and 4.

The same three analyses (PCoA, DAPC and an ancestry matrix) were performed on a subset of the data where Stacks *de novo* was re-run with the *-write-single-snp* command to indicate that only 1 SNP per locus was to be included in the resulting vcf file. This was done to reduce redundancy in the data and account for potentially linked SNPs. After filtering using the same parameters as the full dataset, there were 26 individuals and 18,904 SNPs left for analyses. The analyses were repeated with this smaller dataset to maximize both the ability to detect population structure and cost efficiency based on how many samples were originally sequenced.

2.2.5 Isolation-by-distance

A Mantel test for matrix correspondence was performed using GenAlEx 6.5 to determine if there was a significant correlation ($P < 0.05$) between geographic distance and genetic distance. Genetic distances were previously calculated for the PCoA and were

used again to test for isolation-by-distance. The pairwise geographic distance matrix between all individuals was generated in GenAlEx 6.5 using sample location coordinates. The paired Mantel test was run with 99 permutations in GenAlEx 6.5.

2.2.6 Divergence time estimation

An estimation of the divergence time between east and west populations was calculated using the SNAPP (SNP and AFLP Package for Phylogenetic analysis) package in BEAST2 (Bouckaert et al., 2014; Bryant et al., 2012; Drummond et al., 2002). SNAPP calculates tree likelihoods from SNP markers using Bayesian Markov chain Monte-Carlo (MCMC) and estimates divergence times between populations. The simulation was run for 100,000 MCMC iterations.

2.2.7 Aligning to a reference genome

Prior to performing any genome scans, sequences were aligned to a golden-crowned kinglet reference genome at scaffold assembly (GenBank accession number VWZN000000000) as this was the best reference genome available (Feng et al., 2020). A reference genome is not required for general population analyses; however, it is critical when examining specific areas of the genome that are differentiated between populations. First, the reference genome was indexed using samtools 0.1.2 (Danecek et al., 2021). The Burrows-Wheeler Alignment Tool (bwa) was used to align the forward and reverse reads to the reference genome and mark low quality alignments as secondary (Li and Durbin, 2009). Samtools 0.1.2 was used a second time to sort and index the resulting binary files produced by the alignment process. The sorted and aligned files were run through the Stacks 2.3e *ref_map* pipeline using the default parameters to call variants. The vcf output

was filtered using VCFtools 0.1.16 to remove SNPs and individuals with more than 70% missing data. The aligned and filtered dataset contained 40 individuals and 249,114 SNPs.

2.2.8 Genome scans

All genome scans were performed using the aligned and filtered dataset using 249,114 SNPs. VCFtools 0.1.16 was used to calculate F_{ST} values in a sliding window across the genome between potentially genetically distinct populations using a 20 kb window with a 20 kb step. Other studies looking for differentiated regions have used windows from 10-100 kb, however, finding a suitable window size is often dependent on SNP density and linkage disequilibrium decay (Carlson et al., 2005; Lou et al., 2021; Lowry et al., 2017; Lundberg et al., 2017; Ravinet et al., 2018). East and west groups were compared as well as resident and migratory populations. The same program was also used to calculate π (nucleotide diversity) and Tajima's D in each population of interest (east, west, resident, migratory). A negative Tajima's D suggests population expansion or a selective sweep while a positive Tajima's D can suggest balancing selection or population contraction. Π was calculated using the same window/step size as F_{ST} calculations, while Tajima's D was calculated in bins of 20 kb. All Manhattan plots were visualized in Microsoft Excel. To determine areas of the genome potentially under selection or significantly different between populations, the 99th percentile of windowed F_{ST} values were calculated in R Studio 2021.09.0. This value was selected after examining the Manhattan plot of F_{ST} values to see where there may be significantly different regions, typically seen as peaks in the Manhattan plots. Additionally, the 99th percentile of highest and lowest π values were calculated for each population of interest

(east, west, resident, migratory). Windows that had high F_{ST} values and high or low π values were selected from the annotated reference genome to look at genes located in those areas.

2.2.9 Comparison of golden-crowned kinglets to previously identified candidate genes in birds

While the genes typically associated with migration appear to differ between species, I searched the reference genome for previously found migration candidate genes. This was done to examine if there was any overlap or similarity in genes found related to migration between the golden-crowned kinglets and other migratory species. I searched the golden-crowned kinglet annotated reference genome for the CLOCK, ADCYAP1, CREB1, NPAS2, DRD4, SERT and ADCY8 genes (Contina et al., 2017; Gu et al., 2020; Lugo Ramos et al., 2016; Mueller et al., 2011; Peterson et al., 2013; Saino et al., 2015). The location of the genes was compared to the F_{ST} value in the same or a nearby region as RADseq does not sequence the entire genome. The F_{ST} value was looked at to determine if any of these previously identified genes are potentially different between resident and migratory populations.

2.3 Results

2.3.1. Data processing and filtering

The RADseq process returned 5,942,783 reads with a mean coverage of 9.8x per individual and a minimum coverage of 3.0x per individual. Approximately 95.1% of paired-end reads were successfully aligned to a stack in Stacks 2.3e. The minimum Phred quality scores from sequencing were >30. Prior to filtering, Stacks 2.3 *de novo* called a

total of 1,170,301 SNPs. Post filtering, 217,093 SNPs and 40 individuals were retained for downstream analyses. With the *-write-single-snp* command enabled, 26 individuals and 18,904 SNPs were retained.

2.3.2. Population analyses

The mean observed and expected heterozygosities were similar when compared within each population (Table 2.2). No observed and expected values were greater than 0.06 apart within a single population. The Haida Gwaii and Vancouver Island populations had the lowest observed heterozygosities of all sampled populations. The remainder of the populations had observed heterozygosities ranging from 0.285 to 0.508.

F_{ST} values ranged from 0.00001 to 0.101 (Table 2.3). The highest F_{ST} values (>0.03) were observed between eastern and western populations. The highest F_{ST} value overall was between NL and NWWA (0.101). F_{ST} values were lowest within western populations, with the highest value of 0.048 (VI and NWWA) and lowest of 0.00001 (AK and WA). Between the two eastern populations (Newfoundland and Ontario) the F_{ST} value was 0.00131.

The PCoA using the full dataset with 40 individuals showed the greatest level of clustering along the first PC (3.97%), separating the east and west populations (Figure 2.2). Migratory and resident populations did not show any clear separation along any PC. The second PC (3.07%) clustered VI and HG together but other populations were mixed in with that cluster. The PCoA using the dataset retaining only 1 SNP per locus showed three clusters (Figure 2.2). The first PC (7.07%) explained the strongest clustering, again separating the east and west populations. The second PC (4.41%) showed weak but clear

clustering that divided the western populations into island (HG and VI) and mainland populations (central Alberta and Washington).

The DAPC using the primary dataset showed the same clustering as the PCoA with the same individuals and SNPs, clearly separating east and west populations along the first axis (60.7%) (Figure 2.3). The second axis (27.9%) separated the east and west populations further; however, it did not show any clustering within either group. The DAPC performed with the 1 SNP per locus dataset showed the greatest amount of clustering along the first axis (86.7%), again separating the east and west populations. The second axis (12.1%) separated the east and west groups in addition to separating the mainland (CAB and WA) and island (VI and HG) populations except for one individual from each VI and HG that clustered with the mainland populations.

The ancestry matrix generated with all 40 individuals showed best clustering at $K=2$, with the east and west populations showing strong separation (Figure 2.4). When $K=3$, the ancestry matrix split the east population with no biologically meaningful pattern. The same was true for $K=4$ and $K=5$. The cross-entropy was lowest at $K=2$. With the smaller dataset using 1 SNP per locus, the same patterns were observed, again dividing populations into east and west groups.

2.3.3 Isolation-by-distance

The Mantel test for matrix correspondence indicated that geographic distance and genetic distance showed a weak but positive correlation ($P=0.010$, $R^2=0.1065$). No individuals were obvious outliers along the slope of the trendline.

2.3.4 Divergence time estimation

The BEAST2 software estimated that the eastern and western populations diverged approximately 9,000 years ago. This estimation was consistent between simulation runs.

2.3.5 Reference genome alignment

Approximately 78.9% of reads were successfully mapped to the reference genome and retained for SNP calling by Stacks 2.3e. The remaining reads were removed from analyses due to insufficient mapping qualities or failure to be mapped to the reference genome. Prior to filtering, Stacks 2.3e called 2,033,645 SNPs. After filtering 249,114 SNPs were kept for genome scans and gene ontology analyses.

2.3.6 Genome scans

The F_{ST} genome scan comparing eastern and western populations resulted in a mean F_{ST} across the genome of 0.218. There were 38 windows in the 99th percentile of F_{ST} values (>0.709) containing a total of 81 SNPs (Figure 2.5). Of these 38 windows, four were also associated with the 99th percentile of lowest π values in the east population (on scaffolds VWZN01009326.1, VWZN01000761.1, VWZN01011168.1 and VWZN01019275.1, Table 2.4). A fifth scaffold (VWZN01002418.1, two consecutive windows) contained four SNPs in the 99th percentile of F_{ST} values (Table 2.4). In the western population, these windows were not within the highest or lowest π values. The first window contained a SNP unique to the eastern populations and a Tajima's D value of 1.57. The second window had a Tajima's D value of -0.15 in the eastern population and -0.46 in the western population. The third window had a Tajima's D value of 1.19 in

the eastern population and -0.83 in the western population. The fourth window of interest had a Tajima's D value of -1.05 in the eastern population and -0.53 in the western population. The fifth window containing four SNPs had a Tajima's D value of 1.35 in the eastern population and -0.53 in the western population. Overall, Tajima's D in the western population was more negative.

The F_{ST} genome scan comparing migratory and resident populations resulted in a mean F_{ST} across the genome of 0.113. There were 24 windows in the 99th percentile of F_{ST} values (>0.598) containing 66 SNPs (Figure 2.6). Of these 24 windows, four were also associated with the lowest π values in the migratory populations, and one was associated with the lowest π value in the resident populations (Table 2.4). Overall, the migratory populations had a notably lower π value across the entire genome. The four windows associated with migratory populations were on scaffolds VWZN01007041.1, VWZN01009801.1, VWZN01018308.1 and VWZN01028971.1. The window associated with the lowest π value in the resident populations was on scaffold VWZN01019086.1. The first window had a Tajima's D value of 1.49 in the migratory populations and -0.85 in the resident populations. The next window had a Tajima's D value of 0.54 in the migratory populations and 1.39 in the resident populations. The third window had a Tajima's D value of 0.54 in the migratory populations and -1.35 in the resident populations. The fourth window contained a SNP unique to the migratory populations. The last window had a Tajima's D value of 1.06 in the migratory populations and -0.22 in the resident populations. Overall, the migratory populations had a higher, more positive Tajima's D.

2.3.7 Gene ontology

When examining east and west populations, one of the windows encompassed multiple genes and sequences typically found on the Z chromosome in birds. Three windows contained either a pseudogene or did not have any genes found in the annotated reference genome within or near the region (Table 2.4). Another window was associated with the PALM2 gene. The PALM2 gene codes for a paralemmin protein which is often associated with regulation of cellular shape and plasma membrane dynamics (Hultqvist et al., 2012). In birds, the PALM2 gene is also often located on the Z chromosome (NCBI, 2021).

Within the resident and migratory populations, outlier windows were associated with the CYP19A1, XYLT1, P4HA2, ARRDC4 and FAM169B genes. CYP19A1 encodes for an aromatase which regulates sexual differentiation (Jin et al., 2020). XYLT1 is critical for bone development, P4HA2 is involved in collagen synthesis, ARRDC4 codes for a protein associated with regulation of metabolism, and FAM169B is a pseudogene in humans linked to exercise capacity and deposition of adipose tissue (Du et al., 2015; Sevilla et al., 2020; Patwari et al., 2011; Yuan et al., 2021).

2.3.8 Comparison of golden-crowned kinglets to previously identified candidate genes in birds

The highest F_{ST} value observed within a candidate gene region was 0.228 (DRD4) and the lowest was 0.071 (CREB1). The CLOCK region was not sequenced, but nearby F_{ST} values on the same scaffold were approximately 0.126. The remaining gene regions

(ADCYAP1, NPAS2, SERT, ADCY8) had F_{ST} values in the gene region or nearby region ranging from 0.091 to 0.149.

2.4 Discussion

2.4.1 Population structure and barriers to gene flow

Overall, my findings reflect similar patterns to previous research using different molecular markers in that the primary population differentiation in golden-crowned kinglets is the distinct split of eastern and western populations (Burg et al., 2014; Graham et al., 2020). My PCoA, DAPC and LEA analyses showed these two populations as distinct genetic groupings and F_{ST} values were highest between east and west populations. The most likely explanation for this east-west pattern is the Pleistocene glaciation causing them to reside in separate glacial refugia. BEAST2 estimated that the east and west populations diverged approximately 9,000 years ago. During this time, the LGM had ended, and the majority of the ice sheets would have retreated northward significantly (Dyke, 2004). While the ice sheets would no longer be separating the east and west populations, there was a large expanse of unsuitable habitat left behind by the glaciers (Dyke, 2005). The boreal forest likely did not fully connect across Canada until approximately 7,000 years ago, meaning the eastern and western populations would be unlikely to mix due to the fragmented forest and the kinglet's preference for mature forest (Dyke, 2005). While there was an ice-free corridor between the Laurentide and Cordilleran ice sheets during the glaciation, herbaceous plants likely occupied this area and trees were not present in this corridor for most of the glacial period (Mott and Jackson, 1982). Spruce and pine trees had colonized this area by approximately 8,200 years before present, however, herbaceous plants began to increase, causing the

coniferous habitat to decrease (Mott and Jackson, 1982). If golden-crowned kinglets were able to colonize and remain in this coniferous habitat as they expanded from their western glacial refugia, it would explain why the Rocky Mountains do not appear to inhibit gene flow. The Pleistocene glaciation is also attributed to many other species' population structure, making it a significant factor in the genetic patterns seen today (Burg et al., 2006; Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004).

Graham et al., 2020 suggested that isolation by distance could also explain the distinct east-west split. Distance could maintain the separation by preventing the two groups from exchanging enough gene flow to become one homogenous group, however, we do not have enough evidence to suggest it is the main barrier to gene flow between the two populations and the cause of the observed population structure. The isolation-by-distance test indicated that while geographic distance is a significant contributor to genetic distance, it is only weakly correlated. If distance were the main barrier to gene flow, it may also be expected that Alaska and the Oregon/Washington individuals would also show significant differentiation. It is most likely that the Pleistocene glaciation and the associated habitat changes are the biggest factor in the development of the two genetic groups.

In contrast to the previous two studies, I did not find that populations located on Haida Gwaii clustered together distinctly from other populations. Instead, I found that when accounting for linkage disequilibrium and redundancy, both Haida Gwaii and Vancouver Island populations separated from the mainland populations in the west. The PCoA and DAPC using 1 SNP per locus showed the two island populations separating

from the mainland populations (Figure 2.2; Figure 2.3). These island populations also had the highest number of SNPs that were out of HWE, at least 400 SNPs more than any mainland population. Additionally, both island populations had the lowest observed heterozygosities (Table 2.2). All of the aforementioned results could suggest that golden-crowned kinglet's gene flow or dispersal is limited by bodies of water, although these populations did not separate in the ancestry matrices. Another possibility is that because Vancouver Island and Haida Gwaii were some of the first areas in western Canada to be relatively ice free, this allowed the golden-crowned kinglets to colonize these islands sooner than most of mainland North America (Clague and Ward, 2011). It has also been suggested that Haida Gwaii was home to a glacial refugium inhabited by forest-dependent species (Pruett et al., 2013). This glacial refugium may have been connected to Vancouver Island and coastal British Columbia due to the lowered sea levels at the time, causing the two island populations to form a genetic group (Pruett et al., 2013; Byun et al., 1997). The low heterozygosity combined with the increase in SNPs out of HWE also suggest that there is decreased genetic diversity and few new individuals contributing to the gene pool (Chen et al., 2017; Boessenkool et al., 2007). The previous studies that detected a potential split between Haida Gwaii and mainland populations used mitochondrial DNA and microsatellites, both of which often have a higher mutation rate than SNPs from next-generation sequencing (Anmarkrud et al., 2008; Burg et al., 2014; Graham et al., 2020; Xu et al., 2012). If the split is very recent, it is possible the genomic variation has not yet mutated enough for it to be detectable using RADseq. This is unlikely, however, as studies comparing microsatellites and SNPs typically show that while the two molecular markers show the same genetic structure, RADseq using large numbers of SNPs often uncovers more fine scale population structure than microsatellites

(Hodel et al., 2017; Jefferies et al., 2016; Sunde et al., 2020; Vendrami et al., 2017; Zimmerman et al., 2020). The differences in results using 1 SNP per locus compared to all SNPs demonstrate the importance of linked SNPs in a large dataset. High levels of linkage disequilibrium in SNPs can affect genomic models and many SNPs will be redundant (Calus and Vandenplas, 2018). In this case, the redundancy in SNPs obscures finer scale population structure. Additionally, while the full dataset had drastically more SNPs, the SNPs genotyped were not the same in all individuals which could also affect the ability to detect population structure. Analyses were run with both datasets to increase my ability to find population structure and maximize the sequencing cost, as almost half of the individuals were filtered out when using 1 SNP per locus. Many individuals in this study had poor sequencing coverage and were removed from analyses after filtering. The poor sequencing coverage may be caused by poor initial sample quality as many samples were dried blood on filter paper. Some filter papers are treated or bleached which can degrade DNA. Slightly degraded DNA is often suitable for microsatellites or mtDNA studies, however, next-generation sequencing requires high quality whole genomic DNA. Using SNP markers in this case has revealed slightly different genetic patterns than previous studies.

While east and west populations are consistently different across studies, all present studies have a large sample gap through the middle of Canada. Samples from Manitoba and Saskatchewan are required to fill this gap and determine where the east-west groups split. It is possible that the central groups make up a third distinct genetic grouping, however, it is also possible that they may also make up intermediates between the east and west populations. It will be critical to analyze samples from these areas in the

future to fully understand the genetic patterns we see in golden-crowned kinglets in North America.

2.4.2 Genome scans in east and west populations

The highly differentiated regions between east and west populations were mostly genes or sequences commonly found on the Z chromosome. Birds have a ZW chromosome system where females are the heterogametic sex (ZW) (Battey, 2019). Sex chromosomes contribute disproportionately to speciation and population differentiation and contain many genes that are related to speciation and reproductive isolation (Battey, 2019; Charlesworth et al., 1987; Dhimi et al., 2015; Irwin, 2018; Oyler-McCance et al., 2015). This is partially because mutations are expected to accumulate more readily on the Z chromosome than on autosomes, causing them to diverge faster, which would help explain the relatively recent split in the east and west populations (Irwin, 2018; Oyler-McCance et al., 2015). Additionally, recessive mutations are more likely to become fixed on the Z chromosome because one-third of the mutations will be hemizygous (Dahmi et al., 2015). The differentiation on the Z chromosome between east and west populations is not surprising given the properties of the Z chromosome and the role sex chromosomes play in population differentiation.

2.4.3 Differences in migratory and resident populations

Migratory and resident populations did not show significant differentiation from one another; however, I did find several regions along the genome that were consistently different between the two migratory phenotypes. These regions have high F_{ST} values and low π values, suggesting areas of selection in the genome (Ravinet et al., 2018). Overall,

the migratory populations had higher, more positive Tajima's D values which can suggest balancing selection or a lack of rare alleles (Carlson et al., 2005). Balancing selection suggests that certain alleles are maintained in a population. Specifically, the FAM196B gene and the ARRDC4 gene may be of interest in relation to migratory behaviour because their functions lend themselves to important parts of the migration process. The FAM196B gene has been linked to exercise capacity and deposition of adipose tissue (Du et al., 2015; Sevilla et al., 2020; Patwari et al., 2011; Yuan et al., 2021). During migration birds are required to exert more energy than usual, requiring an increase in their capacity for exercise. Additionally, prior to migration many birds show an increase in fat stores, presumably to help with energy needs during their travels (Dingle and Drake, 2007; Pennycuick, 1998). ARRDC4 is associated with regulation of metabolism and during migration birds' metabolic needs and processes change by increasing food consumption, which, in turn, allows them to put on more muscle and fat as well as expend more energy during long flights (McCabe and Guglielmo, 2019). A third gene, P4HA2, is involved in collagen synthesis. Collagen is critical in strengthening muscle and connective tissues. In many birds, muscle mass increases prior to migration in preparation for long flights (Bauchinger and Biebacj, 2001; Pennycuick, 1998; Swanson and Merkord, 2013). This suggests that differences in this collagen synthesizing gene between migratory phenotypes could be related to migratory physiology. The CYP19A1 gene is involved with aromatase and may also influence the migratory habits as aromatase expression has been found to differ between migratory and resident individuals in other species (Krause, et al., 2022). The aforementioned genes may be candidate genes for golden-crowned kinglet's migration.

When searching for previously identified candidate genes for migration, I did not find that any of these genes were in areas of notably elevated F_{ST} . Even the gene with the highest associated F_{ST} (0.228, *DRD4*) was not involved with any F_{ST} peaks, as the vast majority of SNPs had F_{ST} values below 0.400. Migratory candidate genes do not appear to be universal in bird species, and that may be the case with golden-crowned kinglets as well (Delmore et al., 2020; Delmore et al., 2016; Lundberg et al., 2017; Parody-Merino et al., 2019; Peterson et al., 2013). Another likely explanation is that RADseq may not provide adequate breadth of coverage to detect all candidate genes. RADseq sequences a small portion of the genome causing it to miss signatures of selection or divergence (Lou et al., 2021; Lowry et al., 2016). Whole-genome sequencing, even with low coverage, would ensure that all significant regions were sequenced. Additionally, genome scans using whole-genome sequencing data show much more pronounced and distinct peaks compared to RADseq (Lou et al., 2021; Lowry et al., 2016). In the future, whole-genome sequencing should be utilized to confirm my findings and re-assess both the new candidate genes found as well as candidate genes found in other studies.

Another interesting pattern that I found was that the migratory populations had a notably lower nucleotide diversity throughout the entire genome in comparison to the resident populations (Figure 2.6). Selective sweeps and local adaptation can cause both high and low nucleotide diversity depending on the strength of selection (Jasper and Yeaman, 2020). In the case of selection, differences in nucleotide diversity between populations are typically restricted to specific regions of the genome rather than the entire genome. Low genetic diversity is typically thought to put populations at an increased conservation risk due to lack of genetic variation in the gene pool, influencing how the

population responds to environmental change (Evans and Sheldon, 2008; Harvey et al., 2017). This can be caused by limited gene flow into the population or a sudden population contraction. The migratory populations from this study are from the western edge of Alberta near the Rocky Mountains. It is possible that a population of golden-crowned kinglets was able to colonize this area east of the Rocky Mountains during the Pleistocene forestation. Another possibility is that these migratory individuals are interbreeding with each other more than with the neighbouring resident populations, limiting the number of individuals contributing to the gene pool. This is because resident individuals can establish mate pairs earlier than migratory individuals as they do not have to make the long journey back to their breeding grounds. Additionally, it is often suggested that the farther a population is from the glacial refugia it expanded from, the more reduced the genetic diversity (Fry and Zink, 2002; Hewitt, 1996). If the western populations resided in a coastal glacial refugium, then the Alberta populations may be the farthest away of the sample sites included in this study, leading to reduced nucleotide diversity. The issue with any of these hypotheses, however, is that there remains a large sample gap in the central migratory range that needs to be included in future studies as it is not clear if this pattern is consistent across migratory populations.

The large sample gap is likely limiting my findings and it is critical to fill this gap to understand the migratory genetics of golden-crowned kinglets. We classified individuals from Alberta as migratory, as did the previous study (Graham et al., 2020). An issue with this classification is that range maps from different sources often show that the resident range covers the majority of British Columbia across the Rocky Mountains into southwestern Alberta. The citizen science data accessible through eBird shows small

numbers of golden-crowned kinglets remaining in areas often classified as migratory in December when migratory individuals would have traveled south (eBird, 2021). Even when depicted as migratory, the populations classified as migratory in my study are geographically close to resident populations. This could lead to the presence of mixed populations of migratory and resident birds with few ways aside from satellite tracking to determine the bird's migratory phenotype. All sampling occurred during the breeding season, meaning if there are mixed populations it is possible to have sampled resident birds. Filling in the central sample gap with samples from Saskatchewan and Manitoba will ensure that purely migratory birds are sampled and should be the priority of future research.

2.5 Conclusions

By using genome wide markers, this research has confirmed previous findings on golden-crowned kinglets while also identifying regions of the genome that are significantly different between populations and the genes associated with these regions. In addition to the east west split, this study suggests a third genetic group resides on two islands off the west coast of British Columbia. Several potential candidate genes for migratory behaviour in golden-crowned kinglets were identified in addition to highlighting the Z chromosome as an area of interest for their population divergence between east and west populations. This further enforces that the genetic component of migratory behaviour is different between species and should be studied to understand the multiple ways it can arise. It is critical to evaluate a variety of species as it can be challenging to predict how different taxa respond to processes affecting gene flow. In

identifying genetically unique populations, researchers can use this information to shape conservation efforts by treating them as separate management units.

Table 2.1 Sample locations and classification for this study based on genetic cluster and migratory phenotype.

Population ID	Sample size (primary dataset)	Sample size using 1 SNP per locus	Genetic cluster	Migratory phenotype
OR (Oregon)	1		West	Resident
WA (Washington)	3	3	West	Resident
NWWA (North-west Washington)	2		West	Resident
VI (Vancouver Island)	6	6	West	Resident
HG (Haida Gwaii)	6	6	West	Resident
AK (Alaska)	2		West	Resident
BC (South-east British Columbia)	4		West	Migratory
CAB (Central Alberta)	4	4	West	Migratory
SAB (South Alberta)	4		West	Migratory
ON (Ontario)	4	5	East	Resident
NB (New Brunswick)	1		East	Resident
NL (Newfoundland)	3	3	East	Resident

Table 2.2 Mean expected and observed heterozygosities across all study populations with sample size >1.

Population	Sample size	H_o	H_e
WA	3	0.345	0.397
NWWA	2	0.508	0.528
VI	6	0.220	0.247
HG	6	0.221	0.247
AK	2	0.465	0.532
BC	4	0.285	0.322
CAB	4	0.348	0.398
SAB	4	0.351	0.397
ON	4	0.296	0.327
NL	3	0.359	0.396

Table 2.3 Heatmap of pairwise F_{ST} values between all study populations with sample size >1 . Red indicates higher values and green indicate lower values. Bold F_{ST} values had significant p -values after Benjamini-Hochberg corrections for multiple comparisons (Benjamini and Hochberg, 1995). P -values are provided above the diagonal and are bold if significant.

	WA	NWWA	VI	HG	AK	BC	CAB	SAB	ON	NL
WA		0.0556	0.0667	0.0689	0.0911	0.0578	0.0400	0.0711	0.0133	0.0311
NWWA	0.0006		0.0200	0.0644	0.0733	0.0222	0.0422	0.0444	0.0244	0.0333
VI	0.0001	0.0448		0.0756	0.0556	0.0600	0.0533	0.0622	0.0022	0.0089
HG	0.0003	0.0152	0.0061		0.0889	0.0933	0.0867	0.0511	0.0044	0.0067
AK	0.0000	0.0004	0.0012	0.0021		0.0500	0.0822	0.0844	0.0467	0.0489
BC	0.0001	0.0407	0.0053	0.0008	0.0001		0.0478	0.0800	0.0178	0.0111
CAB	0.0143	0.0136	0.0002	0.0003	0.0002	0.0071		0.0778	0.0378	0.0267
SAB	0.0105	0.0161	0.0048	0.0021	0.0037	0.0034	0.0000		0.0156	0.0289
ON	0.0690	0.0816	0.0885	0.0951	0.0335	0.0877	0.0394	0.0565		0.0489
NL	0.0747	0.1010	0.0994	0.0924	0.0371	0.0861	0.0544	0.0801	0.0013	

Table 2.4 Scaffolds and genes of interest from genome scans between both the east vs. west comparison and the migratory vs. resident comparison.

East vs. West		
Scaffold	Gene	Function
VWZN01009326.1	Spans multiple diverse genes	Multiple genes on the Z-chromosome
VWZN01000761.1	Pseudogene	n/a
VWZN01011168.1	No genes found	n/a
VWZN01019275.1	PALM2	Regulation of cellular shape and plasma membrane dynamics
VWZN01002418.1	No genes found	n/a
Migratory vs. Resident		
Scaffold	Gene	Function
VWZN01007041.1	CYP19A1	Sexual development and behaviour
VWZN01009801.1	XYLT1	Bone development
VWZN01018308.1	P4HA2	Collagen synthesis
VWZN01028971.1	No genes found	n/a
VWZN01019086.1	ARRDC4	Glucose metabolism
	FAM169B	Exercise capacity and deposition of adipose tissue

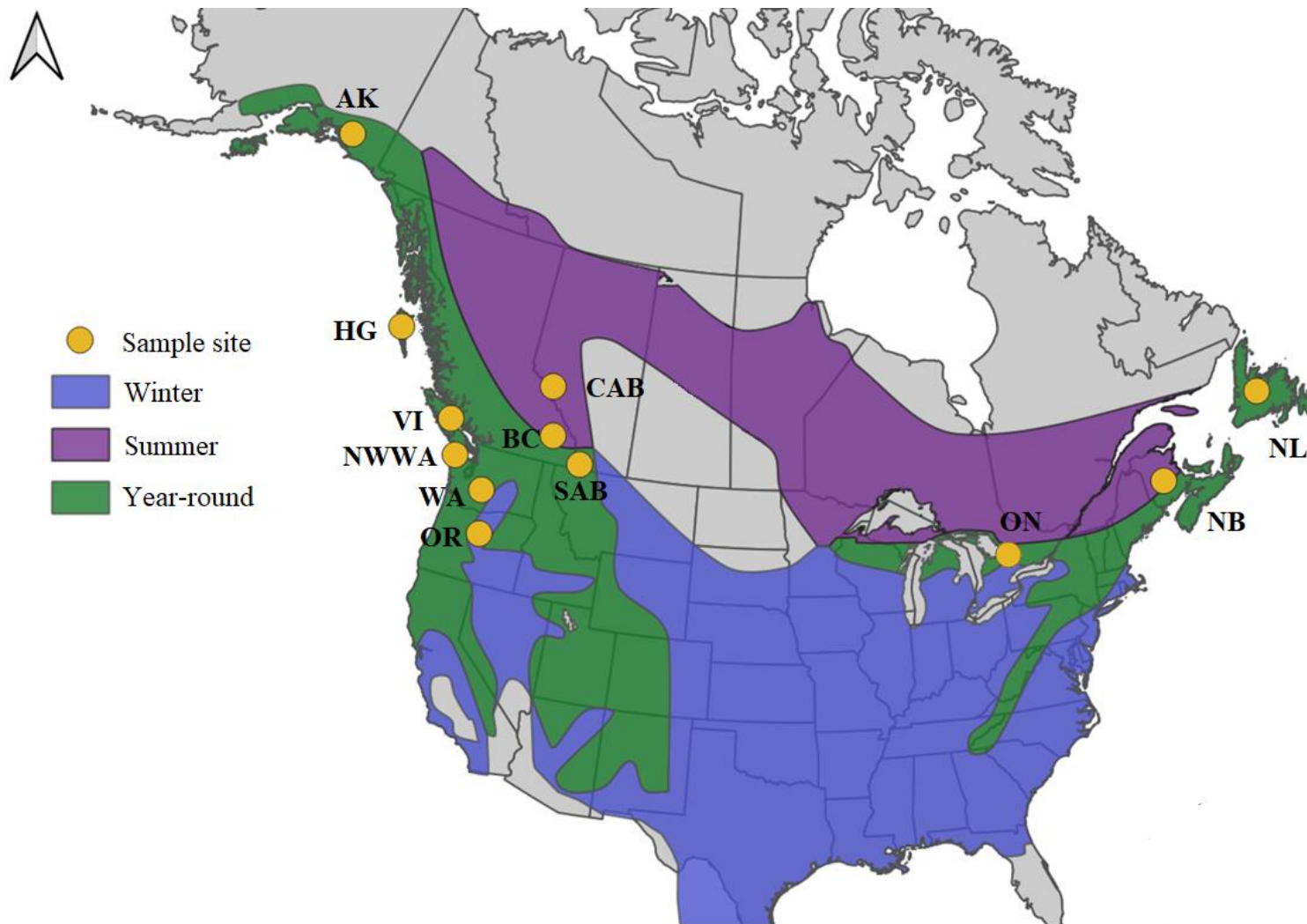


Figure 2.1 Range map of golden-crowned kinglets in North America and sample locations for this study. Range data obtained from BirdLife International (2016). Map created in QGIS 3.2.

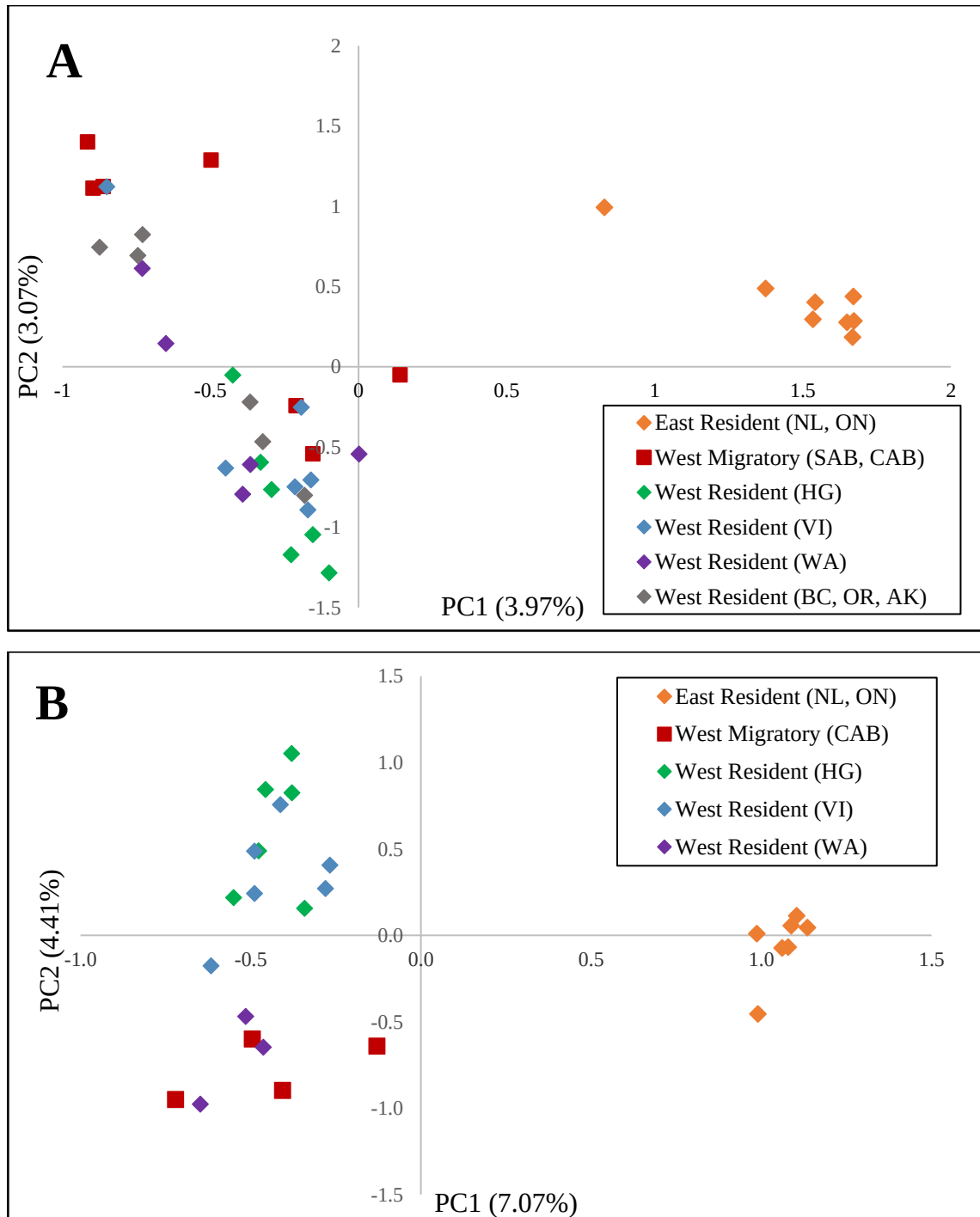


Figure 2.2 PCoA of study populations. A) PCoA using the primary dataset with 40 individuals from 12 populations with 217,093 SNPs. B) PCoA using 1 SNP per locus, retaining 26 individuals from 6 populations with 18,904 SNPs.

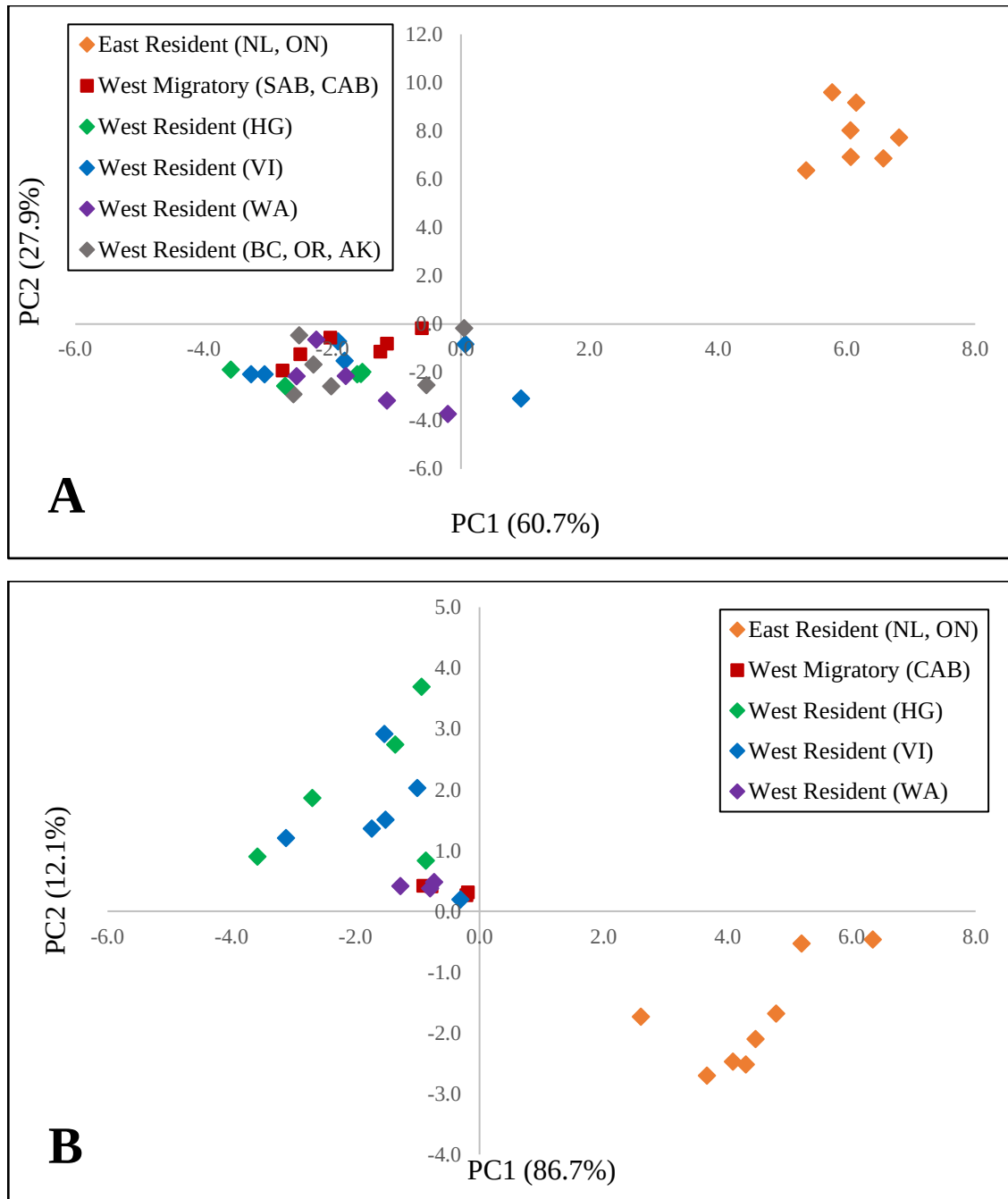


Figure 2.3 DAPC of study populations. A) DAPC using the primary dataset with 40 individuals from 12 populations with 217,093 SNPs. B) DAPC using 1 SNP per locus, retaining 26 individuals from 6 populations with 18,904 SNPs.

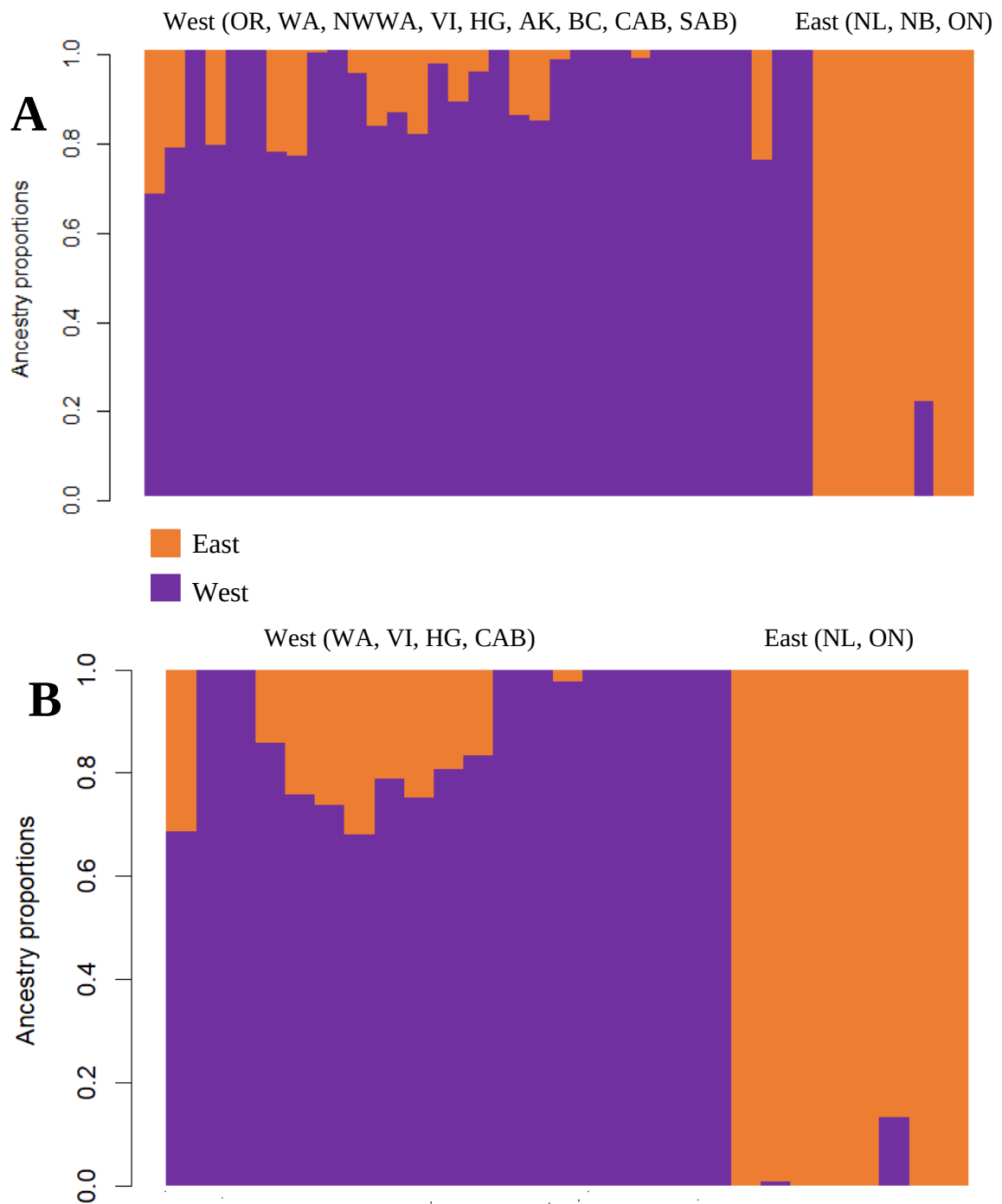


Figure 2.4 Ancestry matrices of study populations. A) LEA plot (K=2) based on all 40 individuals from 12 populations using 217,093 SNPs. B) LEA plot (K=2) using 1 SNP per locus with 26 individuals from 6 populations with 18,904 SNPs.

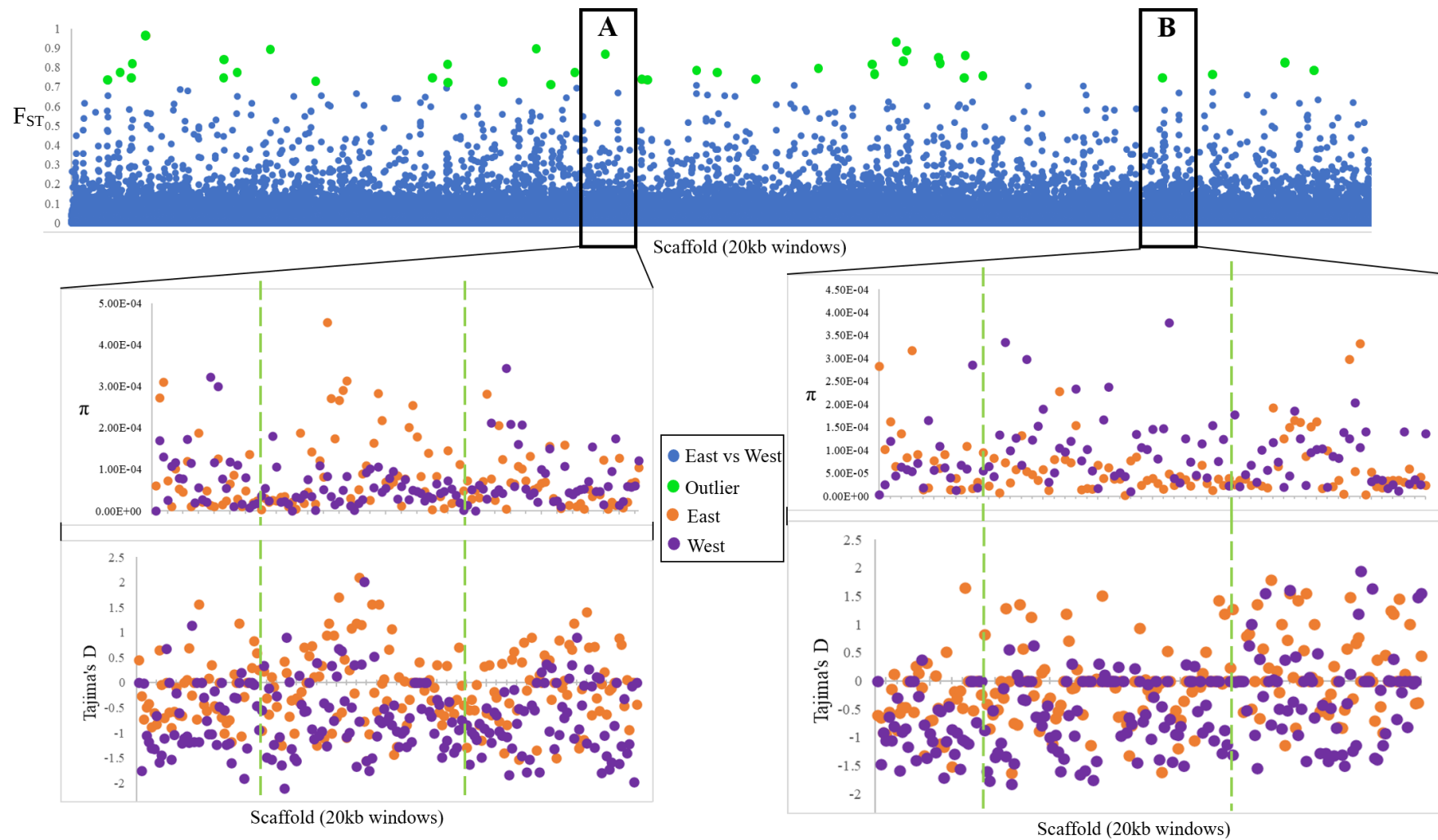


Figure 2.5 Genome scans of F_{ST} from the east and west comparison. Green in the F_{ST} plot indicate the highest SNP windows (outliers). π and Tajima's D are included for the two regions that contain genes of interest. A) Scaffold VWZN01009326.1, associated with many diverse genes on the Z chromosome. B) Scaffold VWZN01019275.1, associated with the PALM2 gene.

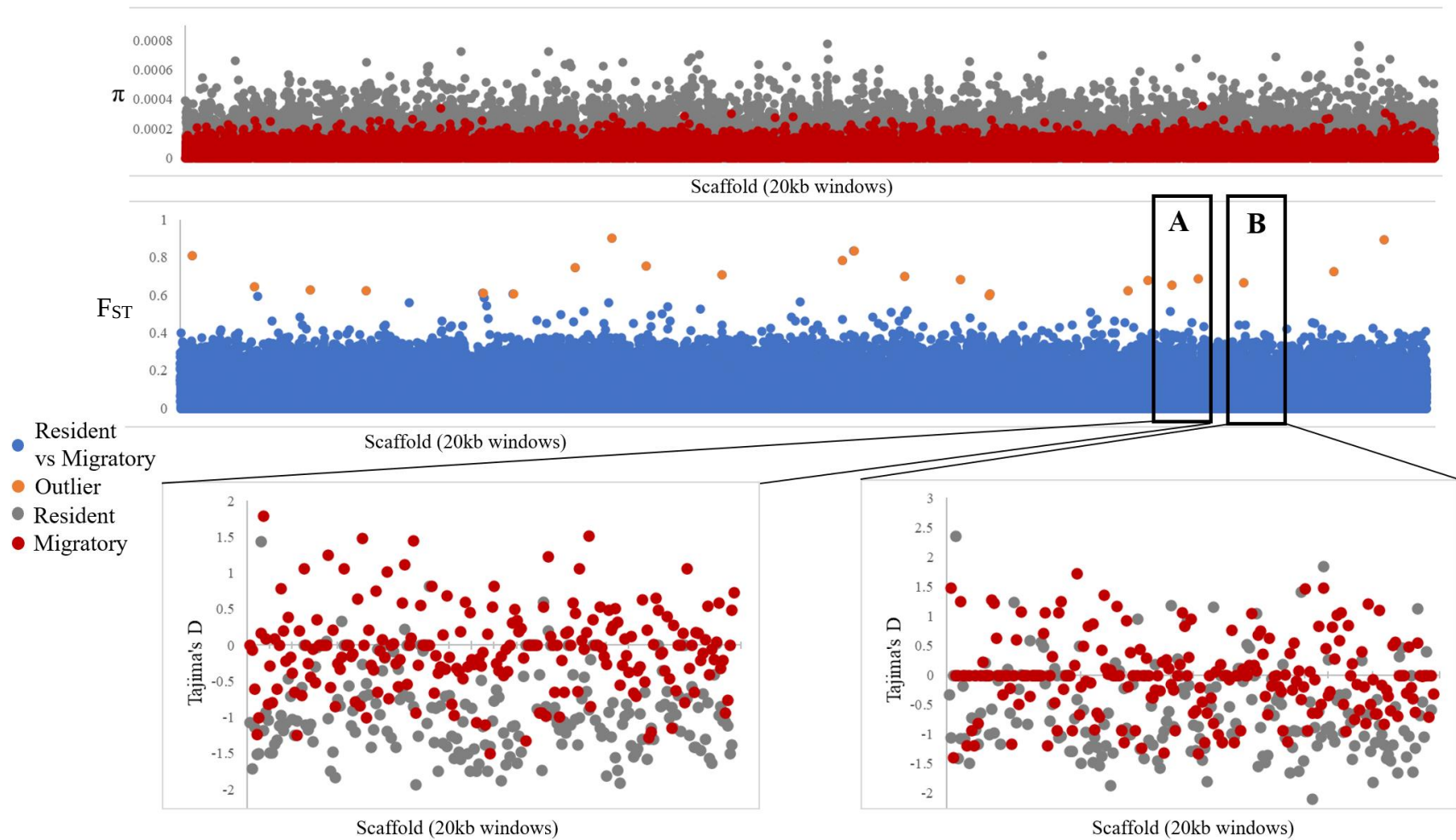


Figure 2.6 Genome scans from the migratory and resident comparisons showing nucleotide diversity (π) and F_{ST} across the genome. Tajima's D is included for two regions containing genes of interest. A) Scaffold VWZN01018308.1, associated with the P4HA2 gene. B) Scaffold VWZN01019086.1, associated with the ARRDC4 and FAM169B genes.

2.6 Literature Cited

- Adams, R. V., and Burg, T. M. (2015). Gene flow of a forest-dependent bird across a fragmented landscape. *PLoS One*, 10, 1-22.
- Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Anmarkrud, J. A., Kleven, O., Bachmann, L., and Lifjeld, J. T. (2008). Microsatellite evolution: Mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. *BMC Evolutionary Biology*, 8, 138.
- Batthey, C. J. (2020). Evidence of linked selection on the Z chromosome of hybridizing hummingbirds. *Evolution*, 74(4), 725–739.
- Bauchinger, U., and Biebach, H. (2001). Differential catabolism of muscle protein in garden warblers (*Sylvia borin*): Flight and leg muscle act as a protein source during long-distance migration. *Journal of Comparative Physiology B*, 171(4), 293–301.
- Birdlife International and Handbook of the Birds of the World. (2016). Bird distribution maps of the world. Version 6.0. Available online at: <https://datazone.birdlife.org/species/requestdis>.
- Boessenkool, S., Taylor, S. S., Tepolt, C. K., Komdeur, J., and Jamieson, I. G. (2007). Large mainland populations of South Island robins retain greater genetic diversity than offshore island refuges. *Conservation Genetics*, 8, 705-714.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., and Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 10(4), e1003537.
- Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., and RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29(8), 1917–1932.
- Burg, T. M., Gaston, A. J., Winker, K., and Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, 15(9), 2409–2419.
- Burg, T. M., Taylor, S., Lemmen, K., Gaston, T., and Friesen, V. (2014). Postglacial population genetic differentiation potentially facilitated by a flexible migratory

- strategy in golden-crowned kinglets (*Regulus satrapa*). *Canadian Journal of Zoology*, 92(2), 163-172.
- Byun, S. A., Koop, B. F., and Reimchen, T. E. (1997). North American black bear mtDNA phylogeography: Implications for morphology and the Haida Gwaii glacial refugium controversy. *Evolution*, 51(5), 1647–1653.
- Calus, M. P. L., and Vandenplas, J. (2018). SNPrune: An efficient algorithm to prune large SNP array and sequence datasets based on high linkage disequilibrium. *Genetics Selection Evolution*, 50(1), 34.
- Carlson, C. S., Thomas, D. J., Eberle, M. A., Swanson, J. E., Livingston, R. J., Rieder, M. J., and Nickerson, D. A. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Research*, 15(11), 1553–1565.
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.
- Charlesworth, B., Coyne, J. A., and Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, 130(1), 113–146.
- Chen, B., Cole, J. W., and Grond-Ginsbach, C. (2017). Departure from Hardy Weinberg Equilibrium and genotyping error. *Frontiers in Genetics*, 8(167), 1-6.
- Clague, J., and Ward, B. (2011). Pleistocene glaciation of British Columbia. *Developments in Quaternary Science*, 15, 563–573.
- Contina, A., Bridge, E. S., Ross, J. D., Shipley, J. R., and Kelly, J. F. (2018). Examination of Clock and Adcyap1 gene variation in a neotropical migratory passerine. *PloS One*, 13(1), e0190859.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008.

- Daniels, S. J., and Walters, J. R. (2000). Inbreeding depression and its effects on natal dispersal in red-cockaded woodpeckers. *The Condor*, 102, 482-491.
- Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., and Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, 9, e54462.
- Delmore, K. E., Toews, D. P. L., Germain, R. R., Owens, G. L., and Irwin, D. E. (2016). The genetics of seasonal migration and plumage color. *Current Biology: CB*, 26(16), 2167–2173.
- Dhami, K. K., Joseph, L., Roshier, D. A., and Peters, J. L. (2016). Recent speciation and elevated Z-chromosome differentiation between sexually monochromatic and dichromatic species of Australian teals. *Journal of Avian Biology*, 47(1), 92–102.
- Dingle, H., and Drake, V. A. (2007). What is migration? *BioScience*, 57(2), 113–121.
- Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., and Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161(3), 1307–1320.
- Dyke, A. S. (2005). Late quaternary vegetation history of northern North America based on pollen, macrofossil, and faunal remains. *Géographie Physique et Quaternaire*, 59(2–3), 211–262.
- Dyke, A. S. (2004). An outline of North American deglaciation with emphasis on central and northern Canada. In J. Ehlers and P. L. Gibbard, *Developments in Quaternary Sciences* (Vol. 2, pp. 373–424). Elsevier.
- eBird. (2021). eBird: An online database of bird distribution and abundance. Cornell Lab of Ornithology. Available online at: <http://ebird.org>
- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567.
- Evans, S. R., and Sheldon, B. C. (2008). Interspecific patterns of genetic diversity in birds: Correlations with extinction risk. *Conservation Biology*, 22(4), 1016–1025.
- Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A. H., Xie, D., Chen, G., Guo, C., Faircloth, B. C., Petersen, B., Wang, Z., Zhou, Q., Diekhans, M., Chen, W., Andreu-Sánchez, S., Margaryan, A., Howard, J. T., Parent, C., ... Zhang, G.

- (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature*, 587(7833), 252–257.
- Frichot, E., and Francois, O. (2015) LEA: and R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925–929.
- Fry, A. J., and Zink, R. M. (1998). Geographic analysis of nucleotide diversity and song sparrow (Aves: Emberizidae) population history. *Molecular Ecology*, 7(10), 1303–1313.
- Graham, B. A., Carpenter, A. M., Friesen, V. L., and Burg, T. M. (2020). A comparison of neutral genetic differentiation and genetic diversity among migratory and resident populations of golden-crowned kinglets (*Regulus satrapa*). *Journal of Ornithology*, 161(2), 509–519.
- Gu, Z., Pan, S., Lin, Z., Hu, L., Dai, X., Chang, J., Xue, Y., Su, H., Long, J., Sun, M., Ganusevich, S., Sokolov, V., Sokolov, A., Pokrovsky, I., Ji, F., Bruford, M. W., Dixon, A., and Zhan, X. (2021). Climate-driven flyway changes and memory-based long-distance migration. *Nature*, 591(7849), 259–264.
- Harvey, M. G., Aleixo, A., Ribas, C. C., and Brumfield, R. T. (2017). Habitat association predicts genetic diversity and population divergence in Amazonian birds. *The American Naturalist*, 190(5), 631–648.
- Helbig, A. J. (1991). Inheritance of migratory direction in a bird species: A cross-breeding experiment with SE- and SW-migrating blackcaps (*Sylvia atricapilla*). *Behavioral Ecology and Sociobiology*, 28(1), 9–12.
- Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58(3), 247–276.
- Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F., and Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54(1), 291–301.
- Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., and Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite

- increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports*, 7(1), 17598.
- Hultqvist, G., Daza, D. O., Larhammar, D., and Kilimann, M. W. (2012). Evolution of the vertebrate paralemmin gene family: Ancient origin of gene duplicates suggests distinct functions. *PLoS One*, 7(7), e41850.
- Irwin, D. E. (2018). Sex chromosomes and speciation in birds and other ZW systems. *Molecular Ecology*, 27(19), 3831–3851.
- Jasper, R. J., and Yeaman, S. (2020). *Local adaptation can cause both peaks and troughs in nucleotide diversity within populations*. bioRxiv.
- Jeffries, D. L., Copp, G. H., Lawson Handley, L., Olsén, K. H., Sayer, C. D., and Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, 25(13), 2997–3018.
- Jin, K., Zuo, Q., Song, J., Zhang, Y., Chen, G., and Li, B. (2020). CYP19A1 (aromatase) dominates female gonadal differentiation in chicken (*Gallus gallus*) embryos sexual differentiation. *Bioscience Reports*, 40(10), BSR20201576.
- Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.
- Kirby, J. S., Stattersfield, A. J., Butchart, S. H. M., Evans, M. I., Grimmett, R. F. A., Jones, V. R., O’Sullivan, J., Tucker, G. M., and Newton, I. (2008). Key conservation issues for migratory land- and waterbird species on the world’s major flyways. *Bird Conservation International*, 18(1), S49–S73.
- Knowles, L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, 10, 691–701.
- Koleček, J., Adamík, P., and Reif, J. (2020). Shifts in migration phenology under climate change: Temperature vs. abundance effects in birds. *Climatic Change*, 159, 177–194.
- Korneliussen, T. S., Moltke, I., Albrechtsen, A., and Nielsen, R. (2013). Calculation of Tajima’s D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics*, 14(1), 289.

- Krause, J. S., Watkins, T., Reid, A. M. A., Cheah, J. C., Perez, J. H., Bishop, V. R., Ramenofsky, M., Wingfield, J. C., and Meddle, S. L. (2022). Gene expression of sex steroid metabolizing enzymes and receptors in the skeletal muscle of migrant and resident subspecies of white-crowned sparrow (*Zonotrichia leucophrys*). *Oecologia*, 199(3), 549-562.
- Lavretsky, P., Dacosta, J. M., Hernández-Baños, B. E., Engilis Jr, A., Sorenson, M. D., and Peters, J. L. (2015). Speciation genomics and a role for the Z chromosome in the early stages of divergence between Mexican ducks and mallards. *Molecular Ecology*, 24(21), 5364–5378.
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589-595.
- Liedvogel, M., Åkesson, S., and Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology and Evolution*, 26(11), 561–569.
- Lou, R. N., Jacobs, A., Wilder, A. P., and Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, 30(23), 5966–5993.
- Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., and Storfer, A. (2017). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, 17(2), 142–152.
- Lugo Ramos, J. S., Delmore, K. E., and Liedvogel, M. (2017). Candidate genes for migration do not distinguish migratory and non-migratory birds. *Journal of Comparative Physiology A*, 203(6), 383–397.
- Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., and Bensch, S. (2017). Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. *Evolution Letters*, 1(3), 155–168.
- MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, S. E., and Nielsen, S. E. (2020). Gene flow and climate-associated genetic variation in a vagile habitat specialist. *Molecular Ecology*, 29(20), 3889-3906.

- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10–12.
- Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16(3), 1215.
- Mott, R. J., and Jackson Jr., L. E. (1982). An 18,000 year palynological record from the southern Alberta segment of the classical Wisconsinan “Ice-free Corridor.” *Canadian Journal of Earth Sciences*, 19(3), 504–513.
- Mueller, J. C., Pulido, F., and Kempnaers, B. (2011). Identification of a gene associated with avian migratory behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 278(1719), 2848–2856.
- National Center for Biotechnology Information (NCBI). (2021). PALM2AKAP2 PALM2 and AKAP2 fusion (*Taeniopygia guttata* (zebra finch)).
- Nazareno, A. G., Bemmels, J. B., Dick, C. W., and Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*, 17(6), 1136–1147.
- Oyler-McCance, S. J., Cornman, R. S., Jones, K. L., and Fike, J. A. (2015). Z chromosome divergence, polymorphism and relative effective population size in a genus of lekking birds. *Heredity*, 115(5), 452–459.
- Parody-Merino, A. M., Battley, P. F., Conklin, J. R., and Fidler, J. R. (2019). No evidence for an association between Clock gene allelic variation and migration timing in a long-distance migratory shorebird (*Limosa lapponica baueri*). *Oecologia*, 191(4), 843–859.
- Patwari, P., Emilsson, V., Schadt, E. E., Chutkow, W. A., Lee, S., Marsili, A., Zhang, Y., Dobrin, R., Cohen, D. E., Larsen, P. R., Zavacki, A. M., Fong, L. G., Young, S. G., and Lee, R. T. (2011). The arrestin domain-containing 3 protein regulates body mass and energy expenditure. *Cell Metabolism*, 14(5), 671–683.
- Peakall, R., and Smouse, P. E. (2006). Genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295.
- Peltier, W. (1994). Ice Age Paleotopography. *Science*, 265, 195–201.

- Pennycuik, C. J. (1998). Computer simulation of fat and muscle burn in long-distance bird migration. *Journal of Theoretical Biology*, 191(1), 47–61.
- Perrin, A., Khimoun, A., Faivre, B., Ollivier, A., Pracontal, N., Théron, F., Loubon, M., Leblond, G., Duron, O., and Garnier, S. (2020). Habitat fragmentation differentially shapes neutral and immune gene variation in a tropical bird species. *Heredity*, 126, 148-162.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135.
- Peterson, M. P., Abolins-Abols, M., Atwell, J. W., Rice, R. J., Milá, B., and Ketterson, E. D. (2013). Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *F1000Research*, 2, 115.
- Pielou, E. C. (1992). *After the Ice Age: The Return of Life to Glaciated North America* (1st edition). University of Chicago Press.
- Pruett, C. L., Topp, C. M., Maley, J. M., McCracken, K. G., Rohwer, S., Birks, S., Sealy, S. G., and Winker, K. (2013). Evidence from the genetics of landbirds for a forested Pleistocene glacial refugium in the Haida Gwaii area. *The Condor*, 115(4), 7255–737.
- Pulido, F. (2007). The genetics and evolution of avian migration. *BioScience*, 57(2), 165–174.
- QGIS Development Team (2022). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.
- Ravinet, M., Elgvin, T. O., Trier, C., Aliabadian, M., Gavrillov, A., and Sætre, G.-P. (2018). Signatures of human-commensalism in the house sparrow genome. *Proceedings of the Royal Society B: Biological Sciences*, 285(1884), 20181246.
- Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., DeSante, D. F., Moore, F., and Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23(23), 5726–5739.

- Rushing, C. S., Royle, J. A., Ziolkowski, D. J., and Pardieck, K. L. (2020). Migratory behavior and winter geography drive differential range shifts of eastern birds in response to recent climate change. *Proceedings of the National Academy of Sciences*, 117(23), 12897–12903.
- Saino, N., Bazzi, G., Gatti, E., Caprioli, M., Cecere, J. G., Possenti, C. D., Galimberti, A., Orioli, V., Bani, L., Rubolini, D., Gianfranceschi, L., and Spina, F. (2015). Polymorphism at the Clock gene predicts phenology of long-distance migration in birds. *Molecular Ecology*, 24(8), 1758–1773.
- Shannon, C. P., Chen, V., Takhar, M., Hollander, Z., Balshaw, R., McManus, B. M., Tebbutt, S. J., Sin, D. D., and Ng, R. T. (2016). SABRE: A method for assessing the stability of gene modules in complex tissues and subject populations. *BMC Bioinformatics*, 17, 460.
- Sevilla, M., Udler, M., Kim, H., Hsu, S., Meigs, J. B., and Manning, A. (2020). Effect of fasting insulin-derived gene variant clusters on cardiometabolic traits. *Diabetes*, 1649-P.
- Sibley, D. A. (2014). *The Sibley Guide to Birds*. Second edition. Alfred A. Knopf, New York.
- Sunde, J., Yildirim, Y., Tibblin, P., and Forsman, A. (2020). Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics*, 11(218), 1-17.
- Swanson, D. L., and Merkord, C. (2013). Seasonal phenotypic flexibility of flight muscle size in small birds: A comparison of ultrasonography and tissue mass measurements. *Journal of Ornithology*, 154(1), 119–127.
- Swanson, D. L., Ingold, J. L., and Galati, R. (2020). Golden-crowned kinglet (*Regulus satrapa*), version 1.0. *Birds of the World*.
- Tallmon, D. A., Luikart, G., and Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, 19(9), 489–496.
- Toews, D. P. L., Campagna, L., Taylor, S. A., Balakrishnan, C. N., Baldassarre, D. T., Deane-Coe, P. E., Harvey, M. G., Hooper, D. M., Irwin, D. E., Judy, C. D., Mason,

- N. A., McCormack, J. E., McCracken, K. G., Oliveros, C. H., Safran, R. J., Scordato, E. S. C., Stryjewski, K. F., Tigano, A., Uy, J. A. C., and Winger, B. M. (2015). Genomic approaches to understanding population divergence and speciation in birds. *The Auk*, 133(1), 13–30.
- Toews, D. P. L., Taylor, S. A., Streby, H. M., Kramer, G. R., and Lovette, I. J. (2019). Selection on *VPS13A* linked to migration in a songbird. *Proceedings of the National Academy of Sciences*, 116(37), 18272–18274.
- Vendrami, D. L. J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M. S., Leese, F., McMinn, C., Moore, H., and Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: Implications for understanding phenotypic plasticity. *Royal Society Open Science*, 4(2), 160548.
- Wang, X., Lu, B., Shao, L., Li, Z., Ali, A., Yu, F., Fu, Z., and Sun, F. (2022). Genome-wide SNPs reveal the fine-scale population structure of *Laodelphax striatellus* in China using double-digest restriction site-associated DNA sequencing. *Genomics*, 114(3), 110329.
- Watson, G. E., M. A. Traylor, Jr., and E. Mayr. (1986). *Family Sylviidae, Old World warblers*. In *Check-list of Birds of the World*, Volume 11. Museum of Comparative Zoology, Cambridge, MA, USA. pp. 3–294
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370.
- Weir, J., and Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society B: Biological Sciences*, 271(1551), 1881–1887.
- Yuan, Y., Zhang, Y., Zhang, P., Liu, C., Wang, J., Gao, H., Hoelzel, A. R., Seim, I., Lv, M., Lin, M., Dong, L., Gao, H., Yang, Z., Caruso, F., Lin, W., da Fonseca, R. R., Wang, D., Wang, X., Rasmussen, M. H., ... Li, S. (2021). Comparative genomics provides insights into the aquatic adaptations of mammals. *Proceedings of the National Academy of Sciences*, 118(37), e2106080118.
- Zimmerman, S. J., Aldridge, C. L., and Oyler-McCance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics*, 21(382), 1-16.

Chapter 3: Population genetics of an irruptive migrant (*Sitta canadensis*) using microsatellites and RADseq

3.1 Introduction

Species can develop genetically different populations over time for numerous reasons. Ultimately, there must be a barrier that is preventing the populations from interbreeding consistently for these genetically isolated populations to develop. These barriers include physical, environmental, or behavioural factors. Knowing what these barriers are is critical to understand population dynamics and identify populations that may be at an increased conservation risk. Genetically isolated populations may experience an increase of inbreeding and genetic drift, both of which can be informative for developing conservation efforts as these processes can decrease genetic diversity (Daniels and Walters, 2000; Haas et al., 2010; Tallmon et al., 2004). Conversely, if a species is homogenous throughout their range, then it can be inferred that there is likely a significant amount of gene flow between populations. This can suggest individuals are able to bridge potential barriers to gene flow, or no contemporary borders exist. Even small breaks in continuous habitat can influence patterns of gene flow, including in species with high dispersal capabilities (Adams and Burg, 2015; MacDonald et al., 2010; Perrin et al., 2020).

Historical barriers, such as glaciation, can have a strong influence on species' populations. The large ice sheets that covered North America until approximately 21,000 years ago forced high-latitude species to remain in glacial refugia on the peripheries of the ice (Peltier, 1994; Pielou, 1991). Numerous species show population structure that can be attributed to genetic drift from a long period in more than one glacial refugia (Burg et al., 2006; Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter,

2004). However, a species may not always reside in multiple glacial refugia. If a species remained in a single glacial refugium during the last glacial maximum (LGM), then the glaciation process was unlikely to be a contributor to any observed population differentiation. In this case, population structure would be caused by one or more other barriers to gene flow.

Areas of unique or isolated habitat can also be home to genetically isolated populations. One example is in the form of sky islands. A sky island refers to an area of elevated habitat that differs from the surrounding lowlands. This can result in the populations residing in the sky island becoming genetically distinct from other populations of the same species. This can be caused by genetic drift and/or differing selective pressures. For example, white-breasted nuthatches (*Sitta carolinensis*) have shown population differentiation in the Santa Rita sky islands in Arizona (Manthey and Moyle, 2015). In Canada, Cypress Hills Interprovincial Park is a sky island that provides a unique habitat to a variety of species. For example, Cypress Hills contains a genetically isolated population of snails and warbling vireos (*Vireo gilvus*) (Carpenter et al., 2022; Dempsey et al., 2020). These previous studies suggest that other species may also have genetically isolated groups located in Cypress Hills and other sky islands.

While physical barriers are typically associated with the development of population structure, behavioural differences can also create barriers to gene flow. Differing migratory strategies within a species can influence their population structure (Delmore et al., 2020; Ralston and Kirchman, 2013; Ruegg et al., 2014). It is hypothesized that migratory behavior typically enhances gene flow between populations (Delmore et al., 2020; Ruegg et al., 2014). This can cause migratory populations to be

difficult to differentiate from one another due to the frequent mixing of alleles which limits the genetic divergence of the populations. The genetic consequences of migratory behaviour are important to study as migratory birds utilize more diverse and widespread habitats than their resident counterparts. Birds that migrate seasonally require breeding grounds, wintering grounds, migratory flyways, and migratory stop-over sites, each of which can influence population structure (Kirby et al., 2008).

One species with interesting migratory habits is the red-breasted nuthatch (*Sitta canadensis*). Red-breasted nuthatches are small, widespread passerines that inhabit a large portion of North America. They are typically found in forests containing spruce and fir trees, however, they also frequent mixed wood forests and suburban habitats (Ghalambor and Martin, 2020). While they are resident across most of their range, the northernmost populations migrate south every winter (Harrap and Quinn, 1995). While not fully understood, these populations are thought to be within approximately 300 km of the farthest northern edge of their range beginning at northwest Alberta and stretching across to Labrador (Sibley, 2014). Furthermore, red-breasted nuthatches are known for their irruptive migration. Every two to five years, when conifer cone production is poor, these birds migrate south in large numbers, even those that are typically resident (Cannings et al., 1987; Dorn and Dorn, 1990). It is not known how these migratory patterns impact the species' population structure. Red-breasted nuthatches also inhabit the Cypress Hills and Black Hills sky islands, which may influence their population structure due to the isolated nature of sky islands.

Despite their large geographic range and varied migratory habits, no regional variation or subspecies have been identified in red-breasted nuthatches. No thorough

genetic studies across their range are currently available, meaning it is possible there is cryptic variation in the species that has not been detected. The goal of my study is to provide the first genetic study of red-breasted nuthatches across their range using microsatellite and genome wide SNP markers from restriction site-associated DNA sequencing (RADseq). Utilizing both microsatellite and SNP markers will allow me to look at both neutral and non-neutral markers. Additionally, utilizing microsatellite markers will allow me to screen hundreds of individuals whereas RADseq is much more cost prohibitive. SNPs from RADseq, however, can be used to investigate areas of the genome that may be under selection (Lundberg et al., 2017; Peterson et al., 2012). Utilizing these two types of molecular markers will allow me to provide a thorough summary of the population genetics of red-breasted nuthatches. I will be able to detect both fine and broad scale population structure, if present, using these markers (Garraway et al., 2013; Wang et al., 2022).

I hypothesize that there is little distinct population genetic structure within the 28 red-breasted nuthatch populations included in my study. They have large dispersal and migratory capabilities as seen by their irruptive migration. If migration enhances gene flow between populations, then irruptive migrations may help prevent reproductive isolation. Red-breasted nuthatches may have resided in a single glacial refugia during the LGM, which would prevent genetic divergence. I do not predict that there will be significant population structure between the sky island populations and the surrounding areas, however, there may be genes or gene regions that are unique to the area due to genetic drift or different selective pressures. Regardless of my findings, I will be able to provide the first study on population genetics of red-breasted nuthatches.

3.2 Methods

3.2.1 Sample collection and DNA extraction

Blood samples were collected during the summer breeding season. Birds were caught using mist nets and song playback and were fitted with numbered aluminum bands to prevent resampling of the same individual. Samples were stored in ethanol at ambient temperature until they were returned to the lab where they were stored at -20°C.

Additional samples in the form of tissue (n=236) were obtained from museums. A total of 420 samples from 28 locations were used in my study (Table 3.1). DNA from most of the samples (n=324) was extracted using a modified Chelex extraction protocol (Walsh et al., 1991). DNA from the remaining samples (n=96) was extracted using a modified salting-out DNA extraction protocol (Miller et al., 1988). This subset of 96 samples was sent for RADseq which requires higher purification of DNA, which the salting-out method provides (Graham et al., 2015).

3.2.2 Microsatellite markers

3.2.2.1 Genotyping

A subset of six samples from geographically distant (>100 km apart) locations were used to optimize and test microsatellite loci for variation prior to large scale screening. I tested these samples at 29 different microsatellite loci. Eight of these loci were polymorphic and therefore used to genotype all available samples. Six of the loci were initially developed for the brown-headed nuthatch (*Sitta pusilla*): Spul 5-6, Spul 5-22, Spul 4-30, Spul 4-31, Spul 4-3 and Spul E19 (Haas et al., 2009). The other two loci, PmaTAGn71 (Pman 71) and Pat 43, were originally developed for other passerine species

(Otter et al., 1998; Saladin et al., 2003). Expected and observed heterozygosity for each locus was calculated in GenAlEx 6.5 using a frequency-based analyses (Peakall and Smouse, 2006, 2012). Arlequin 3.5.2.2 was used to check for deviations from Hardy-Weinberg equilibrium and linkage disequilibrium, and to calculate the average number of alleles per locus (Excoffier and Lisher, 2010).

DNA was amplified in 10 µl PCR reactions (1x FroggaBio buffer, 0.25 mM dNTP, 0.5 µM forward primer, 1 µM reverse primer, 0.05 µM fluorescently labelled M13 tags and 0.5 units of FroggaBio taq polymerase). The PCR consisted of one cycle of 120 s at 94 °C, 7 cycles of 60 s at 94 °C, 45 s at T1, 30 s at 72 °C, 28 cycles of 30 s at 89 °C, 30 s at T2, 45 s at 72 °C and a final cycle of 300 s at 72 °C and 20 s at 4 °C. For the six Spul primers, the annealing temperatures were T1=55 °C and T2=57 °C. Pman 71 and Pat 43 were amplified with T1=50 °C and T2=52 °C. The PCR products were mixed with a loading dye, denatured at 94 °C and run on a 6% acrylamide gel on a Li-Cor 4300 DNA analyzer. Gels were scored independently by two separate people to ensure reliable scoring. Additionally, the same three positive controls were used with each locus for consistent scoring. Samples that amplified poorly even after additional attempts (failed at >50% of loci), were removed from downstream analyses.

3.2.2.2 Population structure

STRUCTURE v2.3.4 was used to examine the population structure using microsatellite genotypes (Hubisz et al., 2009). STRUCTURE assigns individuals to a genetic cluster using a Bayesian model and a Markov chain Monte Carlo (MCMC) method (Pritchard et al., 2000). The number of genetic clusters (K) is pre-defined by the user, and the software output shows the proportion of ancestry (Q) of each individual to

each genetic cluster. In this analysis I applied the admixture model, used correlated alleles, and indicated sampling locations as *locpriors*. *Locpriors* allow the software to utilize sampling location in the analyses which can improve the model and detect weak population structure (Hubisz et al., 2009; Porras-Hurtado et al., 2013). To determine the best value of K, I evaluated values of delta K, the posterior probabilities of K and visualized the STRUCTURE plots as recommended in the STRUCTURE manual (Pritchard et al., 2000). The STRUCTURE analyses were run with a 50,000 burn-in length and 100,000 McMC repetitions for K=1 to 4 (10 runs for each K) with all 420 samples. STRUCTURE results for each K value were averaged and visualized using Structure Plot v.2.0 (Ramasamy et al., 2014).

The GenAlEx 6.5 was used to create a genetic distance matrix that calculated the average genetic distance between each population. The SWBC and NBC populations were removed from the genetic distance matrix as there were only one or two samples in those populations. Using the genetic distance matrix, a principal coordinate analysis (PCoA) was performed in GenAlEx 6.5 to visualize differences between populations. The resulting PCoA was color coded for each population and different point shapes were used to indicate populations east and west of the Rocky Mountains. This was done to visualize if mountain ranges were potentially a geographic barrier to gene flow.

3.2.2.3 F_{ST} and heterozygosity

To examine differentiation between populations, pairwise F_{ST} values were calculated for all populations. F_{ST} refers to the fixation index, where allele frequencies are compared between two populations to evaluate how different they are (Hartl and Clark, 1997). Pairwise F_{ST} and p -values were calculated in Arlequin 3.5.2.2, and the p -values

were corrected using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995). This method controls the false discovery rate when performing multiple comparisons. Arlequin was also used to calculate observed and expected heterozygosity in each population. Populations with five or less individuals were excluded from these analyses. The central and south Manitoba populations were combined for this analysis due to only having four central Manitoba samples. Additionally, these two sites were close together with no obvious geographical barrier separating them that would suggest they are genetically differentiated.

3.2.3. RADseq and SNP analyses

3.2.3.1 RADseq

The RADseq dataset was prepared using three enzymes to digest the DNA before sequencing. Library preparation prior to sequencing was performed at Université Laval before samples were sequenced on an Illumina NovaSeq 6000 with paired-end reads at Génome Québec. During library preparation, DNA was digested with the *PstI* (6-base pair cutter), *NsiI* (5 base-pair cutter) and *MspI* (4 base-pair cutter) restriction enzymes.

3.2.3.2 Data processing and filtering

The raw reads were run through Fastqc 0.11.5 to generate quality reports to ensure that there were no significant errors during sequencing (Andrews, 2011). Raw reads sent from Génome Québec were demultiplexed using Sabre 1.00 to assign reads to individual samples. Adapter and barcode sequences were removed from reads using Cutadapt (Martin, 2011). Cutadapt was also used to trim reads to 80 base pairs in length as the program that calls SNPs requires all reads to be the same length.

Genotyping and variant calling were performed using Stacks 2.3e, which is specifically designed for RADseq data (Catchen et al., 2013). I used the Stacks *de novo* pipeline with default parameters to analyze my data without the use of a reference genome. The *write-single-snp* command was enabled to reduce redundancy in the data and remove potentially linked SNPs. The *populations* function within the *de novo* pipeline indicates number of loci that are not in Hardy-Weinberg equilibrium (HWE). The variant call format (vcf) file produced by Stacks was filtered using VCFtools 0.1.16 (Danecek et al., 2011). SNPs with more than 30% missing data were removed before removing individuals with more than 30% missing data, leaving 77,361 SNPs and 80 individuals for downstream analyses.

3.2.3.3 Population analyses

Mean observed and expected heterozygosity for each population were calculated in Arlequin 3.5.2.2. Arlequin was also used to calculate pairwise F_{ST} and the p -values for these comparisons (Excoffier and Lisher, 2010). The p -values were corrected using the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995).

A discriminant analysis of principal components (DAPC) was performed using Adegnet 2.0.0 to visualize potential population structure in the dataset (Jombart, 2008). First, data were transformed using a PCA to analyze differences between individuals. Then, a discriminant analysis was performed to maximize the distance between clusters as defined by a value of K . This allows detection of even fine-scale population structure in a large SNP dataset. The DAPC was run for $K=1$ to $K=4$ even though the Bayesian Information Criterion indicated that the best value of K was 1. The resulting DAPC was formatted like the PCoA using microsatellite data; different shapes were used to represent

populations east and west of the Rocky Mountains to visualize if this mountain range influenced population structure.

An ancestry matrix was constructed using LEA: An R Package for Landscape and Ecological Association Studies with the cross-entropy function enabled (Frichot and Francois, 2015). The program ran for 100 iterations with the number of assumed populations between 1 and 4. The ancestry matrices were visualized in R-Studio 2021.09.0.

A phylogenetic tree was made using poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction (Kamvar et al., 2014). The tree was based on a genetic distance matrix and was constructed using the UPGMA (unweighted pair group method with arithmetic mean) method with 100 bootstrap supports for each node. A minimum spanning network (MSN) was also built in poppr using a genetic distance matrix. This method visualizes relationships among individuals in the data while minimizing the total genetic distance observed.

3.2.3.4 Isolation-by-distance

A Mantel test for Matrix Correspondence was performed using GenAlEx 6.5 to determine if there was a significant correlation ($P < 0.05$) between geographic distance and genetic distance. Genetic distances were calculated in ADEGENET 2.0.0 and the geographic distance matrix was generated in GenAlEx 6.5 using the sampling coordinates for each sample. The paired Mantel test was then run with 99 permutations in GenAlEx 6.5.

3.2.3.5 Aligning to a reference genome

Prior to performing any genome scans, sequences were aligned to a Eurasian nuthatch (*Sitta europaea*) reference genome at scaffold assembly, as this was the most closely related species that had an annotated reference genome available (GenBank accession number VXBS000000000). A reference genome is not required for general population analyses; however, it is critical when examining specific areas of the genome that are differentiated between populations. The reference genome was indexed using Samtools 0.1.2 (Danecek et al., 2021). The Burrows-Wheeler Alignment Tool (bwa) was used to align the forward and reverse reads to the reference genome and mark low quality alignments as secondary (Li and Durbin, 2009). Samtools 0.1.2 was used a second time to sort and index the resulting binary files produced by the alignment process. The sorted and aligned files were run through the Stacks 2.3e *ref_map* pipeline using the default parameters to call variants for genome scans. The vcf output was filtered using VCFtools 0.1.16 to remove SNPs with more than 40% missing data and individuals with more than 50% missing data. This left 69,860 SNPs and 79 individuals for genome scans and gene ontology analyses.

3.2.3.6 Genome scans

All genome scans were performed using the aligned and filtered dataset using 69,860 SNPs. VCFtools 0.1.16 was used to calculate F_{ST} values in a sliding window across the genome between potentially genetically distinct populations. Specifically, I was interested in investigating the role of sky islands in population differentiation. Each sky island (Cypress Hills and Black Hills) was compared to the three closest populations not separated by a body of water or mountain range. This was done because large

discrepancies in sample size can alter F_{ST} values and introduce false positives when looking for areas of the genome under selection (Bhatia et al., 2013). Cypress Hills samples were compared to SAB, MT, and MB while the Black Hills were compared to MT, CO and MB. The same program was also used to calculate π (nucleotide diversity) and Tajima's D in each population of interest and the surrounding populations. Nucleotide diversity (π) and F_{ST} were calculated in 20 kb windows with a 20 kb step, while Tajima's D was calculated in bins of 20 kb. Previous research looking for differentiated regions have used windows from 10-100 kb, however, determining a suitable window size will vary from study to study based on SNP density and linkage disequilibrium decay (Carlson et al., 2005; Lou et al., 2021; Lowry et al., 2017; Lundberg et al., 2017; Ravinet et al., 2018). High or low π values combined with elevated F_{ST} in an area of the genome can suggest that region is under selection (Ravinet et al., 2018). All Manhattan plots were visualized in Microsoft Excel. To determine areas of the genome potentially under selection or significantly different between populations, the 99th percentile of windowed F_{ST} values were calculated in R Studio 2021.09.0. Additionally, the 99th percentile of highest and lowest π values were calculated for each group of interest (Cypress Hills, Black Hills). The Manhattan plots were also visually examined for any distinct peaks that may fall under the 99th percentile. These peaks, if present, were also considered for outlier analyses. Windows with high F_{ST} values and high or low π values were used to look at their Tajima's D values and any genes in or close to those regions. Outlier windows were also compared between the two sky islands to determine if there was any overlap.

Any window that contained F_{ST} values in the 99th percentile and were also in the highest or lowest π values was extracted from the dataset. These outlier SNPs were subset to a new vcf file using VCFtools 0.1.16 and a PCoA was run using only these outlier SNPs. PCoAs were run instead of DAPCs for outlier SNPs because a DAPC must be provided with more than two clusters to analyze and in this case, I was only interested in two groups. This was done to see if these outlier SNPs represented genotypic differences between the sky islands and the surrounding areas.

3.3 Results

3.3.1 Microsatellite analyses

3.3.1.1 Population structure

A total of 420 red-breasted nuthatches from 28 populations were genotyped using eight variable microsatellite loci (Table 3.2). Arlequin did not detect any significant departures from Hardy-Weinberg equilibrium or any linkage disequilibrium.

The STRUCTURE run with all 28 populations showed that the best number of clusters was $K=1$. When examining populations with *locpriors* at $K=2$, there was no obvious population structure present (Figure 3.2). $K=3$ and $K=4$ also did not show any distinct patterns. Using STRUCTURE, I did not find any clear differentiation between populations.

The PCoA comparing populations showed the majority of the populations clustering relatively close together (Figure 3.3). The first and second axes explained 35.54% and 24.56% of the variation, respectively. A handful of populations (NS/NB, AK, CO, and SD) appeared to separate weakly from the main cluster.

3.3.1.2 F_{ST} and heterozygosity

The pairwise F_{ST} values using eight microsatellite loci had no value exceeding 0.15 (Table 3.3). After Benjamini and Hochberg corrections, only 28 out of 276 comparisons were significant. There was no consistent pattern of which population comparisons had significant F_{ST} values. Observed heterozygosities ranged from 0.519 (MB) to 0.624 (SD) (Table 3.5). The difference between observed and expected heterozygosities within any population was no greater than 0.100. The OR, VI, SEBC and SK populations had the highest observed heterozygosity and MB and LETH had the lowest.

3.3.2. RADseq and SNP analyses

3.3.2.1 Data processing and filtering

The RADseq process returned 546,509,228 paired-end reads with a mean coverage of 20.4x per individual and a minimum coverage of 4.4x per individual. Approximately 96.9% of paired-end reads were successfully aligned to a stack in Stacks 2.3e. All Phred quality scores from sequencing were >30. Prior to filtering, Stacks 2.3 *de novo* called a total of 1,439,968 SNPs. Post filtering, 77,361 SNPs and 80 individuals were retained for population analyses.

3.3.2.2 Population analyses

Overall, the observed and expected heterozygosities within each population were similar, with no two values >0.04 different (Table 3.5). The Black Hill and Cypress Hills populations had the lowest observed heterozygosity of 0.191. Alaska, Nova Scotia/New Brunswick and North Carolina had the highest observed heterozygosity. Observed and

expected heterozygosities were the most different in Alaska, North Carolina, and Vancouver Island populations.

The calculated F_{ST} values ranged from 0.00001 to 0.04. Out of 91 pairwise comparisons, 15 were statistically significant (Table 3.4). From the statistically significant F_{ST} comparisons, 13 of the 14 involved at least one peripheral population. Populations close to the edge of the red-breasted nuthatch range were considered peripheral (AK, VI, CA, CO, NC, NS/NB, NL).

The DAPC with 80 individuals and 13 populations did not show any obvious clustering along any axis (Figure 3.4). Almost all individuals clustered randomly in one group, with a few random individuals pulling farther away from the cluster. These individuals did not come from the same or nearby populations.

The ancestry matrix generated with all individuals did not show any distinct clustering at $K=2-4$. Individuals pulled out randomly and individuals from the same population did not cluster together. The cross-entropy was lowest at $K=1$, indicating that $K=1$ is likely the best number of clusters for the dataset.

The UPGMA tree appeared very shallow with almost all individuals branching off at or before a genetic distance of 0.010 (Figure 3.5). Overall, no population grouped together along the tree tips in a way that would indicate the genetic structure. The MSN was starburst with most individuals branching out from a central node (Figure 3.6). Two individuals from NS/NB pulled away from the central cluster and one individual from NL also separated slightly.

3.3.2.3 Isolation-by-distance

The Mantel test for Matrix Correspondence indicated that geographic distance and genetic distance were not correlated ($R^2=0.0037$, $P=0.230$). This indicates that geographic distance is likely not responsible for any potential genetic differences between red-breasted nuthatch populations.

3.3.2.4 Aligning to a reference genome

Approximately 73% of reads were successfully mapped to the reference genome and kept for SNP calling by Stacks 2.3e. The remaining 27% of reads were removed from analyses due to insufficient mapping qualities or inability to be mapped to the reference genome. This left a total of 1,200,329 loci prior to data filtering. After filtering, 69,860 SNPs and 79 individuals were kept for genome scans and gene ontology analyses. Compared to the dataset developed without the use of a reference genome, this dataset lost two individuals (1 SAB, 1 VI) and retained one sample from CA that was not included in the initial analyses. Apart from these three samples, the remainder of the individuals were consistent between the datasets.

3.3.2.5 Genome scans and outlier PCoA

The F_{ST} genome scan comparing Cypress Hills to the surrounding populations resulted in a mean F_{ST} of 0.0063 across the genome. There were 25 windows in the 99th percentile of F_{ST} values (>0.332). Four of these windows were also associated with 99th percentile of lowest π values in the Cypress Hills populations. These four windows contained 11 SNPs and were considered outliers to analyze using a PCoA. In these four windows the Tajima's D values were 1.53, -0.95, 0.65 and 1.34 in the Cypress Hills

population and 1.54, -1.43, -0.81 and 0.16 in the surrounding populations. Overall, the Tajima's D values were more positive in Cypress Hills than in the surrounding populations.

The four windows that showed separation in the PCoA contained or were close to the RHOBTB1, TMEM26, UR2R, GPC5 and SFRP2 genes (Table 3.6). The RHOBTB1 gene is linked to tumor suppression, TMEM26 is involved in beige fat development, UR2R is thought to affect vasoconstriction and cardiovascular events, GPC5 regulates limb and feather bud development and SFRP2 is related to cardiac development and disease (Kane and Giaid, 2011; Komolka et al., 2017; McKendy et al., 2010; Mukohda et al., 2017; Saad et al., 2017; Town et al., 2011; Wu et al., 2020, Xu et al., 2016).

The F_{ST} genome scan comparing the Black Hills to the surrounding populations gave a mean F_{ST} of 0.012 across the genome. There were 38 windows in the 99th percentile of F_{ST} values (>0.330). Three of these windows were also associated with the lowest π values in the Black Hills populations and were classified as outliers. These were used to construct a PCoA using only these outlier windows. There were 13 SNPs in these three windows. These windows had a Tajima's D value of 1.41, -1.31 and 0.022 in the Black Hills population and -0.93, -1.47 and -0.81 in the surrounding populations.

The three windows that showed separation of Black Hills from the surrounding populations contained or were near the GPR34, GPR82, DUSP10 and PDCD6 genes (Table 3.6). The GPR34 and GPR82 genes were found in the same window and are often arranged together in the genome (Schöneberg et al., 2018). They are typically thought to be expressed in vascular or lymphatic tissues (Schöneberg et al., 2018). The DUSP10

gene is involved in immune response and the PDCD6 gene is critical for programmed cell death (Dar et al., 2022; JiaQing et al., 2013).

The PCoAs constructed with only outlier SNPs from the sky island populations showed separation of the sky islands from their nearby populations along the first PC (30.85% for Cypress Hills and 34.73% for Black Hills) (Figure 3.7). In both sky island PCoAs, a few individuals from the surrounding populations either mixed in with or were not significantly separated from the sky island population. The majority of individuals, however, were distinct.

3.4 Discussion

3.4.1 Population structure and trends

Overall, I did not find any evidence to suggest strong population structure between red-breasted nuthatch populations across North America. Using two types of genetic markers and multiple analyses, no clear population clusters were detected. In my microsatellite analyses, STRUCTURE did not detect any population structure. The significant F_{ST} values using microsatellites could potentially be attributed to some populations having a smaller sample size (≤ 15), as when using fewer genetic markers, a larger sample size is required to accurately estimate allele frequencies (Aguirre-Liguori et al., 2020; Hale et al., 2012; Sunde et al., 2020). Fewer individuals in a population can increase F_{ST} estimates (Aguirre-Liguori et al., 2020). Overall, elevated F_{ST} values did not seem to follow a pattern that would indicate any consistent genetic distinctiveness when using microsatellite markers. The results from STRUCTURE combined with the results

from F_{ST} analyses with microsatellites consistently suggest that there is little population structure in red-breasted nuthatch populations.

My analyses using 77,361 SNPs from RADseq also discovered similar limited genetic structure as the microsatellite analysis. When using the LEA clustering software, the ancestry matrix plots showed random individuals clustering together with populations splitting with no consistent pattern. The DAPCs constructed for all populations did not show any clustering. The vast majority of significant F_{ST} values (13/14) were attributed to at least one peripheral population. Even though several geographically distant populations showed significant differentiation with F_{ST} , large geographic distance between populations did not consistently correspond to a significant F_{ST} value. Additionally, the isolation-by-distance test indicated that genetic distance and geographic distance were not correlated. Sample size was unlikely to be an issue for the SNP F_{ST} calculations because when using large numbers of SNPs (>1,000), sample sizes can be as low as four individuals per population and still generate accurate F_{ST} values (Willing et al., 2012). The aforementioned results suggest that there are high levels of gene flow between populations of red-breasted nuthatches, limiting genetic differentiation.

3.4.2 Molecular markers

Little population structure was detected in red-breasted nuthatches using two types of molecular markers and these markers provided similar results. Previous studies have shown that microsatellite markers and SNPs typically show the same general population structuring (Hodel et al., 2017; Jeffries et al., 2016; Lemopoulos et al., 2019; Sunde et al., 2020; Zimmerman et al., 2020). RADseq using large numbers of SNPs, however, can produce more distinct clustering and is often able to uncover more fine scale population

structure than microsatellites (Hodel et al., 2017; Jefferies et al., 2016; Sunde et al., 2020; Vendrami et al., 2017). This is likely due to the increased number of markers. For example, using over 77,000 SNPs but only eight microsatellite markers in this study. While RADseq heterozygosity measures are typically lower than in microsatellites, SNPs are better at estimating heterozygosity and population level genomic diversity, potentially because SNPs have a wider breadth across the genome (Sunde et al., 2020; Zimmerman et al., 2020). F_{ST} values between microsatellite and SNP analyses are not comparable as the magnitude of differentiation metrics vary substantially between the two datasets (Zimmerman et al., 2020). If strong population structure was present, significant F_{ST} values would likely have been found in the same populations using both molecular markers, however the values themselves would be very different. The similarity in results between the microsatellite data and the RADseq data follow the same pattern as previous studies, although if population structure is present, it is more likely that SNP analyses will uncover it.

3.4.3 Irruptive migration and the LGM

While it may be surprising that a species with such a large geographic range does not have any significant population structure, there are several potential explanations for this pattern. One of the factors that could be facilitating ongoing gene flow between populations is their migratory habits. In addition to the northernmost populations migrating annually, they also show irruptive migration every two to five years in connection with cone production (Cannings et al., 1987; Dorn and Dorn, 1990). This happens in large numbers, as even typically resident individuals will migrate, but it is not known how this influences their population structure. Migratory behaviour normally

enhances gene flow between populations and if large numbers of red-breasted nuthatches are irrupting every few years from across their range, it is possible that these irruptions are facilitating gene flow between populations. Irruptive migration is also thought to have negative population consequences on red-breasted nuthatches (Dunn, 2019). The breeding season after a large irruption often shows a lower population density compared to the season prior to the irruption, suggesting that irruptive migration may increase mortality rates or that this species may be dispersing and colonizing new areas after an irruption (Dunn, 2019). If population density is low, individuals may need to travel farther to find a mate or move into previously occupied territories, both things that would increase genetic exchange between populations. Several other irruptive species have also shown little to no observable population structure despite a large geographic range (Alvarex et al., 2015; Hochachka et al., 1999; Marthinsen et al., 2008). Even the circumpolar snowy owl (*Bubo scandiacus*), another irruptive migrant, demonstrates high levels of gene flow between populations (Marthinsen et al., 2009). The combination of irruptive and mixed migration strategies in red-breasted nuthatches is one of the theories that could explain their limited population structure.

Another explanation is that they likely inhabited a single mainland glacial refugium during the LGM and rapidly expanded once the ice sheets began to melt. Many boreal species show an east-west population split that is attributed to time spent in two separate glacial refugia (Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004). The effects of the LGM differ from species to species, as species with similar distributions can show different contemporary genetic patterns even if they resided in the same glacial refugium (Ralston et al., 2021). Population structure can arise

after the LGM if the species was in one glacial refugium via other factors, meaning the ice sheets do not always cause population structure to arise (Ralston et al., 2021). Red-breasted nuthatches show a shallow phylogenetic tree in combination with a star shaped MSN, both of which suggest a recent population expansion (Davis et al., 2006; Milá et al., 2000). Recent population expansions can also cause populations to have little genetic differentiation, even if there are areas with restricted gene flow (Colbeck et al., 2008; Davis et al., 2006). During the LGM it seems most likely that this species resided in one glacial refugium.

There may have been a northeastern glacial refugium in the Atlantic in addition to a mainland refugium. This idea has been suggested in American redstarts (*Setophaga ruticilla*) as well, as their populations show some divergence in both Newfoundland and Nova Scotia/New Brunswick (Colbeck et al., 2008; Davis et al., 2006). In red-breasted nuthatches, 6/14 of the significant F_{ST} values using RADseq were attributed to NL or NS/NB (Table 3.4). Additionally, the PCoA using microsatellite data showed NS/NB separating slightly from the main cluster and the MSN using RADseq data had several individuals from NL and NS/NB pulling away (Figures 3.3 and 3.6). These results may suggest that there was in fact a second glacial refugium that red-breasted nuthatches inhabited in Atlantic Canada, however, their high dispersal capabilities and varied migratory habits could have caused enough gene flow to prevent these populations from remaining genetically isolated. While the presence of this second glacial refugia is possible, there is more evidence to suggest that this species was in a single mainland refugium.

3.4.4 *Peripheral populations*

In red-breasted nuthatches, populations on the periphery of their range may be of special interest. Using SNPs from RADseq data, 13 of the 14 significant F_{ST} values involved at least one population on the periphery of the species' range (Table 3.4). This pattern has been observed before in other species, with peripheral populations showing genetic differentiation compared to central populations (Botero-Delgadillo et al., 2017; Cassel-Lundhagen et al., 2009; Lami et al., 2001; Mendelsohn et al., 2020; Rossum et al., 2003). Populations on the edge of a species' range are usually less dense and their distribution may be fragmented (Gibson et al., 2009; Mendelsohn et al., 2020). These populations may experience different selective pressures, genetic drift or experienced bottlenecks during range expansion (Cassel-Lundhagen et al., 2009). For example, more northern latitudes are expected to be colder than those farther south, causing different environmental conditions for populations residing in different latitudes. Peripheral populations in both the high and low latitudes are important for conservation management, as they have different responses to climate change (Gibson et al., 2009; Hampe and Petit, 2005). While these populations may show some genetic differentiation in my study based on F_{ST} , my other analyses do not show peripheral populations separating from more central ones. This may be expected, however, because it is thought that gene flow from central populations can prevent peripheral populations from separating completely (Garcia-Ramos and Kirkpatrick, 2017). Even small numbers of migrants each generation into these populations can maintain significant levels of gene flow (Nathan et al., 2017). As the ranges of birds shift with climate changes, it will be

interesting to monitor these peripheral populations over time to see if this genetic pattern is maintained.

3.4.5 Sky islands

There was no evidence to suggest that the sky islands Cypress Hills and the Black Hills are home to strongly separated populations. Sky islands, including the two included in this study, have been found to contain genetically distinct populations of other species (Anderson et al., 2008; Carpenter et al., 2022; Dempsey et al., 2020; Kiesow et al., 2012; Manthey and Moyle 2015). These sky islands are elevated in comparison to the surrounding lowlands which creates a potential barrier to gene flow for some species. They can have drastically different habitat than the surrounding areas which can also prevent species from exchanging genes with other populations. The PCoAs performed with outlier windows showed some separation of the sky island populations and their nearby populations. This suggests that while there may not be significant differentiation between the sky islands and the rest of the populations overall, there may be a few genotypes that are more prevalent in these unique areas. The red-breasted nuthatches irruptive migratory patterns could potentially be preventing these sky islands from becoming genetically isolated. There are currently no studies that adequately report on the site fidelity of red-breasted nuthatches during irruption, although it has been loosely suggested that they have low fidelity (Dunn, 2019). While sky islands may act as isolated habitats for some species, red-breasted nuthatches do not follow this same pattern which could potentially be linked to their high dispersal capabilities.

Even though there may not be much distinct genetic differentiation between the sky islands and the surrounding areas, the suggested outlier windows appear to be

different. The genes in these windows may be different in the sky island populations due to genetic drift or selective pressures specific to the sky islands. Areas of elevated F_{ST} and low nucleotide diversity can suggest that these areas have undergone a selective sweep (Carlson et al., 2005; Ravinet et al., 2018). These patterns were overall observed for the outlier windows identified for the two sky island populations in this study. Multiple genes from both sky island populations were associated with cardiovascular regulation or activity (UR2R, GPR34/82 and SFPR2) (Schöneberg et al., 2018; Wu et al., 2020). These genes may be different due to different stressors and behaviours associated in different habitats. TMEM26 is involved in the development of beige fat which can lend itself to thermoregulation in differing climates (Komolka et al., 2017; Sotome et al., 2021). This could suggest that birds in the sky island populations may have different thermoregulatory requirements than those in the surrounding lowlands. While these windows may be the result of different selective pressures in the sky island populations, the sky island populations are not completely isolated from the surrounding areas. It is also important to acknowledge that RADseq sequences a small portion of the genome, causing important areas to potentially be missed during sequencing. Additionally, whole-genome sequencing provides more distinct peaks in genome scans compared to RADseq, making it easier to visualize and locate differentiated regions (Lou et al., 2021). Whole-genome sequencing may uncover stronger or more important signatures of selection in the sky island populations and should be considered for future studies of red-breasted nuthatches.

3.5 Conclusions

I have conducted the first population genetic analyses of red-breasted nuthatch populations across North America. Using a combination of different molecular markers, I was able to search for both broad and fine-scale population structure within the species. There was no strong population structure found in or between any populations with either molecular marker. This may be at least partially due to their migratory habits and expansion out of a single glacial refugium. To improve the current study, their migration patterns need to be studied in more detail. Sequencing samples with next-generation sequencing from the northern annually migrating red-breasted nuthatches would fill in a critical sample gap, as the RADseq portion of this study did not include any annually migrating populations. Additionally, satellite tracking methods implemented during irruptive years would greatly add to the understanding of irruptive migration of this species. Red-breasted nuthatches demonstrate some of the potential outcomes of widespread migratory behaviour on population genetics.

Table 3.1 Population abbreviations and sample sizes for both microsatellite and RADseq analyses.

Population	Microsatellites	RADseq
CA (California)	28	7
NV (Nevada)	9	
OR (Oregon)	30	
WA (Washington)	31	
NEWA (North-east Washington)	16	
AK (Alaska)	3	2
VI (Vancouver Island, British Columbia)	10	6
SWBC (South-west British Columbia)	2	
CBC (Central British Columbia)	15	
NBC (Northern British Columbia)	1	
SEBC (South-east British Columbia)	22	
ID (Idaho)	20	
UT (Utah)	10	
MT (Montana)	27	5
SAB (Southern Alberta)	18	8
LETH (Lethbridge, Alberta)	8	
CAB (Central Alberta)	16	
SEAB (South-east Alberta, Cypress Hills)	18	8
NEAB (North-east Alberta)	5	
WY (Wyoming)	11	
CO (Colorado)	19	7
SK (Saskatchewan)	12	
SD (South Dakota, Black Hills)	18	9
MB (Manitoba)	18	5
MI (Michigan)	27	7
NC (North Carolina)	7	3
NS/NB (Nova Scotia and New Brunswick)	11	6
NL (Newfoundland)	8	7
Total	420	80

Table 3.2 Microsatellite loci used in this study and the number of alleles found at each loci (n), the size of the products, annealing temperature (TA), observed and expected heterozygosity (H_o, H_e), and the paper and species the primers were originally developed for.

Locus	Primers (5' – 3')	Alleles (n)	Size (bp)	TA (°C)	H _e	H _o	Species	Source
Spul 5-6	F CTCCTTGTGCATGGTTGAA R TCTACTGTCCCACGGGTAAAA	4	233-295	55/57	0.39	0.33	<i>Sitta pusilla</i>	Haas et al., 2009
Spul 5-22	F AAGAAGGGGCTCTGCTTTATG R CCATTGGCAAACAAATTCATCCA	17	144-180	55/57	0.77	0.86	<i>S. pusilla</i>	Haas et al., 2009
Spul 4-30	F ATGCACTGGGTTCTGTGTT R GTTTGTTTACATTGCTGGAAAGG	19	264-312	55/57	0.82	0.84	<i>S. pusilla</i>	Haas et al., 2009
Spul 4-31	F CCCCAAACCCAACTCTGTTA R TGCATTGGTTCATTACTAGATGCT	7	269-295	55/57	0.31	0.37	<i>S. pusilla</i>	Haas et al., 2009
Spul E19	F TCCTGTGAGAGCAGCAAGAA R CTGGCATCAAAGGAAAGCAT	16	235-325	55/57	0.81	0.87	<i>S. pusilla</i>	Haas et al., 2009
Spul 4-3	F AGTCAGCACATGGAACCACA R AAACCCAGCAAACATTCCAC	15	356-402	55/57	0.85	0.90	<i>S. pusilla</i>	Haas et al., 2009
Pat 43	F ACAGGTAGTCAGAAATGGAAAG R GTATCCAGAGTCTTTGCTGATG	22	151-243	50/52	0.80	0.83	<i>Poecile atricapillus</i>	Otter et al., 1998
Pman 71	F TCAGCCTCCAAGGAAAACAG R GCATAAGCAACACCATGCAG	6	186-208	50/52	0.41	0.35	<i>Parus major</i>	Saladin, 2003

Table 3.3 Pairwise population F_{ST} values using eight microsatellite loci. Colors represent F_{ST} values as follows: Grey <0.001, yellow 0.001 to 0.009, orange 0.010 to 0.090 and red >0.090. See Appendix 2.1 for F_{ST} values and the corresponding p -values.

	CA	NV	OR	WA	NEWA	VI	CBC	SEBC	ID	UT	MT	SAB	LETH	CAB	SEAB	WY	CO	SK	SD	MB	MI	NC	NS/NB
CA	Grey																						
NV	Grey	Grey																					
OR	Grey	Grey	Grey																				
WA	Grey	Grey	Grey	Grey																			
NEWA	Grey	Grey	Grey	Grey	Grey																		
VI	Grey	Orange	Grey	Grey	Grey	Grey																	
CBC	Grey	Grey	Grey	Grey	Grey	Grey	Grey																
SEBC	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Grey															
ID	Grey	Grey	Yellow	Grey	Grey	Orange	Grey	Orange	Grey														
UT	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey													
MT	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Yellow	Grey	Grey												
SAB	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Grey											
LETH	Grey	Yellow	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey										
CAB	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey									
SEAB	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Grey	Grey								
WY	Grey	Grey	Grey	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey							
CO	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey						
SK	Grey	Grey	Grey	Orange	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Yellow	Grey	Grey	Yellow	Grey	Yellow	Grey					
SD	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey				
MB	Grey	Grey	Yellow	Grey	Yellow	Grey	Grey	Yellow	Yellow	Yellow	Yellow	Grey	Grey	Grey	Grey	Grey	Orange	Yellow	Grey	Grey			
MI	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey		
NC	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	
NS/NB	Red	Orange	Yellow	Grey	Orange	Orange	Grey	Yellow	Grey	Grey	Orange	Orange	Grey	Orange	Orange	Grey	Grey	Yellow	Yellow	Orange	Grey	Yellow	
NL	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey

Table 3.4 Pairwise population F_{ST} values using RADseq and 77,361 SNPs. Values that are highlighted and bolded were statistically significant after corrections ($P < 0.05$). P -values are provided above the diagonal and are bold if significant.

	CA	AK	VI	MT	SAB	SEAB	CO	SD	MB	MI	NC	NS/NB	NL
CA		0.0964	0.2821	0.1500	0.2571	0.1786	0.1143	0.1000	0.2321	0.1642	0.0929	0.0178	0.2929
AK	0.0110		0.1392	0.0393	0.0857	0.2000	0.0250	0.1750	0.2643	0.2964	0.1607	0.1464	0.0143
VI	0.0038	0.0206		0.1357	0.1857	0.3179	0.0607	0.2214	0.0321	0.1071	0.2036	0.0214	0.1107
MT	0.0033	0.016	0.0056		0.1714	0.0786	0.2179	0.2250	0.1571	0.0821	0.0464	0.0429	0.1929
SAB	0.0034	0.0120	0.0018	0.0017		0.1821	0.0536	0.0750	0.2286	0.2679	0.1036	0.2714	0.1179
SEAB	0.0038	0.0124	0.0023	0.0021	0.0028		0.2143	0.1964	0.2607	0.1893	0.1214	0.3071	0.0714
CO	0.0015	0.0229	0.0026	0.0052	0.0014	0.0015		0.1250	0.0071	0.3179	0.2786	0.0107	0.2857
SD	0.0034	0.0098	0.0015	0.0003	0.0024	0.0040	0.0002		0.1678	0.2750	0.2893	0.0500	0.2071
MB	0.0058	0.0118	0.0066	0.0023	0.0045	0.0049	0.0073	0.0032		0.0286	0.0679	0.2107	0.0643
MI	0.0017	0.0133	0.0028	0.0030	0.0018	0.0023	0.0022	0.0014	0.0053		0.1321	0.0893	0.1429
NC	0.0083	0.0175	0.0119	0.0108	0.0077	0.0077	0.0122	0.0062	0.0102	0.0077		0.1286	0.1536
NS/NB	0.0074	0.0183	0.0089	0.0141	0.0117	0.0101	0.0207	0.0124	0.0128	0.0115	0.0149		0.0571
NL	0.0046	0.0216	0.0037	0.0021	0.0017	0.0012	0.0046	0.0027	0.0037	0.0017	0.0115	0.0157	

Table 3.5 Observed and expected heterozygosities for the microsatellite and RADseq datasets.

	Microsatellites			RADseq		
Population	H_o	H_e	n	H_o	H_e	n
CA	0.574	0.548	28	0.209	0.228	7
AK				0.495	0.537	2
NV	0.561	0.514	9			
OR	0.618	0.576	30			
WA	0.581	0.578	31			
NEWA	0.572	0.545	16			
VI	0.610	0.570	10	0.224	0.250	6
CBC	0.589	0.543	15			
SEBC	0.604	0.598	22			
ID	0.571	0.551	20			
UT	0.592	0.554	10			
MT	0.584	0.576	27	0.266	0.282	5
SAB	0.596	0.569	18	0.191	0.210	8
LETH	0.538	0.533	8			
CAB	0.553	0.550	16			
SEAB	0.560	0.544	18	0.191	0.207	8
WY	0.566	0.566	11			
CO	0.597	0.561	19	0.201	0.221	7
SK	0.608	0.581	12			
SD	0.624	0.554	18	0.175	0.193	9
MB	0.519	0.529	18	0.264	0.285	5
MI	0.550	0.552	27	0.205	0.220	7
NC	0.583	0.534	7	0.368	0.402	3
NS/NB	0.564	0.560	11	0.296	0.326	6
NL	0.592	0.540	8	0.203	0.225	7

Table 3.6 Scaffolds that contained outlier windows in the sky island analyses and the genes associated with these windows.

Cypress Hills		
Scaffold	Gene	Function
VXBS01002960.1	RHOBTB1	Tumor suppression
	TMEM26	Beige fat development
VXBS01002275.1	UR2R	Cardiovascular activity
VXBS01000712.1	GPC5	Limb and feather development
VXBS01001979.1	SFRP2	Cardiac development
Black Hills		
Scaffold	Gene	Function
VXBS01003073.1	GPR34 and GPR82	Vascular or lymphatic tissues
VXBS01000199.1	DUSP10	Immune response
VXBS01001928.1	PDCD6	Programmed cell death

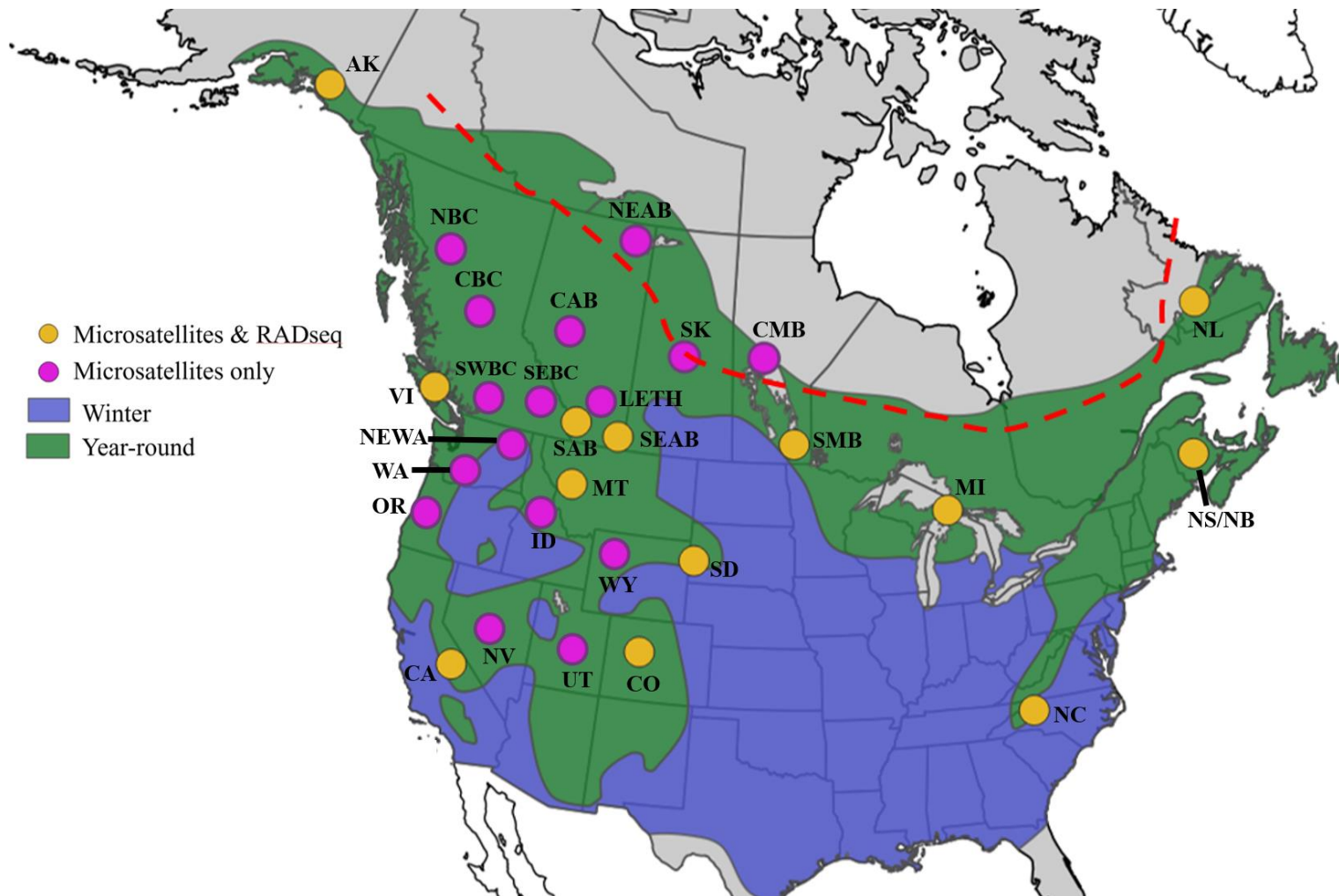


Figure 3.1 Range map and sample sites of Red-breasted nuthatches for this study using microsatellites and RADseq. Populations north of the red dashed line were considered migratory. This border was approximated from Sibley (2014). Range data was obtained from BirdLife International (2016) and the map was created in QGIS 3.

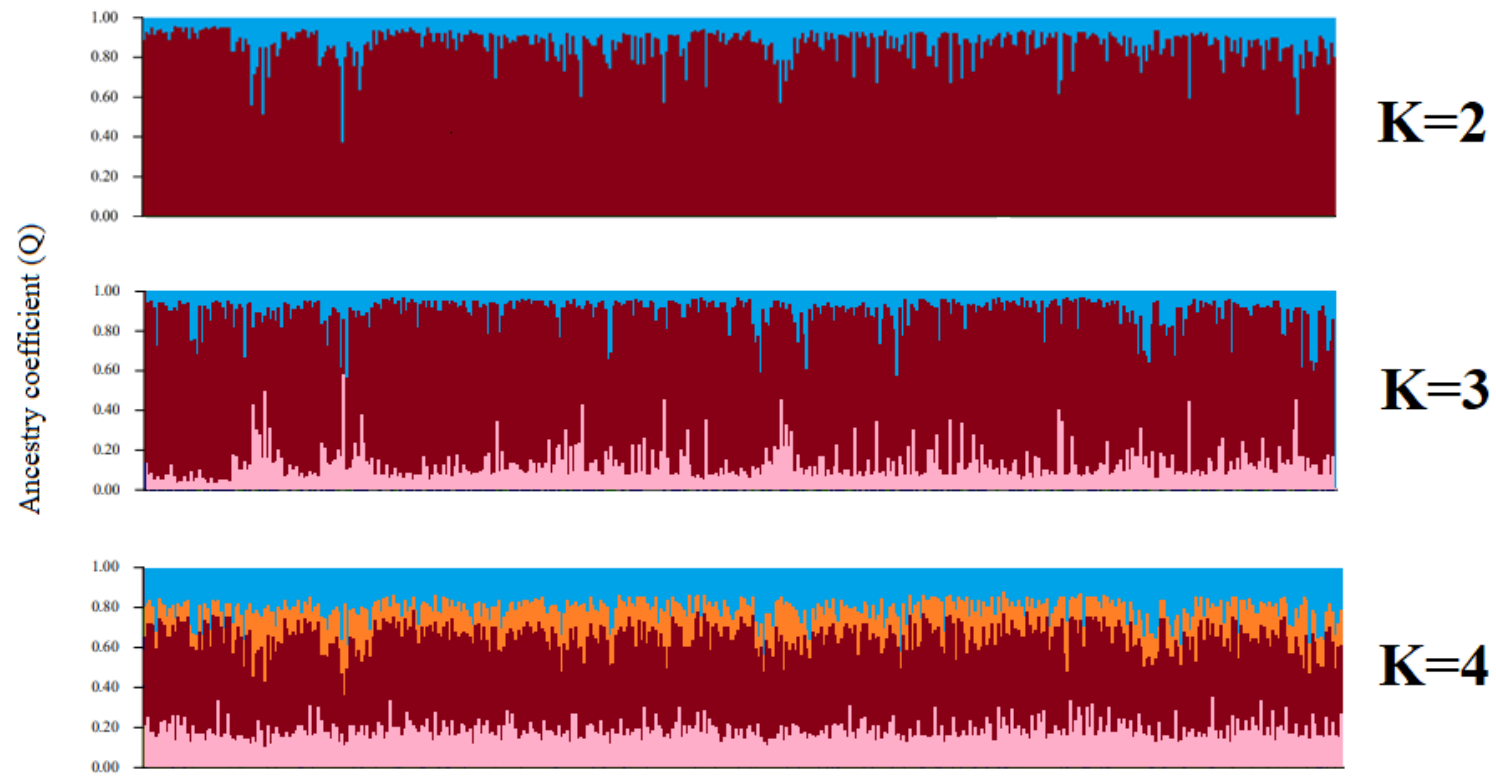


Figure 3.2 Hierarchical STRUCTURE plots for K=2 to K=4 of 420 red-breasted nuthatches using eight microsatellite loci.

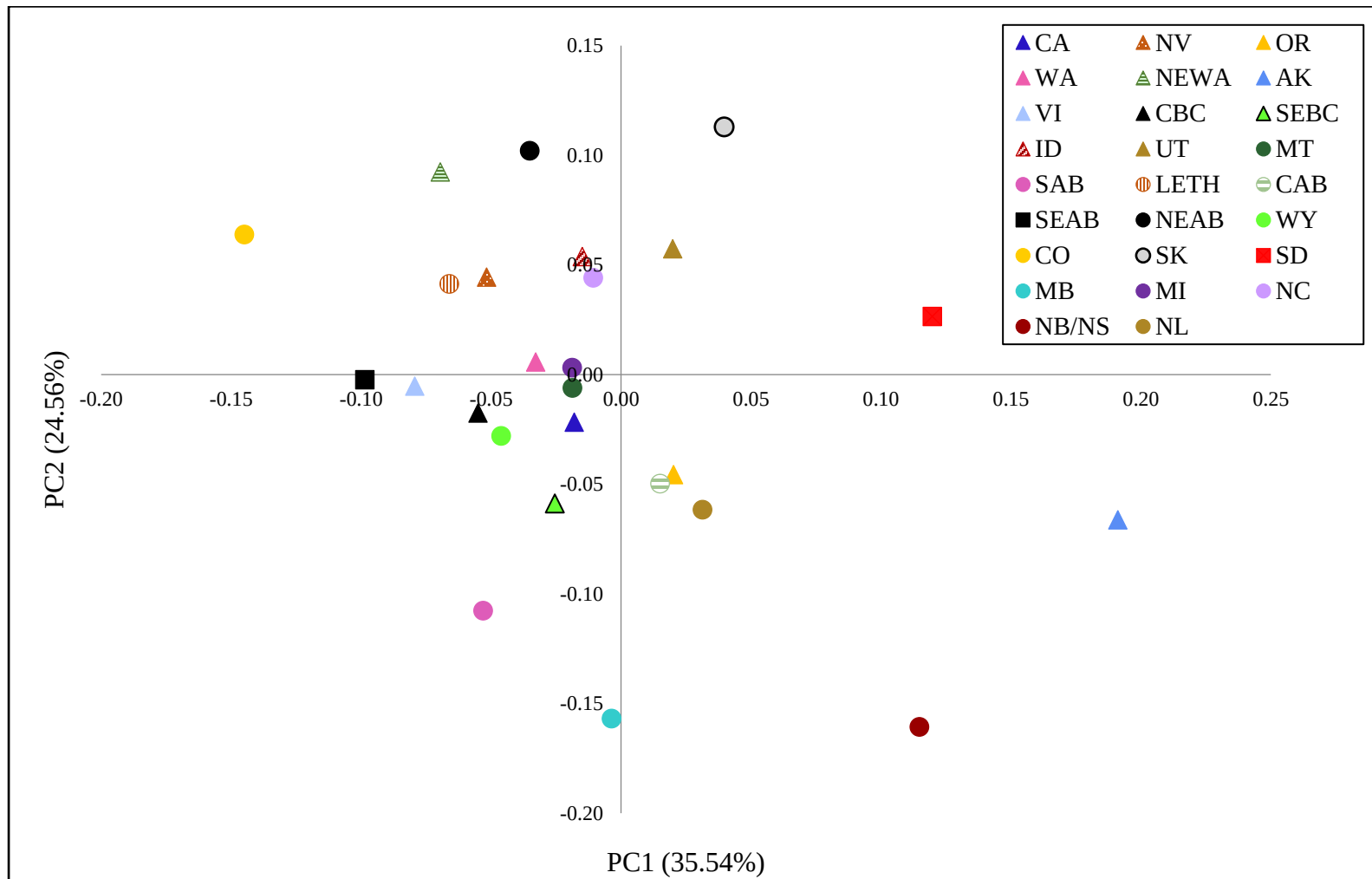


Figure 3.3 Principal coordinate analysis (PCoA) of 26 populations using eight microsatellite markers. Populations west of the Rocky Mountains are indicated by triangles while populations east of the Rocky Mountains are indicated by circles. The two sky islands (Cypress Hills and Black Hills) are shown as squares.

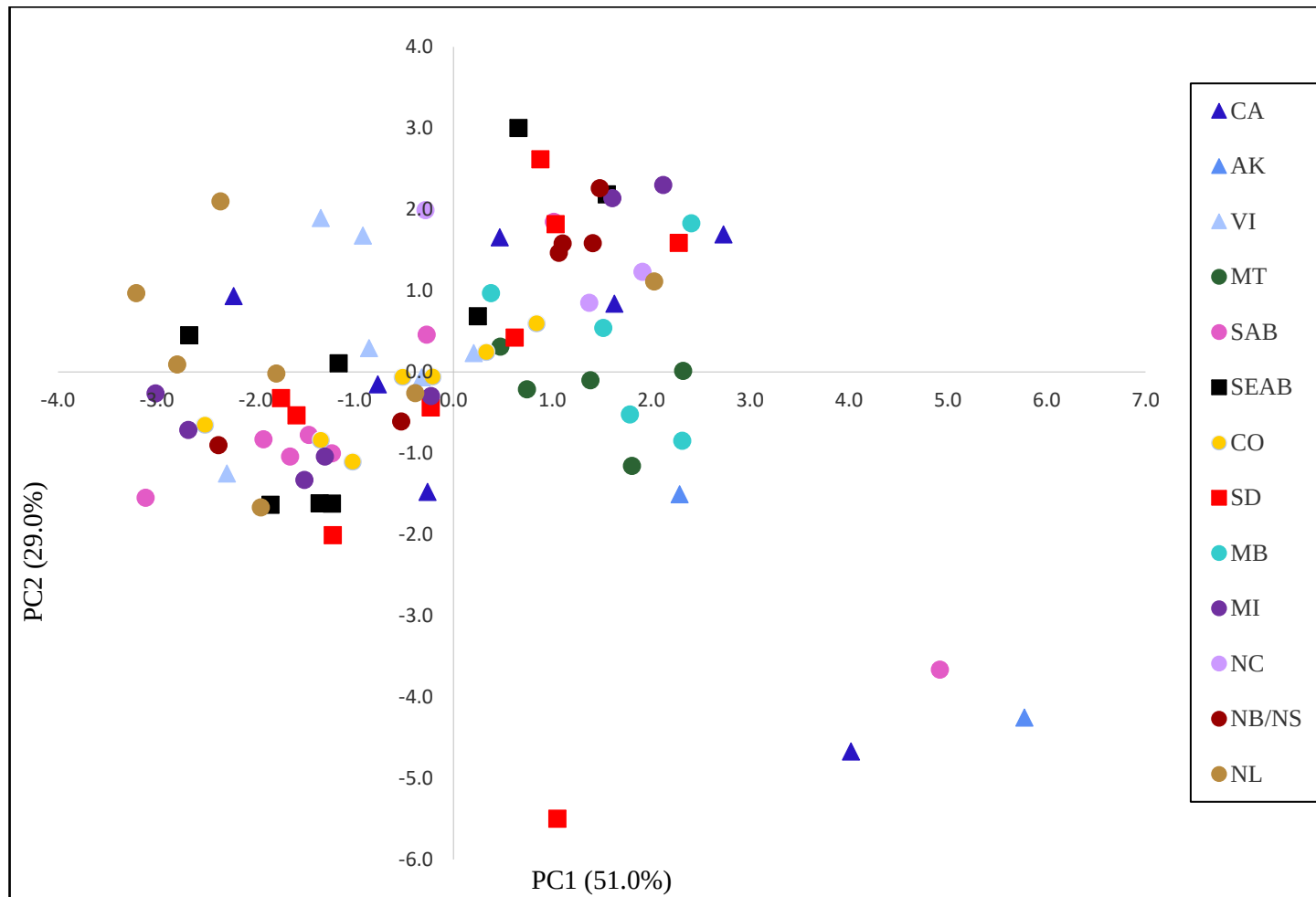


Figure 3.4 Discriminant analysis of principal components (DAPC) using 77,361 SNPs from RADseq with 80 individuals. Populations west of the Rocky Mountains are indicated by triangles while populations east of the Rocky Mountains are indicated by circles. The two sky islands (Cypress Hills and Black Hills) are shown as squares.

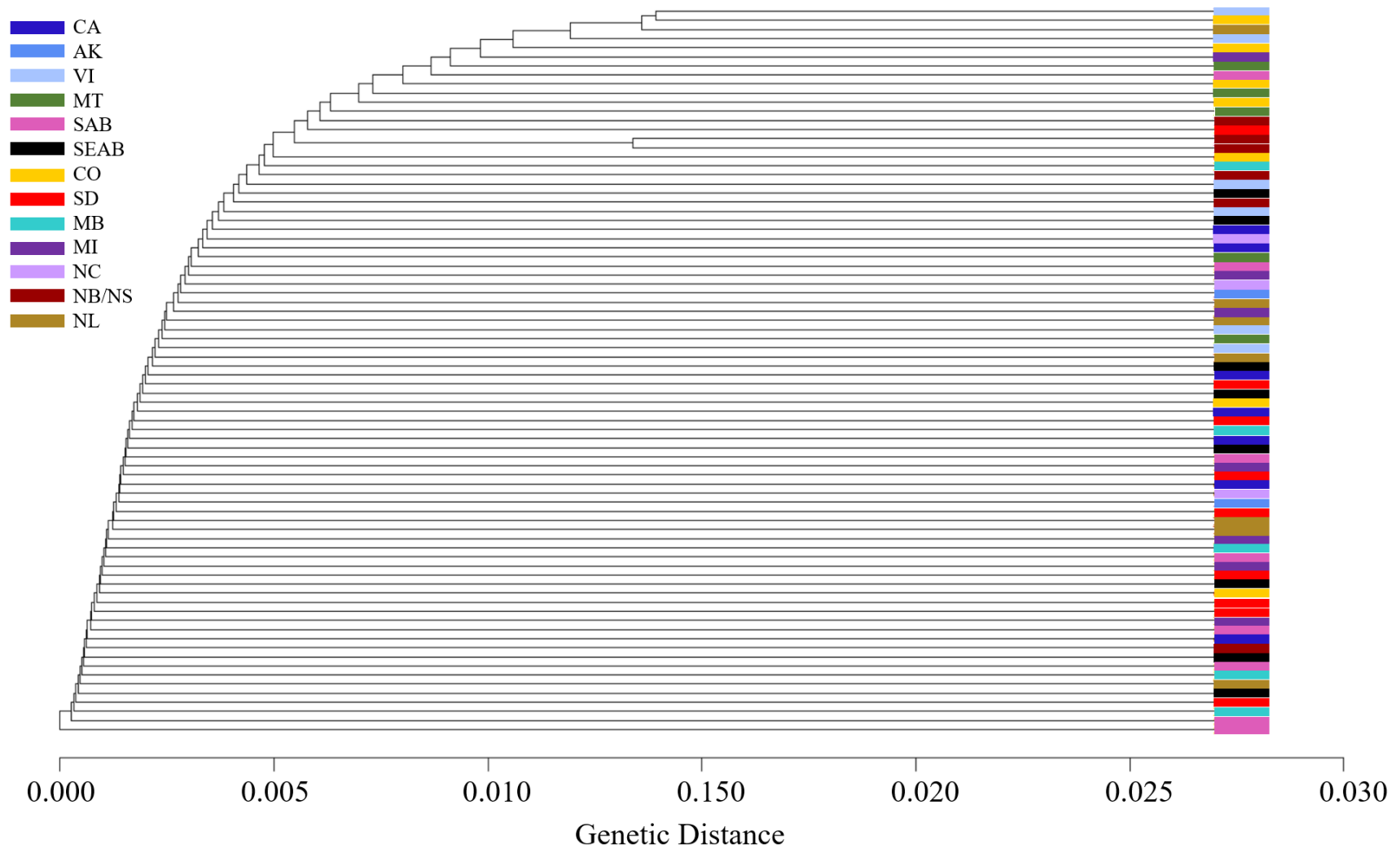


Figure 3.5 UPGMA tree of 80 red-breasted nuthatches individuals using genetic distances from RADseq data. Each node has a minimum of 100 bootstrap support.

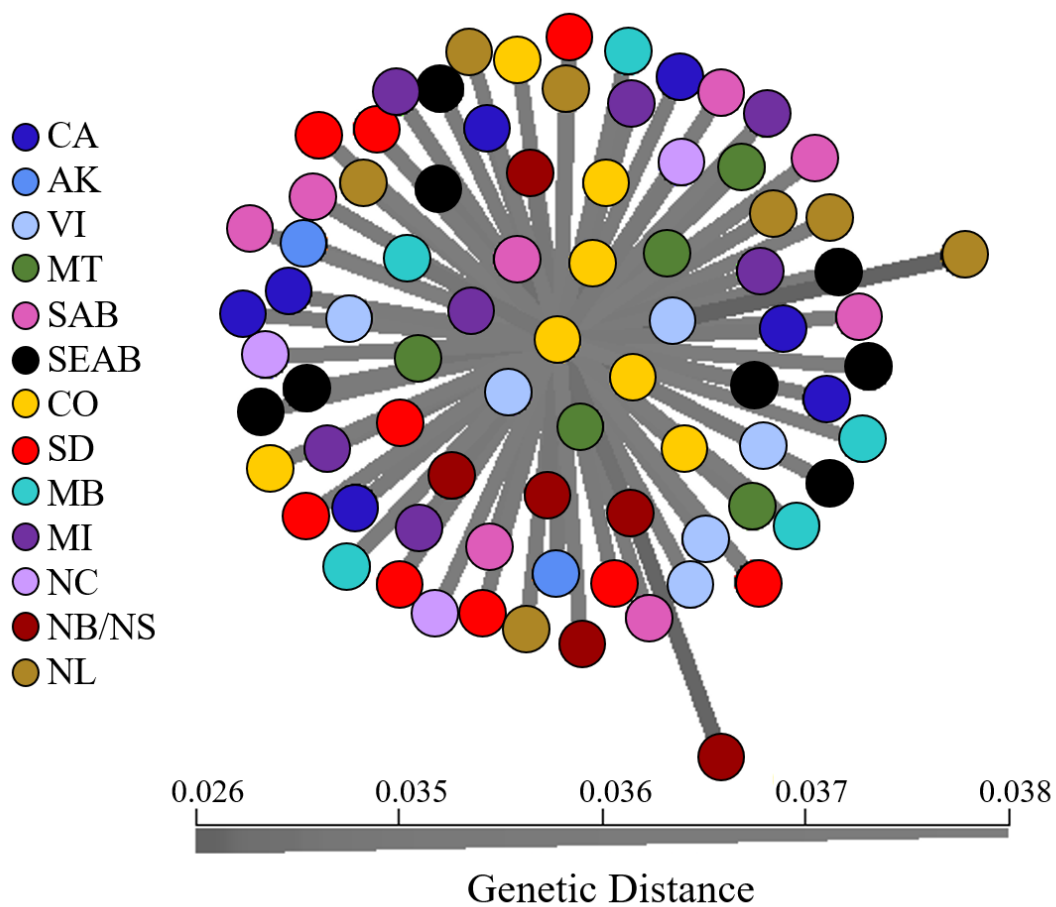


Figure 3.6 Minimum spanning network (MSN) created using pairwise genetic distances to describe the genetic similarity among 80 red-breasted nuthatches from 13 populations using SNPs from RADseq.

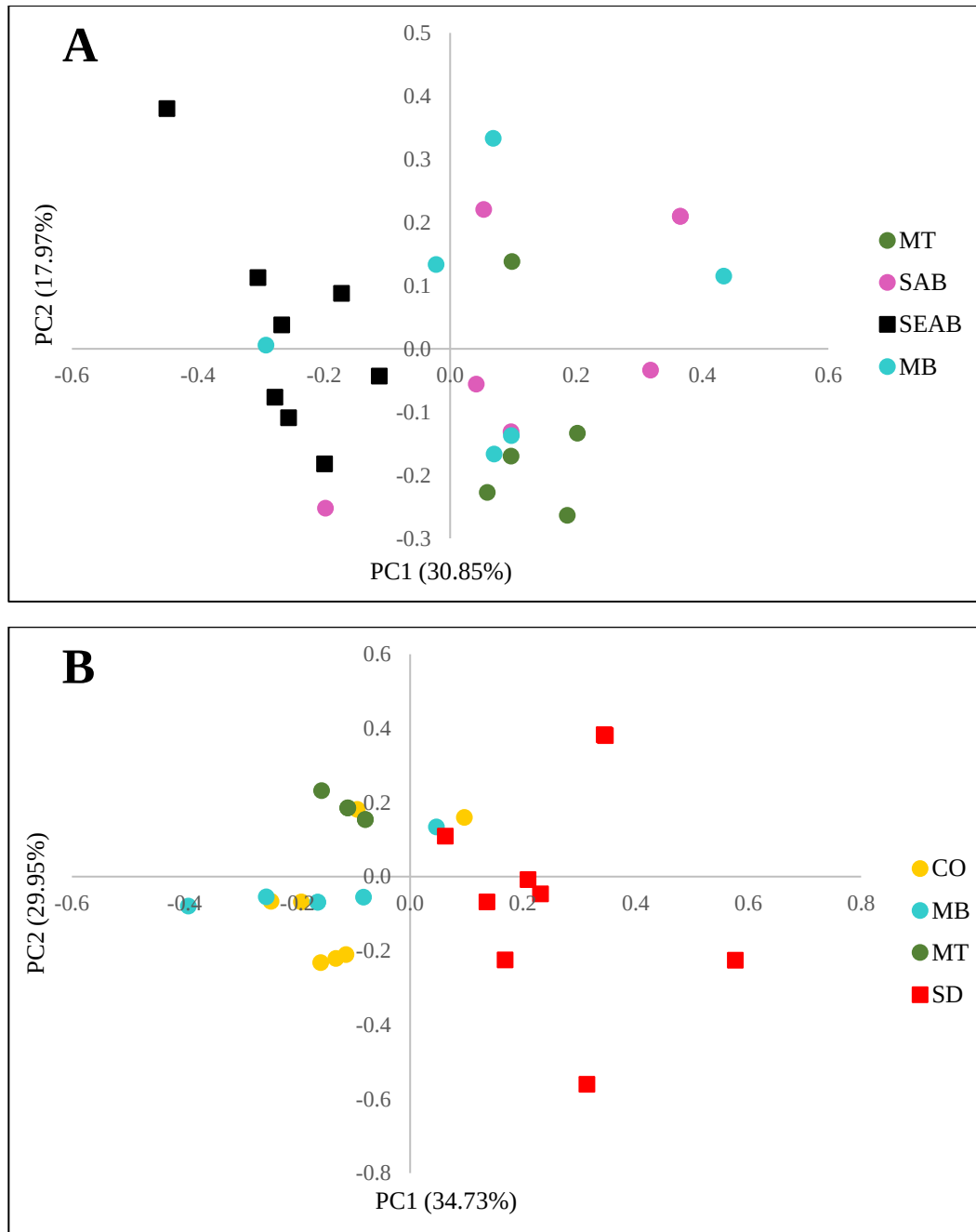


Figure 3.7 PCoAs using SNPs from potential outlier windows. A) Cypress Hills (SEAB) compared to MT, SAB and MB using 11 outlier SNPs. B) Black Hills (SD) compared to the CO, MB and MT populations using 13 outlier SNPs.

3.6 Literature Cited

- Adams, R. V., and Burg, T. M. (2015). Gene flow of a forest-dependent bird across a fragmented landscape. *PLoS One*, 10, 1-22.
- Alvarez, S., Salter, J. F., McCormack, J. E., and Milá, B. (2016). Speciation in mountain refugia: Phylogeography and demographic history of the pine siskin and black-capped siskin complex. *Journal of Avian Biology*, 47(3), 335–345.
- Anderson, C., Spellman, G., Ferrell, C., Strickler, K., and Sarver, S. (2008). Conservation genetics of American dipper (*Cinclus mexicanus*): The genetic status of a population in severe decline. *Conservation Genetics*, 9, 939–944.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
- Bhatia, G., Patterson, N., Sankararaman, S., and Price, A. L. (2013). Estimating and interpreting F_{ST} : The impact of rare variants. *Genome Research*, 23(9), 1514-1521.
- Burg, T. M., Gaston, A. J., Winker, K., and Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, 15(9), 2409–2419.
- Burg, T. M., Taylor, S., Lemmen, K., Gaston, T., and Friesen, V. (2014). Postglacial population genetic differentiation potentially facilitated by a flexible migratory strategy in golden-crowned kinglets (*Regulus satrapa*). *Canadian Journal of Zoology*, 92(2), 163-172.
- Carlson, C. S., Thomas, D. J., Eberle, M. A., Swanson, J. E., Livingston, R. J., Rieder, M. J., and Nickerson, D. A. (2005). Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Research*, 15(11), 1553–1565.
- Carpenter, A. M., Graham, B. A., Spellman, G. M., Klicka, J., and Burg, T. M. (2022). Genetic, bioacoustic and morphological analyses reveal cryptic speciation in the warbling vireo complex (*Vireo gilvus*: Vireonidae: Passeriformes). *Zoological Journal of the Linnean Society* 195(1), 45-64.
- Cannings, R. A., Cannings, R. J., and Cannings S. G. (1987). *Birds of the Okanagan Valley, British Columbia*. Royal British Columbia Museum, Canada.

- Colbeck, G. J., Gibbs, H. L., Marra, P. P., Hobson, K., and Webster, M. S. (2008). Phylogeography of a widespread North American migratory songbird (*Setophaga ruticilla*). *Journal of Heredity*, 99(5), 453–463.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008.
- Daniels, S. J., and Walters, J. R. (2000). Inbreeding depression and its effects on natal dispersal in red-cockaded woodpeckers. *The Condor*, 102, 482-491.
- Dar, M. A., Ahmad, S. M., Bhat, B. A., Dar, T. A., Haq, Z., Wani, B. A., Shabir, N., Kashoo, Z. A., Shah, R. A., Ganai, N. A., and Heidari, M. (2022). Comparative RNA-Seq analysis reveals insights in *Salmonella* disease resistance of chicken; and database development as resource for gene expression in poultry. *Genomics*, 114(5), 110475.
- Davis, L. A., Roalson, E. H., Cornell, K. L., Mcclanahan, K. D., and Webster, M. S. (2006). Genetic divergence and migration patterns in a North American passerine bird: Implications for evolution and conservation. *Molecular Ecology*, 15(8), 2141–2152.
- Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., and Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, 9, e54462.
- Dempsey, Z. W., Goater, C. P., and Burg, T. M. (2020). Living on the edge: Comparative phylogeography and phylogenetics of *Oreohelix* land snails at their range edge in Western Canada. *BMC Evolutionary Biology*, 20(1), 3.
- Dorn, J. L., and Dorn, R. D. (1999) *Wyoming Birds Second Edition*. Mountain West Publishing, Wyoming.
- Drovetski, S. V., Zink, R. M., Ericson, P. G. P., and Fadeev, I. V. (2010). A multilocus study of pine grosbeak phylogeography supports the pattern of greater intercontinental divergence in Holarctic boreal forest birds than in birds inhabiting other high-latitude habitats. *Journal of Biogeography*, 37(4), 696–706.
- Dunn, E. H. (2019). Dynamics and population consequences of irruption in the red-breasted nuthatch (*Sitta canadensis*). *The Auk*, 136(2), 1-10.

- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567.
- Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A. H., Xie, D., Chen, G., Guo, C., Faircloth, B. C., Petersen, B., Wang, Z., Zhou, Q., Diekhans, M., Chen, W., Andreu-Sánchez, S., Margaryan, A., Howard, J. T., Parent, C., ... Zhang, G. (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature*, 587(7833), 252–257.
- García-Ramos, G., and Kirkpatrick, M. (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, 51(1), 21–28.
- Garroway, C. J., Radersma, R., Sepil, I., Santure, A. W., De Cauwer, I., Slate, J., and Sheldon, B. C. (2013). Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. *Evolution*, 67(12), 3488–3500.
- Ghalambor, C. K., and Martin, T. E. (2020). Red-breasted nuthatch (*Sitta canadensis*), version 1.0. *Birds of the World*.
- Gibson, S. Y., Van Der Marel, R. C., and Starzomski, B. M. (2009). Climate change and conservation of leading-edge peripheral populations. *Conservation Biology*, 23(6), 1369–1373.
- Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., and Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme-associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, 15(6), 1304–1315.
- Haas, S. E., Smith, J., Kimball, R., and Clark, A. (2009). Isolation and characterization of polymorphic microsatellite markers for the brown-headed nuthatch (*Sitta pusilla*). *Conservation Genetics*, 10, 1393–1395.
- Haas, S. E., Cox, J. A., Smith, J. V., and Kimball, R. T. (2010). Fine-scale spatial genetic structure in the cooperatively breeding brown-headed nuthatch (*Sitta pusilla*). *Southeastern Naturalist*, 9(4), 743–756.
- Hampe, A., and Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, 8(5), 461–467.

- Harrap, S., and D. Quinn. (1995). *Chickadees, tits, nuthatches and treecreepers*. Princeton University Press, Princeton.
- Hartl, D. L., and Clark, A. G. (1997). *Principles of population genetics*. Sinauer Associates, Sunderland.
- Hochachka, W. M., Wells, J. V., Rosenberg, K. V., and Tessaglia-Hymes, D. L. (1999). Irruptive migration of common redpolls. *The Condor*, 101(2), 195–204.
- Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., and Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports*, 7(1), 17598.
- Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9(5), 1322–1332.
- Jeffries, D. L., Copp, G. H., Lawson Handley, L., Olsén, K. H., Sayer, C. D., and Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, 25(13), 2997–3018.
- JiaQing, W., ZhenGang, L., Jia, B., DaiZong, L., Shuang, M., and GuangNing, N. (2013). Cloning and function prediction of PDCD6 full-length cDNA of *Oncorhynchus mykiss*. *Journal of Shanghai Jiaotong University - Agricultural Science*, 31(6), 30–36.
- Jombart, T. (2008). Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281.
- Kane, E., and Giaid, A. (2011). Potential clinical implications of the urotensin II receptor antagonists. *Frontiers in Pharmacology*, 2(38), 1–11.
- Kiesow, A. M., Monroe, E. M., and Britten, H. B. (2012). Genetic structure of the arboreal squirrels (*Glaucomys sabrinus* and *Tamiasciurus hudsonicus*) in the North American Black Hills. *Canadian Journal of Zoology*, 90(9), 1191–1200.

- Kirby, J. S., Stattersfield, A. J., Butchart, S. H. M., Evans, M. I., Grimmett, R. F. A., Jones, V. R., O'Sullivan, J., Tucker, G. M., and Newton, I. (2008). Key conservation issues for migratory land and waterbird species on the world's major flyways. *Bird Conservation International*, 18(1), S49–S73.
- Knowles, L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, 10, 691–701.
- Komolka, K., Albrecht, E., Gotoh, T., and Maak, S. (2017). Abundance of beige and brown adipocyte markers in different adipose depots of cattle at 26 months of age. *Advances in Animal Biosciences*, 8(S1), S38–S41.
- Lammi, A., Siikamäki, P., and Mustajärvi, K. (1999). Genetic diversity, population size, and fitness in central and peripheral populations of a rare plant *Lychnis viscaria*. *Conservation Biology*, 13(5), 1069–1078.
- Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M., Koskiniemi, J., and Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness—Implications for brown trout conservation. *Ecology and Evolution*, 9(4), 2106–2120.
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595.
- Lou, R. N., Jacobs, A., Wilder, A. P., and Therikildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, 30(23), 5966–5993.
- Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., and Bensch, S. (2017). Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. *Evolution Letters*, 1(3), 155–168.
- MacDonald, Z. G., Dupuis, J. R., Davis, C. S., Acorn, S. E., and Nielsen, S. E. (2020). Gene flow and climate-associated genetic variation in a vagile habitat specialist. *Molecular Ecology*, 29(20), 3889–3906.

- Macfarlane, C. B. A., Natola, L., Brown, M. W., and Burg, T. M. (2016). Population genetic isolation and limited connectivity in the purple finch (*Haemorhous purpureus*). *Ecology and Evolution*, 6(22), 8304–8317.
- Manthey, J. D., and Moyle, R. G. (2015). Isolation by environment in white-breasted nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology*, 24(14), 3628–3638.
- Marthinsen, G., Wennerberg, L., and Lifjeld, J. T. (2008). Low support for separate species within the redpoll complex (*Carduelis flammea–hornemanni–cabaret*) from analyses of mtDNA and microsatellite markers. *Molecular Phylogenetics and Evolution*, 47(3), 1005–1017.
- Marthinsen, G., Wennerberg, L., Solheim, R., and Lifjeld, J. T. (2009). No phylogeographic structure in the circumpolar snowy owl (*Bubo scandiacus*). *Conservation Genetics*, 10(4), 923–933.
- Mendelsohn, B., Bedrosian, B., Love Stowell, S. M., Gagne, R. B., LaCava, M. E. F., Godwin, B. L., Hull, J. M., and Ernest, H. B. (2020). Population genomic diversity and structure at the discontinuous southern range of the great gray owl in North America. *Conservation Genetics*, 21(4), 693–706.
- Milá, B., Girman, D. J., Kimura, M., and Smith, T. B. (2000). Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society B: Biological Sciences*, 267(1447), 1033–1040.
- Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16(3), 1215.
- Mukohda, M., Ibeawuchi, S.-R. C., Hu, C., Lu, K.-T., Nair, A. R., Agbor, L. N., Wu, J., Quelle, F. W., and Sigmund, C. D. (2017). RhoBTB1 is a novel gene protecting against hypertension. *The FASEB Journal*, 31(S1), 1034.
- Otter, K., Ratcliffe, L., Michaud, D., and Boag, P. T. (1998). Do female black-capped chickadees prefer high-ranking males as extra-pair partners? *Behavioral Ecology and Sociobiology*, 43(1), 25–36.

- Nathan, L. R., Kanno, Y., and Vokoun, J. C. (2017). Population demographics influence genetic responses to fragmentation: A demogenetic assessment of the ‘one migrant per generation’ rule of thumb. *Biological Conservation*, 210, 261–272.
- Pasquet, E., Barker, F. K., Martens, J., Tillier, A., Cruaud, C., and Cibois, A. (2014). Evolution within the nuthatches (Sittidae: Aves, Passeriformes): molecular phylogeny, biogeography, and ecological perspectives. *Journal of Ornithology*, 155(3), 755–765.
- Peakall, R., and Smouse, P. E. (2006). Genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295.
- Peltier, W. (1994). Ice Age Paleotopography. *Science*, 265, 195–201.
- Perrin, A., Khimoun, A., Faivre, B., Ollivier, A., Pracontal, N., Théron, F., Loubon, M., Leblond, G., Duron, O., and Garnier, S. (2020). Habitat fragmentation differentially shapes neutral and immune gene variation in a tropical bird species. *Heredity*, 126, 148–162.
- Pielou, E. C. (1992). *After the Ice Age: The Return of Life to Glaciated North America* (1st edition). University of Chicago Press.
- Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., and Lareu, M. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, 4(98), 1–13.
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Ralston, J., FitzGerald, A. M., Burg, T. M., Starkloff, N. C., Warkentin, I. G., and Kirchman, J. J. (2021). Comparative phylogeographic analysis suggests a shared history among eastern North American boreal forest birds. *Ornithology*, 138(3), ukab018.
- Ralston, J., and Kirchman, J. J. (2013). Predicted range shifts in North American boreal forest birds and the effect of climate change on genetic diversity in blackpoll warblers (*Setophaga striata*). *Conservation Genetics*, 14(2), 543–555.

- Ramasamy, R., Ramasamy, S., Bindroo, B., and Naik, V. (2014). STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus*, 3, 431.
- Ravinet, M., Elgvin, T. O., Trier, C., Aliabadian, M., Gavrilov, A., and Sætre, G.-P. (2018). Signatures of human-commensalism in the house sparrow genome. *Proceedings of the Royal Society B: Biological Sciences*, 285(1884), 20181246.
- Rossum, F. V., Vekemans, X., Gratia, E., and Meerts, P. (2003). A comparative study of allozyme variation of peripheral and central populations of *Silene nutans* L. (Caryophyllaceae) from Western Europe: implications for conservation. *Plant Systematics and Evolution*, 242, 49-61.
- Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., DeSante, D. F., Moore, F., and Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23(23), 5726–5739.
- Saad, K., Theis, S., Otto, A., Luke, G., and Patel, K. (2017). Detailed expression profile of the six Glypicans and their modifying enzyme, Notum during chick limb and feather development. *Gene*, 610, 71–79.
- Saladin, V., Bonfils, D., Binz, T., and Richner, H. (2003). Isolation and characterization of 16 microsatellite loci in the great tit *Parus major*. *Molecular Ecology Notes*, 3(4), 520–522.
- Schöneberg, T., Meister, J., Knierim, A. B., and Schulz, A. (2018). The G protein-coupled receptor GPR34 – The past 20 years of a grownup. *Pharmacology and Therapeutics*, 189, 71–88.
- Sibley, D. A. (2014). *The Sibley Guide to Birds*. Second edition. Alfred A. Knopf, New York.
- Sotome, R., Hirasawa, A., Kikusato, M., Amo, T., Furukawa, K., Kuriyagawa, A., Watanabe, K., Collin, A., Shirakawa, H., Hirakawa, R., Tanitaka, Y., Takahashi, H., Wu, G., Nochi, T., Shimmura, T., Warden, C. H., and Toyomizu, M. (2021). In vivo emergence of beige-like fat in chickens as physiological adaptation to cold environments. *Amino Acids*, 53(3), 381–393.

- Spellman, G. M., and Klicka, J. (2007). Phylogeography of the white-breasted nuthatch (*Sitta carolinensis*): Diversification in North American pine and oak woodlands. *Molecular Ecology*, 16(8), 1729–1740.
- Strong, C., Zuckerberg, B., Betancourt, J. L., and Koenig, W. D. (2015). Climatic dipoles drive two principal modes of North American boreal bird irruption. *Proceedings of the National Academy of Sciences*, 112(21), E2795–E2802.
- Sunde, J., Yildirim, Y., Tibblin, P., and Forsman, A. (2020). Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for Pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics*, 11(218), 1-17.
- Tallmon, D. A., Luikart, G., and Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, 19(9), 489–496.
- Town, L., McGlinn, E., Davidson, T.-L., Browne, C. M., Chawengsaksophak, K., Koopman, P., Richman, J. M., and Wicking, C. (2011). Tmem26 is dynamically expressed during palate and limb development but is not required for embryonic survival. *PLOS ONE*, 6(9), e25228.
- Vendrami, D. L. J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M. S., Leese, F., McMinn, C., Moore, H., and Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: Implications for understanding phenotypic plasticity. *Royal Society Open Science*, 4(2), 160548.
- Walsh, P. S., Metzger, D. A., and Higuchi, R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, 10(4), 506–513.
- Walstrom, V. W., Klicka, J., and Spellman, G. M. (2012). Speciation in the white-breasted nuthatch (*Sitta carolinensis*): A multilocus perspective. *Molecular Ecology*, 21(4), 907–920.
- Wang, X., Lu, B., Shao, L., Li, Z., Ali, A., Yu, F., Fu, Z., and Sun, F. (2022). Genome-wide SNPs reveal the fine-scale population structure of *Laodelphax striatellus* in

- China using double-digest restriction site-associated DNA sequencing. *Genomics*, 114(3), 110329.
- Weir, J., and Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings. Biological Sciences/The Royal Society*, 271, 1881–1887.
- Willing, E.-M., Dreyer, C., and van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLOS ONE*, 7(8), e42649.
- Wu, Y., Liu, X., Zheng, H., Zhu, H., Mai, W., Huang, X., and Huang, Y. (2020). Multiple roles of sFRP2 in cardiac development and cardiovascular disease. *International Journal of Biological Sciences*, 16(5), 730–738.
- Xu, Z., Ji, C., Zhang, Y., Zhang, Z., Nie, Q., Xu, J., Zhang, D., and Zhang, X. (2016). Combination analysis of genome-wide association and transcriptome sequencing of residual feed intake in quality chickens. *BMC Genomics*, 17(1), 594.
- Zimmerman, S. J., Aldridge, C. L., and Oyler-McCance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics*, 21(382), 1-16.

Chapter 4: General Discussion

4.1 Introduction and main findings

Different species can show unique genetic patterns based historical and contemporary barriers to gene flow. Even species with a similar range can be affected in different ways by these physical, behavioural, or environmental barriers. It is important to evaluate the genetic patterns in different species to help shape conservation efforts. Genetically distinct populations may need to be considered as separate conservation units rather than protecting the species as one homogenous entity. My thesis provides an excellent example of how two species with some similar traits can show different genetic patterns. Both study species are small passerines with migratory populations and are widespread throughout North America. I was able to examine populations from both species to study their population genetics and investigate potential barriers or facilitators to gene flow.

The Pleistocene glaciation has been attributed to many species' population structure (Burg et al., 2006; Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004). The impact of residing in different glacial refugia is critical to consider when examining population differentiation (Peltier, 1994; Pielou, 1991; Weir and Schluter, 2004). In my study species, population structure can suggest the number and potentially the location of refugia they resided in during the LGM. Golden-crowned kinglets show the signature east west split of species that were in two glacial refugia; one in the east and one in the west. This is seen in a number of other species, including non avian taxa (Burg et al., 2014; Knowles, 2001; Macfarlane et al., 2016; Weir and Schluter, 2004). Red-breasted nuthatches, however, do not show any patterns that would strongly

suggest they were in more than one glacial refugium. There is weak evidence to suggest they may have been in a glacial refugium in Atlantic Canada in addition to a mainland one, however there is stronger evidence to indicate that they expanded rapidly out of a one refugium. One of the factors that may have influenced these patterns is that golden-crowned kinglets prefer mature conifer forests, while red-breasted nuthatches are more adept to living in a variety of forest habitats (Ghalambor and Martin, 2020; Swanson et al., 2020). This would cause golden-crowned kinglets to be more restricted to certain areas during the LGM when forest and plant landscapes were different than present day. The impacts of the LGM can often explain a large portion of the genomic landscapes in high latitude species.

Differences in migratory behaviour can also explain some of the genetic patterns found in my study species. Golden-crowned kinglets have a large expanse of migratory populations in the center of their range. This may limit gene flow between the east and west resident populations. Resident populations do not have to travel prior to the breeding season, suggesting they may be able to breed slightly earlier than their migratory counterparts. Different breeding times could prevent some genetic exchange between migratory and resident populations. On the other hand, red-breasted nuthatches show significant amounts of gene flow between populations. This may be facilitated by migratory behaviour, as large numbers irrupt south every two to five years (Cannings et al., 1987; Dorn and Dorn, 1990). Migratory behaviour is typically thought to enhance gene flow between populations, so if individuals are migrating every few years, then it is not surprising that there is significant gene flow between populations (Delmore et al., 2020; Ruegg et al., 2014). In addition to this irruptive migration, the northernmost

populations appear to migrate every year (Harrap and Quinn, 1995). There is little research that indicates how far these populations migrate and their site fidelity, which may be important for understanding these migratory populations (Dunn, 2019). While migration is seen in many taxa, nearly one fifth of the world's bird species migrate in some way, making it critical to understand when examining the population structure of different bird species.

I also investigated the role of sky islands in population differentiation, specifically in red-breasted nuthatches as I had samples from two different sky island populations. Sky islands are unique habitats that can be home to genetically unique populations of a variety of taxa including flying squirrels and snails (Dempsey et al., 2020; Kiesow et al., 2012). While I did not find that red-breasted nuthatches residing in sky islands were strongly separated from the rest of populations, I did find certain genotypes are more prevalent in the sky islands compared to the surrounding populations. This may suggest that there is some local adaptation occurring in the sky island populations due to different selective pressures. Even though other bird species have shown genetic differentiation in sky islands, red-breasted nuthatches in sky islands may be prevented from separating completely due to their high dispersal capabilities (Manthey and Moyle, 2015; Carpenter et al., 2022). The unique nature of sky islands makes them important to consider, as they seem to support genetically unique populations of some species.

In my study I was able to use two types of molecular markers to research red-breasted nuthatch populations in North America which had never been analysed for population structure. Golden-crowned kinglets have been studied using microsatellites in previous research, therefore I added to the current knowledge of the species using

RADseq data (Burg et al., 2014; Graham et al., 2020). Generally, microsatellites and RADseq show similar patterns of population differentiation and structure (Hodel et al., 2017; Jeffries et al., 2016; Lemopoulos et al., 2019; Sunde et al., 2020; Zimmerman et al., 2020). However, RADseq can sometimes provide better resolution due to the sheer number of markers that it provides (Hodel et al., 2017; Jefferies et al., 2016; Sunde et al., 2020; Vendrami et al., 2017). In both species, the broad scale patterns found by both types of markers were consistent. With golden-crowned kinglets, although the strongest population structure remains the east west split, I also found evidence to suggest that two island populations off the west coast are also distinct using RADseq. This contrasts with previous research using microsatellites which only show one of the islands, Haida Gwaii, separating from mainland populations (Burg et al., 2014; Graham et al., 2020). In red-breasted nuthatches, RADseq uncovered a pattern of differentiation that suggests peripheral populations may be somewhat different than the central ones, although this was less apparent in the microsatellite analyses. RADseq also allows investigation of markers that may be under selection, as microsatellites are neutral markers and do not code for traits that selection can act upon. The microsatellites, however, allowed me to analyse 420 red-breasted nuthatches, a number that would not be feasible with RADseq due to sequencing costs. There are trade offs with using various types of molecular markers and using two types of molecular markers has allowed me to conduct a thorough study of the population genetics of these two North American birds.

4.2 Limitations and future directions

While I was able to provide an overview of the population genetics of my two study species, there remains critical gaps that should be prioritized in future research.

There are important sample sites that I was not able to include in my thesis due to covid restrictions during the field season. In golden-crowned kinglets, it is important to sequence individuals from their central migratory populations. The populations I considered migratory in my study were geographically close to resident populations. It is very possible that there are mixed resident and migratory populations, suggesting that I was not able to sequence purely migratory populations. Additionally, some range maps show the resident range in the west extending across the Rocky Mountains into Alberta, indicating that my migratory populations may be resident instead. The central golden-crowned kinglet populations are also important for determining where the east and west populations split, as there could be a third genetic group in the middle of their range. It would also be interesting to sequence golden-crowned kinglet samples from the Cypress Hills sky island as they have been reported there in low numbers. For red-breasted nuthatches, it is necessary to sequence individuals from the northern part of their range using next-generation sequencing because these are annually migrating populations. I was able to include some of these migratory individuals in my microsatellite analyses, however, they had small sample sizes which could prevent them from separating even if the migratory populations are distinct. Additionally, differences in migratory behaviour are often restricted to just a few areas of the genome, meaning differences in migratory behaviour may not be detected using microsatellites (Delmore et al., 2020; Delmore et al., 2016; Lundberg et al., 2017; Parody-Merino et al., 2019; Peterson et al., 2013).

Although my thesis focused on the genetic aspects of population structure and migratory behaviour, it would be useful to incorporate information on the flyways and stopover sites my study species use during migration. Birds migrating with different

directionality or utilizing different flyways have been shown to either be genetically distinct or have areas of the genome that may explain these differences (Gu et al., 2021; Lundberg et al., 2017). With new tracking methods, even birds as small as the golden-crowned kinglet can be tracked during their migration (Williamson and Witt, 2021). Both of my study species would benefit from satellite tracking. In red-breasted nuthatches, it would be particularly useful to track individuals during irruptive years as it has been suggested that irruption increases mortality and that those that do return may not return to the same location in the following breeding season (Dunn, 2019). The limitations of satellite tracking include cost of trackers and the ability to catch and fit a sufficient number of individuals with trackers.

While RADseq and microsatellites can provide useful information on the population structure of a species, they both have limitations when it comes to looking for areas of selection in the genome. Microsatellites can not provide any information regarding selection, and RADseq samples a small, random portion of the genome. This means it is possible to simply miss portions of the genome that are currently experiencing a selective sweep. Utilizing whole-genome sequencing is very informative in looking for signatures of selection. Manhattan plots from genome scans using whole-genome sequencing data show much clearer peaks and valleys than plots generated using RADseq data (Lou et al., 2021). In the future, it will be important to use whole-genome sequencing with these species to confidently identify areas of the genome under selection in different populations. This can point to populations experiencing local adaptation or different selective pressures, potentially suggesting that these populations are in the process of

diverging. In high latitude birds it is important to consider different approaches to ensure adequate conservation measures are implemented.

4.3 Closing statement

My thesis investigated the population genetic structure of red-breasted nuthatches and golden-crowned kinglets. Red-breasted nuthatches show limited population structure throughout their range, likely due to a rapid expansion from a single glacial refugium and ongoing gene flow due to their migratory habits. Golden-crowned kinglets have a distinct east west genetic split from their time in two separate glacial refugia during the LGM. I have identified potentially important gene regions in both species that could contribute to diversification and/or be the result of selective pressures.

4.4 Literature Cited

- Burg, T. M., Gaston, A. J., Winker, K., and Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, 15(9), 2409–2419.
- Burg, T. M., Taylor, S., Lemmen, K., Gaston, T., and Friesen, V. (2014). Postglacial population genetic differentiation potentially facilitated by a flexible migratory strategy in golden-crowned kinglets (*Regulus satrapa*). *Canadian Journal of Zoology*, 92(2), 163–172.
- Cannings, R. A., Cannings, R. J., and Cannings S. G. (1987). *Birds of the Okanagan Valley, British Columbia*. Royal British Columbia Museum, Canada.
- Carpenter, A. M., Graham, B. A., Spellman, G. M., Klicka, J., and Burg, T. M. (2021). Genetic, bioacoustic and morphological analyses reveal cryptic speciation in the warbling vireo complex (*Vireo gilvus*: Vireonidae: Passeriformes). *Zoological Journal of the Linnean Society* 195(1), 45–64.
- Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., and Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, 9, e54462.
- Delmore, K. E., Toews, D. P. L., Germain, R. R., Owens, G. L., and Irwin, D. E. (2016). The genetics of seasonal migration and plumage color. *Current Biology: CB*, 26(16), 2167–2173.
- Dempsey, Z. W., Goater, C. P., and Burg, T. M. (2020). Living on the edge: Comparative phylogeography and phylogenetics of *Oreohelix* land snails at their range edge in Western Canada. *BMC Evolutionary Biology*, 20(1), 3.
- Dorn, J. L., and Dorn, R. D. (1999) *Wyoming Birds Second Edition*. Mountain West Publishing, Wyoming.
- Dunn, E. H. (2019). Dynamics and population consequences of irruption in the red-breasted nuthatch (*Sitta canadensis*). *The Auk*, 136(2), 1–10.
- Ghalambor, C. K., and Martin, T. E. (2020). Red-breasted nuthatch (*Sitta canadensis*), version 1.0. *Birds of the World*.
- Graham, B. A., Carpenter, A. M., Friesen, V. L., and Burg, T. M. (2020). A comparison of neutral genetic differentiation and genetic diversity among migratory and

- resident populations of golden-crowned kinglets (*Regulus satrapa*). *Journal of Ornithology*, 161(2), 509–519.
- Gu, Z., Pan, S., Lin, Z., Hu, L., Dai, X., Chang, J., Xue, Y., Su, H., Long, J., Sun, M., Ganusevich, S., Sokolov, V., Sokolov, A., Pokrovsky, I., Ji, F., Bruford, M. W., Dixon, A., and Zhan, X. (2021). Climate-driven flyway changes and memory-based long-distance migration. *Nature*, 591(7849), 259–264.
- Harrap, S., and D. Quinn. (1995). *Chickadees, tits, nuthatches and treecreepers*. Princeton University Press, Princeton.
- Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., and Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports*, 7(1), 17598.
- Jeffries, D. L., Copp, G. H., Lawson Handley, L., Olsén, K. H., Sayer, C. D., and Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, 25(13), 2997–3018.
- Kiesow, A. M., Monroe, E. M., and Britten, H. B. (2012). Genetic structure of the arboreal squirrels (*Glaucomys sabrinus* and *Tamiasciurus hudsonicus*) in the North American Black Hills. *Canadian Journal of Zoology*, 90(9), 1191–1200.
- Knowles, L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, 10, 691–701.
- Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M., Koskiniemi, J., and Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness—Implications for brown trout conservation. *Ecology and Evolution*, 9(4), 2106–2120.
- Lou, R. N., Jacobs, A., Wilder, A. P., and Therkildsen, N. O. (2021). A beginner’s guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, 30(23), 5966–5993.
- Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., and Bensch, S. (2017). Genetic differences between willow warbler migratory

- phenotypes are few and cluster in large haplotype blocks. *Evolution Letters*, 1(3), 155–168.
- Macfarlane, C. B. A., Natola, L., Brown, M. W., and Burg, T. M. (2016). Population genetic isolation and limited connectivity in the purple finch (*Haemorhous purpureus*). *Ecology and Evolution*, 6(22), 8304–8317.
- Manthey, J. D., and Moyle, R. G. (2015). Isolation by environment in white-breasted nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology*, 24(14), 3628–3638.
- Parody-Merino, A. M., Battley, P. F., Conklin, J. R., and Fidler, J. R. (2019). No evidence for an association between Clock gene allelic variation and migration timing in a long-distance migratory shorebird (*Limosa lapponica baueri*). *Oecologia*, 191(4), 843–859.
- Peltier, W. (1994). Ice Age Paleotopography. *Science*, 265, 195–201.
- Peterson, M. P., Abolins-Abols, M., Atwell, J. W., Rice, R. J., Milá, B., and Ketterson, E. D. (2013). Variation in candidate genes CLOCK and ADCYAP1 does not consistently predict differences in migratory behavior in the songbird genus *Junco*. *F1000Research*, 2, 115.
- Pielou, E. C. (1992). *After the Ice Age: The Return of Life to Glaciated North America* (1st edition). University of Chicago Press.
- Ralston, J., and Kirchman, J. J. (2013). Predicted range shifts in North American boreal forest birds and the effect of climate change on genetic diversity in blackpoll warblers (*Setophaga striata*). *Conservation Genetics*, 14(2), 543–555.
- Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., DeSante, D. F., Moore, F., and Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, 23(23), 5726–5739.
- Sunde, J., Yıldırım, Y., Tibblin, P., and Forsman, A. (2020). Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for Pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics*, 11(218), 1–17.

- Swanson, D. L., Ingold, J. L., and Galati, R. (2020). Golden-crowned kinglet (*Regulus satrapa*), version 1.0. *Birds of the World*. Vendrami 2017
- Weir, J., and Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings. Biological Sciences/The Royal Society*, 271, 1881–1887.
- Williamson, J. L., and Witt, C. C. (2021). A lightweight backpack harness for tracking hummingbirds. *Journal of Avian Biology*, 52(9), e02802.
- Zimmerman, S. J., Aldridge, C. L., and Oyler-McCance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics*, 21(382), 1-16.

Appendix 1: Supplementary information for Chapter 2

Population genomics of golden-crowned kinglets (*Regulus satrapa*) using RADseq

Appendix 1.1 Individual sample details for the golden-crowned kinglet (GCKI) samples included in this study.

Location	Sample ID	Species	Band Number/Museum ID	Latitude	Longitude
Chilkat, AK	AK 002	GCKI	2300-69198		
Juneau, AK	AK 005	GCKI	2310-74952		
Bow Valley, AB	CAB 001	GCKI	2680-74815	51.0848	-115.0915
Pine Grove, AB	CAB 002	GCKI	2680-74814	51.0380	-114.7755
Pine Grove, AB	CAB 003	GCKI	2680-74819	51.0380	-114.7755
Barrier Lake, AB	CAB 004	GCKI	2680-74818	49.5043	115.7239
Haida Gwaii, BC	HG 006	GCKI	1840-71686		
Haida Gwaii, BC	HG 007	GCKI	2050-72286		
Haida Gwaii, BC	HG 012	GCKI	2260-93132		
Haida Gwaii, BC	HG 015	GCKI	2260-93229		
Haida Gwaii, BC	HG 020	GCKI	2300-69210		
Haida Gwaii, BC	HG 023	GCKI	2300-69249		
Fundy National Park, NB	NB 001	GCKI	gcki 1	45.6150	-65.0356
Richard Squires Provincial Park, NL	NL 001	GCKI	gcki 1	49.3472	-57.3353
Eagle Mountain, NL	NL 003	GCKI	gcki 3	49.8372	-57.2575
Eagle Mountain, NL	NL 004	GCKI	gcki 4	49.8372	-57.2575
Olympic National Park, WA	NWWA 001	GCKI	2310-74915	48.0610	-123.5930
Olympic National Park, WA	NWWA 002	GCKI	2310-74916	48.0610	-123.5930
Toronto Area, ON	ON 003	GCKI	1B-615		
Toronto Area, ON	ON 011	GCKI	1B-1068		
Toronto Area, ON	ON 014	GCKI	1B-1315		
Toronto Area, ON	ON 018	GCKI	1B-583		
Cape Lookout, OR	OR 115	GCKI	231074908	45.3400	-123.9920
West Castle, AB	SAB 002	GCKI	2680-74813	49.4529	114.4097
West Castle, AB	SAB 003	GCKI	2680-74817	49.3864	114.3441

West Castle, AB	SAB 004	GCKI	2680-74806	49.3839	114.3473
West Castle, AB	SAB 803	GCKI	2680-74808	49.3839	114.3473
Mount MacPherson, BC	SBC 002	GCKI	2300-69021	50.9900	-118.1900
Mount MacPherson, BC	SBC 003	GCKI	2300-69022	50.9900	-118.1900
Mount MacPherson, BC	SBC 004	GCKI	2300-69023	50.9900	-118.1900
Sandon, BC	SBC 116	GCKI	2300-69035	49.5500	-117.1800
Vancouver Island, BC	VI 005	GCKI	2120-19872	49.8400	-125.1100
Vancouver Island, BC	VI 006	GCKI	2120-19885	49.8400	-125.1100
Vancouver Island, BC	VI 004	GCKI	2120-19873	49.8400	-125.1100
Vancouver Island, BC	VI 007	GCKI	2120-19876	49.8400	-125.1100
Vancouver Island, BC	VI 014	GCKI	2300-69003	49.6800	-124.9800
Vancouver Island, BC	VI 025	GCKI	2300-69019	49.1900	-125.1100
Mount Rainier, WA	WA 001	GCKI	2300-69086	46.7500	-121.8110
Mount Rainier, WA	WA 002	GCKI	2300-69095	46.7500	-121.8110
Mount Rainier, WA	WA 004	GCKI	2310-74901	46.7650	-121.8110

Appendix 2: Supplementary information for Chapter 3

Population genetics of an irruptive migrant (*Sitta canadensis*) using microsatellites and RADseq

Appendix 2.1 Pairwise F_{ST} comparisons using microsatellite data in red-breasted nuthatches. * = $F_{ST} < 0.0001$. Corrected p -values are above the diagonal and F_{ST} values are below. Values highlighted and bolded were considered statistically significant.

	CA	NV	OR	WA	NEWA	VI	CBC	SEBC	ID	UT	MT	SAB	LETH	CAB
CA		0.0964	0.0571	0.0732	0.3285	0.6651	0.2321	0.1697	0.7138	0.1089	0.0625	0.0928	0.3625	0.2946
NV	0.0001		0.5705	0.8518	0.7654	0.0196	0.2964	0.2304	0.5139	0.4115	0.5214	0.1428	0.3742	0.5321
OR	*	*		0.1392	0.8845	0.6220	0.0785	0.3071	0.0411	0.1411	0.1846	0.3321	0.0214	0.2357
WA	*	*	*		0.2839	0.0607	0.1160	0.4071	0.0589	0.1303	0.1851	0.1125	0.5089	0.0660
NEWA	*	*	*	*		0.4423	0.3535	0.0946	0.9375	0.1142	0.0696	0.2750	0.2785	0.1232
VI	*	0.0144	*	0.0001	*		0.2803	0.7652	0.0268	0.7428	0.2025	0.2928	0.5500	0.6227
CBC	*	*	*	*	*	*		0.7822	0.4816	0.3160	0.7138	0.1250	0.3375	0.6410
SEBC	*	0.0053	*	*	*	*	*		0.0176	0.4212	0.3000	0.6988	0.0875	0.5940
ID	*	*	0.0077	*	*	0.0125	*	0.0171		0.6192	0.04821	0.1433	0.5344	0.5160
UT	*	*	*	*	*	*	*	*	*		0.6108	0.6691	0.7661	0.2982
MT	*	*	*	*	*	0.0017	*	*	0.0022	*		0.1285	0.3303	0.3142
SAB	*	*	*	*	*	*	*	*	0.0066	*	*		0.1339	0.3517
LETH	*	0.0028	0.0193	*	*	*	*	*	*	*	*	*		0.3642
CAB	*	*	*	*	*	*	*	*	*	*	*	*	*	
SEAB	*	*	*	*	*	*	*	*	*	*	*	*	0.0153	*
WY	*	*	*	*	0.0442	*	*	*	*	*	*	*	*	*
CO	*	0.0043	*	*	*	*	*	*	0.0086	*	*	*	*	*
SK	0.0009	*	*	0.0140	*	*	*	*	0.0046	*	*	0.0085	*	*
SD	*	*	*	*	*	*	*	0.0021	0.0050	*	*	*	*	*
MB	*	*	0.0016	*	0.0031	*	*	0.0019	0.0074	0.0011	0.0017	*	*	*
MI	*	*	*	*	*	0.0031	*	0.0017	*	*	*	*	*	*
NC	*	*	*	*	*	*	*	*	*	*	*	*	*	*
NS/NB	0.1230	0.0262	0.0074	0.0140	0.0259	0.0277	*	0.0057	0.0004	*	0.0167	0.0262	0.0003	0.0156
NL	*	*	*	*	*	*	*	0.0442	*	*	*	*	*	*

	SEAB	WY	CO	SK	SD	MB	MI	NC	NS/NB	NL
CA	0.2375	0.1267	0.7373	0.3458	0.3607	0.6807	0.2803	0.2875	0.0179	0.3232
NV	0.4975	0.6604	0.0357	0.5789	0.1000	0.5757	0.1446	0.7842	0.0054	0.5214
OR	0.4777	0.2911	0.7898	0.4094	0.1214	0.0125	0.6866	0.7973	0.1227	0.3464
WA	0.2696	0.5428	0.4316	0.0268	0.5943	0.3446	0.2821	0.2143	0.0892	0.2857
NEWA	0.2196	0.0500	0.3178	0.52341	0.5551	0.3333	0.3035	0.1196	0.1053	0.7601
VI	0.7759	0.7925	0.7199	0.5246	0.0821	0.4459	0.0393	0.4295	0.0054	0.4329
CBC	0.5817	0.0839	0.7067	0.4594	0.3726	0.2392	0.7516	0.6214	0.7396	0.2232
SEBC	0.5546	0.1107	0.6635	0.6083	0.2301	0.3473	0.3172	0.6431	0.2320	0.0446
ID	0.6532	0.7720	0.0411	0.2279	0.1421	0.1747	0.5660	0.3089	0.1071	0.4708
UT	0.7495	0.0642	0.3107	0.2892	0.0982	0.0321	0.5017	0.1018	0.3789	0.1375
MT	0.6935	0.5517	0.5340	0.6785	0.6219	0.1491	0.3678	0.6500	0.0474	0.2410
SAB	0.3053	0.2339	0.3125	0.3584	0.5462	0.1575	0.0803	0.3214	0.0089	0.4250
LETH	0.0429	0.5428	0.3357	0.1321	0.4321	0.1627	0.4000	0.3660	0.0241	0.6737
CAB	0.2428	0.0857	0.1464	0.5125	0.0767	0.4053	0.4500	0.3857	0.0303	0.3250
SEAB		0.1482	0.4410	0.2859	0.3564	0.4699	0.7612	0.4018	0.0035	0.3912
WY	*		0.3267	0.6712	0.6988	0.3875	0.3196	0.1035	0.5881	0.1500
CO	*	*		0.2712	0.1844	0.0454	0.5117	0.6922	0.1357	0.5929
SK	0.0026	*	0.0019		0.3482	0.0357	0.5603	0.7274	0.0375	0.6375
SD	*	*	0.0026	*		0.4930	0.7462	0.6510	0.0183	0.6372
MB	*	*	0.0139	0.0072	*		0.3803	0.4303	0.0170	0.4786
MI	*	*	*	*	*	*		0.3339	0.0107	0.3018
NC	*	*	*	*	*	*	*		0.1534	0.1179
NS/NB	0.0416	*	0.0009	0.0167	0.0184	0.0279	0.0282	0.0131		0.4036
NL	*	*	*	*	*	*	*	*	*	

Appendix 2.2 Individual sample details for the red-breasted nuthatches (RBNU) included in this study. An asterisk next to the sample ID indicates the sample was extracted using the salted-out method and sent for RADseq. NF=National Forest.

Location	Sample ID	Species	Band Number/Museum ID	Latitude	Longitude
Mount Hanley, NS	NS 001	RBNU	rbnu 1	44.7536	-65.1311
Mount Hanley, NS	NS 002	RBNU	rbnu 2	44.7536	-65.1311
Middleton, NS	NS 003	RBNU	rbnu 3	44.9614	-65.0669
Alysford, NS	NS 004	RBNU	rbnu 4	45.0292	-64.8375
Alysford, NS	NS 005	RBNU	rbnu 5	45.0292	-64.8375
Terra Nova National Park, NL	NL 008	RBNU	rbnu 1	48.5197	-53.9672
Lynx Creek, AB	SAB 004	RBNU	rbnu 1	49.4719	-114.4625
Lynx Creek, AB	SAB 005	RBNU	rbnu 2	49.4719	-114.4625
Revelstoke, BC	SEBC 001	RBNU	rbnu 1	50.9828	-118.1789
Revelstoke, BC	SEBC 002	RBNU	rbnu 2	51.0066	-118.1913
Revelstoke, BC	SEBC 003	RBNU	rbnu 3	51.0139	-118.2028
Revelstoke, BC	SEBC 004	RBNU	rbnu 4	50.8607	-118.0958
Revelstoke, BC	SEBC 005	RBNU	rbnu 5	50.9706	-118.1752
Anchorage, AK	AK 001	RBNU	rbnu 1	61.5496	-149.5499
Waterton, AB	SAB 006	RBNU	rbnu 3	49.0133	-113.6700
Helena, MT	MT 007	RBNU	rbnu 1	46.5423	-112.1111
Bozeman, MT	MT 008	RBNU	rbnu 2	45.6938	-110.9046
Olds, AB	CAB 001	RBNU	2520-39801	51.7916	-114.2862
Smithers, BC	CBC 011	RBNU	rbnu 1	54.7594	-127.3617
Smithers, BC	CBC 012	RBNU	rbnu 1	54.7594	-127.3617
Mt. Rainier, WA	WA 030	RBNU	rbnu 1	46.7847	-121.7418
Mt. Rainier, WA	WA 031	RBNU	rbnu 2	46.7847	-121.7418
Mt. Rainier, WA	WA 032	RBNU	rbnu 3	46.7835	-121.7495
Seattle Forest Park, WA	WA 033	RBNU	rbnu 4	47.7433	-122.2933

Seattle Forest Park, WA	WA 034	RBNU	rbnu 5	47.7433	-122.2933
Seattle Forest Park, WA	WA 035	RBNU	rbnu 6	47.7433	-122.2933
Seattle Forest Park, WA	WA 036	RBNU	rbnu 7	47.7433	-122.2933
Walton Lake Campground, OR	OR 027	RBNU	rbnu 1	44.9638	-118.2333
Walton Lake Campground, OR	OR 028	RBNU	rbnu 1	44.4674	-119.9353
Walton Lake Campground, OR	OR 029	RBNU	rbnu 2	44.4341	-120.3187
Walton Lake Campground, OR	OR 030	RBNU	rbnu 3	44.4341	-120.3187
Crystal Springs, OR	OR 031	RBNU	rbnu 1	42.5674	-122.0854
Crater Lake, OR	OR 033	RBNU	rbnu 2	42.8840	-122.1354
Crater Lake, OR	OR 034	RBNU	rbnu 3	42.8507	-122.1520
Cache Forest, UT	UT 009	RBNU	RBNU 1	41.4067	-111.5427
Cache Forest, UT	UT 010	RBNU	RBNU 2	41.4544	-111.5040
Thornhill, BC	NBC 001	RBNU	RBNU1	54.5060	-128.5430
Tyee Lake, BC	CBC 013	RBNU	RBNU2	54.7074	-127.0397
Revelstoke, BC	SEBC 006	RBNU	RBNU6	50.9769	-118.2276
Moscow, ID	ID 010	RBNU	RBNU1	46.7741	-116.8622
Quispamsis, NB	NB 006*	RBNU	RBNU6	45.4793	-65.9192
Quispamsis, NB	NB 007*	RBNU	RBNU7	45.4793	-65.9192
Second North River, NB	NB 005*	RBNU	RBNU8	46.0628	-65.0494
Second North River, NB	NB 232	RBNU	RBNU9	46.0628	-65.0494
Musquodoboit, NS	NS 010*	RBNU	RBNU10	45.0347	-63.0846
Musquodoboit, NS	NS 011*	RBNU	RBNU11	45.0036	-63.0140
Prince Albert National Park, SK	SK 001	RBNU	RBNU1	53.9820	-106.2957
Custer State Park, SD	SD 007	RBNU	rbnu1	43.8114	-103.4555
Custer State Park, SD	SD 008	RBNU	rbnu2	43.7989	-103.4402
Custer State Park, SD	SD 009	RBNU	rbnu3	43.7788	-103.4491
Jasper National Park, AB	CAB 002	RBNU	rbnu2	52.8491	-118.0797

Jasper National Park, AB	CAB 003	RBNU	rbnu3	52.9112	-118.0926
Jasper National Park, AB	CAB 004	RBNU	rbnu4	52.9112	-118.0926
Crownest Pass, AB	SAB 006	RBNU	RBNU 4	49.5883	-114.2142
Fort St James, BC	CBC 013	RBNU	RBNU 3	54.4160	-124.2204
Francois Lake, BC	CBC 014	RBNU	RBNU 4	54.0398	-125.9867
Ponderosa State Park, ID	ID 011	RBNU	RBNU 1	44.9411	-116.0805
Ponderosa State Park, ID	ID 012	RBNU	RBNU 2	44.9411	-116.0805
Ponderosa State Park, ID	ID 013	RBNU	RBNU 3	44.9411	-116.0805
Ponderosa State Park, ID	ID 014	RBNU	RBNU 4	44.9411	-116.0805
Ponderosa State Park, ID	ID 015	RBNU	RBNU 5	44.9411	-116.0805
Ponderosa State Park, ID	ID 016	RBNU	RBNU 6	45.0069	-116.1480
Ponderosa State Park, ID	ID 017	RBNU	RBNU 7	44.9819	-116.1128
Little Payette Lake, ID	ID 018	RBNU	RBNU 8	44.9268	-116.0431
Little Payette Lake, ID	ID 019	RBNU	RBNU 9	44.9268	-116.0431
Little Payette Lake, ID	ID 020	RBNU	RBNU 10	44.9268	-116.0431
Revelstoke, BC	SEBC 006	RBNU	RBNU 9	50.9963	-118.3148
Revelstoke, BC	SEBC 007	RBNU	RBNU 10	50.9963	-118.3148
Revelstoke, BC	SEBC 008	RBNU	RBNU 11	50.9963	-118.3148
Revelstoke, BC	SEBC 009	RBNU	RBNU 12	50.9963	-118.3148
Woolford Provincial Park, AB	SAB 006	RBNU	2710-78423	49.1750	-113.1876
Woolford Provincial Park, AB	SAB 007	RBNU	2710-78424	49.1750	-113.1876
Red Deer, AB	SAB 008	RBNU	2710-78443	52.3163	-113.7769
Starvation Lake, WA	NEWA 001	RBNU	2540-22988	48.4928	-117.7091
Starvation Lake, WA	NEWA 002	RBNU	2540-22989	48.4928	-117.7091
North Gorge, WA	NEWA 003	RBNU	2540-22992	48.7865	-118.0013
North Gorge, WA	NEWA 004	RBNU	2540-22993	48.7865	-118.0013
Big Meadow Lake, WA	NEWA 005	RBNU	2540-22998	48.7257	-117.5619

Kelowna, BC	SEBC 010	RBNU	2710-78332	49.8760	-119.4304
Redmond, WA	WA 037	RBNU	2580-49308	47.6556	-121.9956
Lacey, WA	WA 038	RBNU	2580-49317	47.0090	-122.8325
Steven's County, WA	NEWA 006	RBNU	2580-49327	48.5536	-117.5784
Steven's County, WA	NEWA 007	RBNU	2580-49328	48.5536	-117.5784
Steven's County, WA	NEWA 008	RBNU	2580-49330	48.5757	-117.6459
Steven's County, WA	NEWA 009	RBNU	2580-49331	48.5757	-117.6459
Steven's County, WA	NEWA 010	RBNU	2580-49338	48.6113	-117.5538
Steven's County, WA	NEWA 011	RBNU	2580-49339	48.6113	-117.5538
Steven's County, WA	NEWA 012	RBNU	2580-49341	48.6113	-117.5538
Steven's County, WA	NEWA 013	RBNU	2580-49342	48.6113	-117.5538
Steven's County, WA	NEWA 014	RBNU	2580-49344	48.6113	-117.5538
Steven's County, WA	NEWA 015	RBNU	2580-49345	48.6113	-117.5538
Steven's County, WA	NEWA 016	RBNU	2580-49346	48.6113	-117.5538
Helena-Lewis and Clark NF, MT	MT 009	RBNU	2580-49359	46.1953	-112.4476
Helena-Lewis and Clark NF, MT	MT 010	RBNU	2580-49364	46.1953	-112.4476
Helena-Lewis and Clark NF, MT	MT 011	RBNU	2580-49371	46.2833	-112.4738
Helena-Lewis and Clark NF, MT	MT 012	RBNU	2580-49373	46.5407	-112.2483
Helena, MT	MT 013	RBNU	2580-49375	46.5407	-112.2483
Cooper's Lake, MT	MT 014	RBNU	2580-49389	47.0702	-112.9223
Cooper's Lake, MT	MT 015	RBNU	2580-49390	47.0702	-112.9223
Gulliver Lake, MI	MI 004	RBNU	2320-54101	45.9599	-86.0310
Gulliver Lake, MI	MI 005	RBNU	2320-54103	45.9599	-86.0310
Harlow Lake, MI	MI 006	RBNU	2320-54125	46.6324	-87.4862
Harlow Lake, MI	MI 007	RBNU	2320-54127	46.6324	-87.4862
Harlow Lake, MI	MI 008	RBNU	2320-54128	46.6324	-87.4862
Harlow Lake, MI	MI 009	RBNU	2320-54131	46.6324	-87.4862

Harlow Lake, MI	MI 010	RBNU	2320-54132	46.6324	-87.4862
Harlow Lake, MI	MI 011	RBNU	2320-54133	46.6324	-87.4862
Ottawa NF, MI	MI 012	RBNU	2320-54135	46.0772	-88.7584
Ottawa NF, MI	MI 013	RBNU	2320-54142	46.0772	-88.7584
Ottawa NF, MI	MI 014	RBNU	2320-54144	46.1291	-88.7358
Curt's Place, MI	MI 015	RBNU	2320-54150	46.6533	-90.0497
Curt's Place, MI	MI 016	RBNU	2320-54153	46.6533	-90.0497
Curt's Place, MI	MI 017	RBNU	2320-54154	46.6533	-90.0497
Curt's Place, MI	MI 018	RBNU	2320-54155	46.6533	-90.0497
Curt's Place, MI	MI 019	RBNU	2320-54156	46.6533	-90.0497
Ottawa NF, MI	MI 001*	RBNU	2320-54139	46.0772	-88.7584
Ottawa NF, MI	MI 003*	RBNU	2320-54145	46.1291	-88.7358
Ottawa NF, MI	MI 146*	RBNU	2320-54146	46.1291	-88.7358
Ottawa NF, MI	MI 148*	RBNU	2320-54148	46.1291	-88.7358
Osoyoos, BC	SEBC 011	RBNU	RBNU 1	49.0669	-119.5658
Banff, AB	SAB 001*	RBNU	RBNU 1	51.1981	-115.5926
Jasper, AB	CAB 005	RBNU	RBNU 5	52.9133	-118.0951
Manning, BC	SWBC 001	RBNU	RBNU 1	49.1288	-120.6239
West Castle, AB	SAB 009	RBNU	2710-78293	49.3839	-114.3473
West Castle, AB	SAB 294*	RBNU	2710-78294	49.3839	-114.3473
West Castle, AB	SAB 295*	RBNU	2710-78295	49.3839	-114.3473
West Castle, AB	SAB 296*	RBNU	2710-78296	49.3839	-114.3473
West Castle, AB	SAB 011	RBNU	2710-78297	49.4016	-114.3368
West Castle, AB	SAB 298*	RBNU	RBNU12	49.4449	-114.3225
West Castle, AB	SAB 012*	RBNU	2710-78298	49.4016	-114.3368
West Castle, AB	SAB 010	RBNU	2710-78299	49.3864	-114.3441
Flathead NF, MT	MT 016	RBNU	2710-42701	48.3764	-114.5390
Flathead NF, MT	MT 017	RBNU	2710-42702	48.3593	-114.5006

Tally Lake, MT	MT 018	RBNU	2710-42703	48.4144	-114.5832
Columbia Mountain Trail, MT	MT 019	RBNU	2710-42704	48.3816	-114.1148
Columbia Mountain Trail, MT	MT 020	RBNU	2710-42705	48.3816	-114.1148
Custer NF, MT	MT 021	RBNU	2710-42706	45.1520	-109.3468
Basin Campground, MT	MT 022	RBNU	2710-42707	45.1629	-109.3931
Basin Campground, MT	MT 023	RBNU	2710-42708	45.1629	-109.3931
Custer NF, MT	MT 024	RBNU	2710-42709	45.1508	-109.3430
Wild Bill Lake, MT	MT 025	RBNU	2710-42710	45.1551	-109.3694
Cranbrook Community Forest, BC	SEBC 012	RBNU	2710-78617	49.5043	-115.7238
Cranbrook Community Forest, BC	SEBC 013	RBNU	2710-78300	49.5043	-115.7238
Cranbrook Community Forest, BC	SEBC 014	RBNU	2710-78612	49.5043	-115.7238
Cranbrook Community Forest, BC	SEBC 015	RBNU	2710-78613	49.5043	-115.7238
Moyie Lake, BC	SEBC 016	RBNU	2710-78614	49.3736	-115.8301
Cranbrook Community Forest, BC	SEBC 017	RBNU	2710-78615	49.5043	-115.7238
Cranbrook Community Forest, BC	SEBC 018	RBNU	2710-78616	49.5043	-115.7238
Weaselhead Flats, AB	CAB 006	RBNU	2710-78618	50.9949	-114.1571
Kamloops, BC	SEBC 019	RBNU	2710-78619	50.5916	-120.3864
Cypress Hills, AB	SEAB 001*	RBNU	2710-78620	49.6573	-110.1021
Cypress Hills, AB	SEAB 002*	RBNU	2710-78621	49.6573	-110.1021
Cypress Hills, AB	SEAB 003*	RBNU	2710-78622	49.6618	-110.0699
Cypress Hills, AB	SEAB 004	RBNU	2710-78625	49.6601	-110.0830
Cypress Hills, AB	SEAB 005	RBNU	2710-78627	49.8705	-110.0372
Cypress Hills, AB	SEAB 623*	RBNU	2710-78623	49.6618	-110.0699
Cypress Hills, AB	SEAB 624*	RBNU	2710-78624	49.6618	-110.0699
Cypress Hills, AB	SEAB 626*	RBNU	2710-78626	49.8705	-110.0372
Cypress Hills, AB	SEAB 628*	RBNU	2710-78628	49.8705	-110.0372
Cypress Hills, AB	SEAB 629*	RBNU	2710-78629	49.8705	-110.0372

Red Deer, AB	CAB 007	RBNU	2710-78443	52.3083	-113.7628
Olds, AB	CAB 008	RBNU	2520-39801	51.7781	-114.1190
Castle Bridge, AB	SAB 011	RBNU	2710-78297	49.4016	114.3368
Revelstoke, BC	SEBC 020	RBNU	RBNU 7	50.9887	-118.2778
Revelstoke, BC	SEBC 021	RBNU	RBNU 8	50.9887	-118.2778
Manning, BC	SWBC 002	RBNU	RBNU 1	50.9887	-118.2778
Lethbridge, AB	LETH 001	RBNU	2010-78634	49.7007	-112.8591
Lethbridge, AB	LETH 002	RBNU	2010-78635	49.7007	-112.8591
Lethbridge, AB	LETH 003	RBNU	2710-78630	49.7007	-112.8591
Lethbridge, AB	LETH 004	RBNU	2010-78636	49.7007	-112.8591
Lethbridge, AB	LETH 005	RBNU	2010-78639	49.7007	-112.8591
Lethbridge, AB	LETH 006	RBNU	2010-78640	49.7007	-112.8591
Lethbridge, AB	LETH 007	RBNU	2010-78641	49.7007	-112.8591
Lethbridge, AB	LETH 008	RBNU	2010-78642	49.7007	-112.8591
Crownest Pass, AB	SAB 002*	RBNU	RBNU 5	49.5757	-114.2114
Goose Bay State Game Refuge, AK	AK 002*	RBNU	DAB 738	61.2900	-149.9500
Goose Bay State Game Refuge, AK	AK 003*	RBNU	DAB 737	61.2900	-149.9500
Prince George, BC	CBC 001	RBNU	JMD 033	53.9387	-122.7750
Prince George, BC	CBC 002	RBNU	GMS 858	53.9387	-122.7750
Prince George, BC	CBC 003	RBNU	GMS 856	53.9387	-122.7750
Prince George, BC	CBC 004	RBNU	GMS 857	53.9387	-122.7750
Prince George, BC	CBC 005	RBNU	JMD 031	53.9387	-122.7750
Prince George, BC	CBC 006	RBNU	JMD 034	53.9387	-122.7750
Prince George, BC	CBC 007	RBNU	GMS 859	53.9387	-122.7750
Prince George, BC	CBC 008	RBNU	JMD 035	53.9387	-122.7750
Prince George, BC	CBC 009	RBNU	JMD 036	53.9387	-122.7750
Prince George, BC	CBC 010	RBNU	GMS 860	53.9387	-122.7750

Nanaimo, BC	VI 001*	RBNU	JMD 039	48.5550	-124.4090
Nanaimo, BC	VI 002*	RBNU	GMS 853	48.5550	-124.4090
Nanaimo, BC	VI 003*	RBNU	JMD 038	48.5550	-124.4090
Nanaimo, BC	VI 004	RBNU	GMS 854	48.5550	-124.4090
Nanaimo, BC	VI 862*	RBNU	GMS 862	48.5550	-124.4090
Nanaimo, BC	VI 861*	RBNU	GMS 861	48.5550	-124.4090
Nanaimo, BC	VI 040*	RBNU	JMD 040	48.5550	-124.4090
Nanaimo, BC	VI 863*	RBNU	GMS 863	48.5550	-124.4090
Nanaimo, BC	VI 041*	RBNU	JMD 041	48.5550	-124.4090
Nanaimo, BC	VI 864*	RBNU	GMS 864	48.5550	-124.4090
Sequoia NF, CA	CA 001*	RBNU	GMS 623	43.8000	-119.0300
Shasta NF, CA	CA 002*	RBNU	GMS 337	40.5672	-122.3106
Shasta NF, CA	CA 003*	RBNU	GMS 345	40.5672	-122.3106
Shasta NF, CA	CA 004	RBNU	GMS 338	40.5672	-122.3106
Shasta NF, CA	CA 005	RBNU	GMS 340	40.5672	-122.3106
Shasta NF, CA	CA 006	RBNU	GMS 341	40.5672	-122.3106
Shasta NF, CA	CA 007	RBNU	GMS 343	40.5672	-122.3106
Shasta NF, CA	CA 008	RBNU	GMS 336	40.5672	-122.3106
Shasta NF, CA	CA 009	RBNU	RAF 046	40.5672	-122.3106
Shasta NF, CA	CA 010	RBNU	GMS 339	40.5672	-122.3106
Sequoia NF, CA	CA 011	RBNU	MM 041	43.8000	-119.0300
Sequoia NF, CA	CA 012	RBNU	MM 036	43.8000	-119.0300
Sequoia NF, CA	CA 015	RBNU	GMS 624	43.8000	-119.0300
Sequoia NF, CA	CA 013	RBNU	GMS 621	43.8000	-119.0300
Sequoia NF, CA	CA 014	RBNU	GMS 622	43.8000	-119.0300
Sequoia NF, CA	CA 016	RBNU	GMS 625	43.8000	-119.0300
Sequoia NF, CA	CA 017	RBNU	GMS 626	43.8000	-119.0300
Sequoia NF, CA	CA 018	RBNU	JK 03 369	43.8000	-119.0300

Shasta NF, CA	CA 019	RBNU	GMS 342	40.5672	-122.3106
Inyo NF, CA	CA 020	RBNU	GMS 363	37.4377	-119.0057
Inyo NF, CA	CA 021	RBNU	GMS 365	37.4377	-119.0057
Inyo NF, CA	CA 022	RBNU	GMS 368	37.4377	-119.0057
Inyo NF, CA	CA 023	RBNU	GMS 369	37.4377	-119.0057
Inyo NF, CA	CA 024	RBNU	GMS 370	37.4377	-119.0057
Inyo NF, CA	CA 025*	RBNU	GMS 364	37.4377	-119.0057
Shasta NF, CA	CA 026*	RBNU	GMS 344	40.5672	-122.3106
Shasta NF, CA	CA 027*	RBNU	GMS 346	40.5672	-122.3106
Shasta NF, CA	CA 028*	RBNU	DHB 5014	41.1189	-121.5970
Cour d'Alene, ID	ID 001	RBNU	GMS 637	47.5833	-116.5000
Cour d'Alene, ID	ID 002	RBNU	GMS 640	47.5833	-116.5000
Cour d'Alene, ID	ID 003	RBNU	GMS 649	47.5833	-116.5000
Cour d'Alene, ID	ID 004	RBNU	GMS 638	47.5833	-116.5000
Cour d'Alene, ID	ID 005	RBNU	GMS 639	47.5833	-116.5000
Cour d'Alene, ID	ID 006	RBNU	GMS 641	47.5833	-116.5000
Cour d'Alene, ID	ID 007	RBNU	GMS 642	47.5833	-116.5000
Cour d'Alene, ID	ID 008	RBNU	GMS 650	47.5833	-116.5000
Cour d'Alene, ID	ID 009	RBNU	GMS 647	47.5833	-116.5000
Sprague, MB	MB 001*	RBNU	GMS 867	49.2911	-95.1567
Sprague, MB	MB 002*	RBNU	GMS 868	49.2911	-95.1567
Sprague, MB	MB 003*	RBNU	JK 03 658	49.2911	-95.1567
Sprague, MB	MB 004*	RBNU	JK 03 657	49.2911	-95.1567
Sprague, MB	MB 005*	RBNU	GMS 869	49.2911	-95.1567
Sprague, MB	MB 006*	RBNU	GMS 874	49.2911	-95.1567
Sprague, MB	MB 007	RBNU	GMS 870	49.2911	-95.1567
Sprague, MB	MB 008	RBNU	GMS 873	49.2911	-95.1567

Sprague, MB	MB 009	RBNU	GMS 871	49.2911	-95.1567
Sprague, MB	MB 010	RBNU	GMS 872	49.2911	-95.1567
Helena NF, MT	MT 001*	RBNU	DHB 5131	46.4139	-111.4421
Helena NF, MT	MT 002*	RBNU	DHB 5133	46.4139	-111.4421
Helena NF, MT	MT 003*	RBNU	DHB 5132	46.4139	-111.4421
Helena NF, MT	MT 004*	RBNU	DHB 5130	46.4139	-111.4421
Helena NF, MT	MT 005*	RBNU	GMS 422	46.4139	-111.4421
Helena NF, MT	MT 006*	RBNU	GMS 423	46.4139	-111.4421
Helena NF, MT	MT 007	RBNU	GMS 420	46.4139	-111.4421
Helena NF, MT	MT 008	RBNU	GMS 421	46.4139	-111.4421
Mohawk Canyon, NV	NV 001	RBNU	MM 051	39.1590	-115.3310
Mount Moriah, NV	NV 002	RBNU	GMS 427	39.1900	-114.1540
Mount Moriah, NV	NV 003	RBNU	GMS 431	39.1900	-114.1540
Mount Moriah, NV	NV 004	RBNU	GMS 432	39.1900	-114.1540
Mount Moriah, NV	NV 005	RBNU	GMS 424	39.1900	-114.1540
Mount Moriah, NV	NV 006	RBNU	GMS 430	39.1900	-114.1540
Mount Moriah, NV	NV 007	RBNU	GMS 428	39.1900	-114.1540
Mount Moriah, NV	NV 008	RBNU	GMS 429	39.1900	-114.1540
Mohawk Canyon, NV	NV 009	RBNU	MM 048	39.1590	-115.3310
Avalon Peninsula, NL	NL 001*	RBNU	EVL 457	47.4095	-53.3826
Deer Lake, NL	NL 002*	RBNU	GMS 881	49.4700	-57.3170
Deer Lake, NL	NL 003*	RBNU	GMS 882	49.4700	-57.3170
Deer Lake, NL	NL 004*	RBNU	MM 087	49.4700	-57.3170
Deer Lake, NL	NL 005*	RBNU	JK 04 001	49.4700	-57.3170
Avalon Peninsula, NL	NL 006*	RBNU	EVL 459	47.4095	-53.3826
Burnsville, NL	NL 007*	RBNU	WLK 272	48.7634	-56.0962
Cane River Gap, NC	NC 001*	RBNU	RBB 553	35.8098	-82.3483
Walker Knob, NC	NC 002*	RBNU	RBB 535	35.7480	-82.3470

Walker Knob, NC	NC 003*	RBNU	RBB 537	35.7480	-82.3470
Cane River Gap, NC	NC 004*	RBNU	RBB 551	35.8000	-82.3400
Reinhart Gap, NC	NC 005*	RBNU	RBB 549	35.3390	-82.9750
Rough Butt Bald, NC	NC 006*	RBNU	RBB 544	35.3160	-82.9680
Rough Butt Bald, NC	NC 007	RBNU	RBB 545	35.3160	-82.9680
Deschutes NF, OR	OR 004	RBNU	GMS 362	43.4462	-121.3566
Deschutes NF, OR	OR 005	RBNU	GMS 361	43.4462	-121.3566
Deschutes NF, OR	OR 006	RBNU	GMS 360	43.4462	-121.3566
Deschutes NF, OR	OR 007	RBNU	GMS 359	43.4462	-121.3566
Prairie Farm Springs, OR	OR 008	RBNU	GHL 114	44.5306	-121.5469
Selma, OR	OR 009	RBNU	AJK 683	42.3895	-123.6347
Prairie Farm Springs, OR	OR 010	RBNU	CAF 029	44.5306	-121.5469
Deschutes NF, OR	OR 011	RBNU	DHB 4980	43.5236	-121.2165
Deschutes NF, OR	OR 012	RBNU	DHB 4984	43.5236	-121.2165
Whitman NF, OR	OR 013	RBNU	GMS 416	45.2086	-118.2019
Whitman NF, OR	OR 014	RBNU	GMS 415	45.2086	-118.2019
Prairie Farm Springs, OR	OR 015	RBNU	MLD 186	44.5306	-121.5469
Wallowa NF, OR	OR 016	RBNU	GMS 373	45.4661	-117.1727
Whitman NF, OR	OR 017	RBNU	GMS 417	45.2086	-118.2019
Wallowa NF, OR	OR 018	RBNU	GMS 371	45.4661	-117.1727
Wallowa-Whitman NF, OR	OR 019	RBNU	JK 06 671	45.2733	-117.0902
Wallowa-Whitman NF, OR	OR 020	RBNU	JK 09 007	45.2733	-117.0902
Green Ridge, OR	OR 021	RBNU	NAM 74	44.6000	-121.5667
Wallowa-Whitman NF, OR	OR 022	RBNU	JK 06 670	45.2733	-117.0902
Davis Mountain, OR	OR 023	RBNU	PJG 198	43.6487	-121.7630
Wallowa NF, OR	OR 024	RBNU	GMS 372	45.4661	-117.1727
Black Hills NF, SD	SD 001*	RBNU	GMS 395	43.4812	-103.4720

Black Hills NF, SD	SD 002*	RBNU	DHB 4956	43.4326	-103.3423
Black Hills NF, SD	SD 003*	RBNU	DHB 4966	43.4326	-103.3423
Spearfish, SD	SD 004	RBNU	DMNS 45524	44.4833	-103.9333
Spearfish, SD	SD 005	RBNU	DMNS 45526	44.4833	-103.9333
Spearfish, SD	SD 006	RBNU	DMNS 45489	44.4833	-103.9333
Black Hills NF, SD	SD 4965*	RBNU	DHB 4965	43.4326	-103.3423
Black Hills NF, SD	SD 4957*	RBNU	DHB 4957	43.4326	-103.3423
Black Hills NF, SD	SD 4958*	RBNU	DHB 4958	43.4326	-103.3423
Black Hills NF, SD	SD 4959*	RBNU	DHB 4959	43.4326	-103.3423
Black Hills NF, SD	SD 4960*	RBNU	DHB 4960	43.4326	-103.3423
Black Hills NF, SD	SD 394*	RBNU	GMS 394	43.4812	-103.4720
Black Mountain, UT	UT 001	RBNU	MM 063	41.2650	-111.5770
Black Mountain, UT	UT 002	RBNU	MM 046	41.2650	-111.5770
Three Mile Creek Canyon, UT	UT 003	RBNU	GMS 845	41.2650	-111.5770
Public Grove Canyon, UT	UT 004	RBNU	MM 056	41.2680	-111.5160
Public Grove Canyon, UT	UT 005	RBNU	JMD 025	41.2560	-111.5200
Public Grove Canyon, UT	UT 006	RBNU	MM 064	41.2560	-111.5200
Public Grove Canyon, UT	UT 007	RBNU	JMD 024	41.2560	-111.5200
Public Grove Canyon, UT	UT 008	RBNU	MM 068	41.2560	-111.5200
North Fork Scatter Creek, WA	WA 008	RBNU	CDS 3522	48.5542	-118.8197
Colville NF, WA	WA 009	RBNU	GMS 315	48.3725	-118.3291
Colville NF, WA	WA 010	RBNU	GMS 316	48.3725	-118.3291
Colville NF, WA	WA 011	RBNU	GMS 318	48.3725	-118.3291
Factoria, WA	WA 012	RBNU	EVL 1261	47.5767	-122.1579
Bainbridge, WA	WA 013	RBNU	EVL 1511	47.6918	-122.5620
Robinson Canyon, WA	WA 014	RBNU	SVE 1080	47.0182	-120.7594
Iron Creek, WA	WA 015	RBNU	JK 14 003	48.1338	-118.6523

Olympic NF, WA	WA 016	RBNU	DHB 5062	48.3725	-118.3291
Buck Meadows, WA	WA 017	RBNU	GWS 3913	47.0287	120.9359
Robinson Canyon, WA	WA 018	RBNU	CDS 5060	47.0182	-120.7594
Colville NF, WA	WA 019	RBNU	GMS 319	48.3725	-118.3291
Colville NF, WA	WA 020	RBNU	GMS 317	48.3725	-118.3291
Cle Elum, WA	WA 021	RBNU	GAV 83	47.1954	-120.8331
Republic, WA	WA 022	RBNU	DAB 516	48.6400	-118.4500
Tacoma, WA	WA 023	RBNU	TNL 288	47.2628	-122.2929
Tacoma, WA	WA 024	RBNU	BLE 027	47.7066	-122.6377
Mount Baker-Snoqualmie NF, WA	WA 025	RBNU	GMS 352	48.2595	-121.3238
Iron Creek, WA	WA 026	RBNU	KLE 777	48.1338	-118.6523
Rabbit Creek, WA	WA 027	RBNU	CSW 3873	48.5554	-118.5989
Redmond, WA	WA 028	RBNU	AKW 054	47.6418	-122.0804
Mount Rainier, WA	WA 029	RBNU	RBH 100	47.0682	-120.6719
Laramie Mountains, WY	WY 001	RBNU	DMNS 46045	42.4241	-105.3365
Laramie Mountains, WY	WY 002	RBNU	DMNS 46066	42.4241	-105.3365
Laramie Mountains, WY	WY 003	RBNU	DMNS 46070	42.4241	-105.3365
Laramie Mountains, WY	WY 004	RBNU	DMNS 46064	42.4241	-105.3365
Laramie Mountains, WY	WY 005	RBNU	DMNS 35185	41.1689	-104.8268
Laramie Mountains, WY	WY 006	RBNU	DMNS 46068	42.4241	-105.3365
Laramie Mountains, WY	WY 007	RBNU	DMNS 46046	42.4241	-105.3365
Laramie Mountains, WY	WY 008	RBNU	DMNS 46065	42.4241	-105.3365
Laramie Mountains, WY	WY 009	RBNU	DMNS 43217	41.3068	-105.6159
Laramie Mountains, WY	WY 010	RBNU	DMNS 46067	42.4241	-105.3365
Laramie Mountains, WY	WY 011	RBNU	DMNS 46069	42.4241	-105.3365
Prineville, OR	OR 025	RBNU	DMNS 47941	44.5506	-120.6142
Summer Lake, OR	OR 026	RBNU	DMNS 48073	42.7423	-121.3828

Pine Mountain, OR	OR 027	RBNU	DMNS 48005	43.8026	-120.9359
Brighton, CO	CO 001*	RBNU	DMNS 35293	39.9811	-104.7874
Boulder, CO	CO 002*	RBNU	DMNS 48666	40.0456	-105.3072
Fort Collins, CO	CO 003*	RBNU	DMNS 49116	40.5255	-105.0234
Salida, CO	CO 004	RBNU	DMNS 45937	38.6574	-106.1133
Denver, CO	CO 005	RBNU	DMNS 45628	39.7481	-104.9438
Golden, CO	CO 006	RBNU	DMNS 49780	39.8168	-105.3911
Lakewood, CO	CO 007	RBNU	DMNS 45714	39.7008	-105.1546
Lakewood, CO	CO 008	RBNU	DMNS 48665	39.6776	-105.0973
Denver, CO	CO 009	RBNU	DMNS 43265	39.7311	-104.9352
Boulder, CO	CO 010	RBNU	DMNS 44444	40.0227	-105.2842
Highlands Ranch, CO	CO 011	RBNU	DMNS 49795	39.5392	-104.9615
Longmont, CO	CO 012	RBNU	DMNS 49005	40.1730	-105.1144
Evergreen, CO	CO 013	RBNU	DMNS 42697	39.6187	-105.2842
Fort Collins, CO	CO 014	RBNU	DMNS 43096	40.5871	-105.1124
San Isabel NF, CO	CO 015	RBNU	DMNS 39997	37.2500	-105.1083
Denver, CO	CO 016*	RBNU	DMNS 49774	39.7476	-104.9429
Franktown, CO	CO 017*	RBNU	DMNS 42795	39.3707	-104.7821
Fort Collins, CO	CO 018*	RBNU	DMNS 49004	40.5445	-105.0825
Denver, CO	CO 019*	RBNU	DMNS 45295	39.7474	-104.9437
Wood Buffalo National Park, AB	NEAB 001	RBNU	2091-71926	59.5937	-112.2766
Wood Buffalo National Park, AB	NEAB 002	RBNU	2091-71927	59.5127	-112.2789
Wood Buffalo National Park, AB	NEAB 003	RBNU	2091-71928	59.5127	-112.2789
Wood Buffalo National Park, AB	NEAB 004	RBNU	2091-71929	59.5127	-112.2789
Wood Buffalo National Park, AB	NEAB 005	RBNU	2091-71930	59.5127	-112.2789
Twin Falls Lake, AB	CAB 009	RBNU	2091-71931	57.4554	-117.5133
Prince Albert National Park, SK	SK 002	RBNU	2091-71932	53.9150	-106.0196

Prince Albert National Park, SK	SK 003	RBNU	2091-71933	53.6384	-106.0666
Prince Albert National Park, SK	SK 004	RBNU	2091-71934	53.7716	-106.2930
Prince Albert National Park, SK	SK 005	RBNU	2091-71935	53.7716	-106.2930
Prince Albert National Park, SK	SK 006	RBNU	2091-71936	54.0345	-106.4157
Prince Albert National Park, SK	SK 007	RBNU	2091-71937	54.0345	-106.4157
Prince Albert National Park, SK	SK 008	RBNU	2091-71938	53.6384	-106.0666
Prince Albert National Park, SK	SK 009	RBNU	2091-71939	53.6384	-106.0666
Prince Albert National Park, SK	SK 010	RBNU	2091-71940	53.6384	-106.0666
Prince Albert National Park, SK	SK 011	RBNU	2091-71941	53.6384	-106.0666
Prince Albert National Park, SK	SK 012	RBNU	2091-71942	53.9017	-106.1579
Cypress Hills, AB	SEAB 010	RBNU	2091-71943	49.6793	-110.1856
Cypress Hills, AB	SEAB 011	RBNU	2091-71944	49.6793	-110.1856
Cypress Hills, AB	SEAB 012	RBNU	2091-71945	49.6793	-110.1856
Cypress Hills, AB	SEAB 013	RBNU	2091-71946	49.6456	-110.0302
Cypress Hills, AB	SEAB 014	RBNU	2091-71947	49.6631	-110.3039
Cypress Hills, AB	SEAB 015	RBNU	2091-71948	49.6573	-110.1021
Cypress Hills, AB	SEAB 016	RBNU	2091-71949	49.6573	-110.1021
Cypress Hills, AB	SEAB 017	RBNU	2091-71950	49.6573	-110.1021
Ottawa NF, MI	MI 020	RBNU	76959	46.6414	-88.8660
Lesser Slave Lake, AB	CAB 010	RBNU	74892	55.1415	-114.3230
Lesser Slave Lake, AB	CAB 011	RBNU	74893	55.1415	-114.3230
Lesser Slave Lake, AB	CAB 012	RBNU	74891	55.1415	-114.3230
Lesser Slave Lake, AB	CAB 013	RBNU	74890	55.3042	-114.8820
Lesser Slave Lake, AB	CAB 014	RBNU	74894	55.3042	-114.8820
Cold Lake, AB	CAB 015	RBNU	74889	54.3991	-110.0799
Cold Lake, AB	CAB 016	RBNU	74895	54.3991	-110.0799
Lac du Bonnet, MB	MB 011	RBNU	TMP19-034	50.4179	-95.6042

Thompson Airport, MB	MB 012	RBNU	MMF19-159	55.2801	-98.3863
Flin Flon, MB	MB 013	RBNU	TMP19-045	54.9686	-101.4581
Flin Flon, MB	MB 014	RBNU	TMP19-049	54.9686	-101.4581
Flin Flon, MB	MB 015	RBNU	MMF19-179	54.9417	-101.4554
Flin Flon, MB	MB 016	RBNU	CLB 2527	54.9686	-101.4581
Lac du Bonnet, MB	MB 017	RBNU	ERGC-23	50.6270	-95.5373
Lac du Bonnet, MB	MB 018	RBNU	BMW 714	50.6306	-95.5668
Ottawa NF, MI	MI 021	RBNU	245862	46.6602	-88.8693
Ottawa NF, MI	MI 022	RBNU	245893	46.6602	-88.8693
Ottawa NF, MI	MI 023	RBNU	245894	46.6602	-88.8693
Ottawa NF, MI	MI 024	RBNU	245912	46.6602	-88.8693
Ottawa NF, MI	MI 025	RBNU	245914	46.6602	-88.8693
Ottawa NF, MI	MI 026	RBNU	245918	46.6602	-88.8693