

**LANDSCAPE GENETICS OF TWO DECLINING NORTH AMERICAN
AERIAL INSECTIVORES**

THILINI THAKSHILA MADDEGAMGODA LEKAMLAGE
Bachelor of Science (Hons), University of Colombo, Sri Lanka, 2020

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

MASTER OF SCIENCE

in

BIOLOGICAL SCIENCES

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

© Thilini Thakshila Maddegamgoda Lekamlage, 2023

LANDSCAPE GENETICS OF TWO DECLINING NORTH AMERICAN AERIAL
INSECTIVORES

THILINI THAKSHILA M.L

Date of Defence: August 17, 2023

Dr. T. Burg Thesis Supervisor	Professor	Ph.D.
----------------------------------	-----------	-------

Dr. K. Hobson Thesis Examination Committee Member Environment and Climate Change Canada Department of Biological Science University of Western Ontario	Professor	Ph.D.
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------	-------

Dr. A. Iwaniuk Thesis Examination Committee Member	Professor	Ph.D.
----------------------------------------------------------	-----------	-------

Dr. J. McCune Thesis Examination Committee Member	Assistant Professor	Ph.D.
---------------------------------------------------------	---------------------	-------

Dr. D. Yevtushenko Chair, Thesis Examination Committee	Associate Professor	Ph.D.
--------------------------------------------------------------	---------------------	-------

DEDICATION

To my dear pet Teena, thank you for the love and happiness you brought into my life. Your memory will forever remain in my heart, and I am grateful for the time we spent together.

ABSTRACT

I used molecular markers to examine the population genetic structure of two declining North American aerial insectivores; barn swallows and bank swallows. Both species are long-distance migrants and have a wide distribution. The findings revealed similar overall genetic clusters corresponding to western, central and eastern regions of North America. Strong population differentiation of the eastern population in both species was observed making this population a priority for conservation, especially given the large decline in the region. The results provide evidence that both species inhabited multiple refugia during the Pleistocene and interplay between migratory pathways and geographical barriers also play an important role in shaping genetic differences between populations. This study provides insights into selecting appropriate genetic techniques in future studies of migratory species and assists with developing conservation strategies.

ACKNOWLEDGEMENTS

First and foremost, I extend my heartfelt thanks to my supervisor, Dr. Theresa Burg. Her continuous support and belief in my abilities have been instrumental in shaping my research and academic endeavours. Dr. Burg's unwavering dedication to excellence has inspired me to strive for the best, even during challenging times. I am truly grateful for her patience, guidance, and unwavering commitment to my success. Thank you so much for believing in me and providing me with the best experiences throughout my academic journey. Being selected as one of their graduate students is a true honour and a testament to their confidence in my abilities.

I would also like to express my appreciation to Dr. Keith Hobson and Mr. Kevin Kardynal from Environment and Climate Change Canada for their invaluable support and regular meetings. Their expertise and insightful feedback have been crucial in refining my research and enhancing the quality of my work. I am grateful for their willingness to share their knowledge and for their encouragement throughout this project.

Furthermore, I would like to acknowledge the members of my committee Dr. A. Iwaniuk, Dr. J. McCune and Dr. K. Hobson for their valuable inputs and guidance. Their collective expertise and comments have helped shape the direction of my research and enriched its overall quality. I am thankful for their time, effort, and commitment to ensuring the success of my project. I am also grateful to Dr. Dmytro Yevtushenko and Dr. Robert Laird for their feedback during the seminar course at the University of Lethbridge in my first year.

I would also like to express my gratitude to my undergraduate supervisor, Dr. Sampath Seneviratne, who ignited my passion for avian science and guided me onto

this path. His mentorship and enthusiasm for the subject matter have been a driving force behind my continued interest and growth in the field. I would like to extend my sincere appreciation to Dr. Elizabeth Scordato, my undergraduate co-supervisor, for her invaluable contribution to my academic journey. You both are remarkable and passionate research scientists whose guidance and mentorship have greatly influenced the path I have chosen. I also like to thank Dr. Tegan Barry and Ms. Laurie Pacarynuk, as well as all the co-Tas of BIOL 1010 labs, for their exceptional TA experience.

I would like to thank all the past and present members of Burg Lab. I would like to start with thanking Dilini Akka (Dr. Dilini Abeyrama) for being there for me right from the start, teaching me the ropes of NGS work. And let me not forget to mention the amazing breakfasts she made for me on those early mornings when I rolled in at 5:30 am, knowing I'd forget to eat otherwise. I cannot express enough how much your kindness meant to me. I would also like to thank the postdocs in the lab, Dr. Juan Enciso, for all the help and inspiration in learning coding. His patience and guidance have ignited my passion for coding and have expanded my skill set in analysis. I am grateful for his contributions and all his help. I would like to express my gratitude to Dr. Linda Lait for her valuable inputs and contributions throughout my time in the lab. I appreciate her support and guidance. Aaron and Bryan thanks for all your patience, guidance, and willingness to assist have been invaluable throughout my time in the grad office. I can't forget to mention the past lab members, Danika, Bukola and Dr. Brendan Graham. You were instrumental in helping me get my feet off the ground. Thanks for all Friday lunches and Tims runs. And let's not forget the awesome field days I had with Danika, Aaron and Linda. Those were some great times and great experiences.

I would like to express my heartfelt gratitude to the Beaty Biodiversity Museum, Denver Museum, Burke Museum, RBC Museum, RSK Museum, Sarah Fensore and all the bird banding stations for their invaluable contributions to this study. I would also like to thank Mr. Crosby Adam from ECCC for sending samples and all volunteers who helped me with sample collections. Your willingness to provide samples has been essential in advancing our research and understanding of avian science. A special thanks goes to Ildiko Szabo, collections curator for her generous accommodations and exceptional hospitality during the sample collection phase at UBC.

I would like to express my deepest appreciation for the financial support that made this project possible and facilitated my stipend throughout the research. I am immensely grateful to the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, the University of Lethbridge Graduate Research Award, and Environment and Climate Change Canada for their generous funding and stipend support. Additionally, I would like to extend my heartfelt thanks to the University of Lethbridge for accepting me as a graduate student.

I would like to extend a heartfelt thank you to my family (my mom, dad, sister, and grandma) for their unwavering support and encouragement. Moving to Canada from my home country Sri Lanka was undoubtedly a challenging transition, not just for me but for my family as well. Despite the difficulties, my family has always stood by my side, providing endless support and belief in my abilities. I am eternally grateful for the immense love, strength, and guidance they have showered upon me. Thank you, from the bottom of my heart, for being the pillars of strength and the driving force behind my pursuit of my goals and dreams. I am forever indebted to you for the support, guidance, and love that you continue to provide me.

I would like to extend a heartfelt thank you to my loving husband, Kanchu, for his unwavering support and understanding. Throughout my journey, he has been my support, always standing by my side, listening to my tears and stress, and providing endless encouragement. I cannot express how grateful I am for his presence in my life.

I would also like to express my heartfelt gratitude to all my friends back home and my friends and roommates here in Lethbridge. Your unwavering support, comforting words, and shared understanding of my stress have been invaluable throughout this journey.

I appreciate and acknowledge the contributions and support of all individuals involved in this study, whether they provided assistance, guidance, or even offered a single word of encouragement. I extend my heartfelt thanks to all of you for being a part of my success.

Last but not least, heartfelt thanks to all the incredible birds who have donated their blood and feathers for science. Keep being amazing, feathered friends, and may your populations soar high and thrive!

TABLE OF CONTENTS

Dedication	iii
Abstract	iv
Acknowledgement	v
List of Tables	xii
List of Figures	xiii
List of Abbreviations, Acronyms and Symbols	xv
Chapter 1: General Introduction	1
1.1 Population Genetics	1
1.1.1 Genetic Diversity	1
1.1.2 Barriers to Gene Flow	1
1.2 Migration and Migratory Divides	3
1.3 Study Species	5
1.3.1 Barn Swallow (<i>Hirundo rustica</i>)	6
1.3.2 Bank Swallow (<i>Riparia riparia</i>)	7
1.4 Molecular Markers	8
1.4.1 Microsatellite Markers	8
1.4.2 Next Generation Sequencing	9
1.5 Thesis Overview	10
References	11
Chapter 2: Landscape Genetic Structure of Barn Swallows Breeding in North America	18
Abstract	18
2.1 Introduction	19
2.2 Methods	22

2.2.1 Sample Collection and DNA Extraction	22
2.2.3 RADseq Analysis (Restriction Site Associated DNA Sequencing)	24
2.3 Results	29
2.3.1 Microsatellite Analyses	29
2.3.2 RADseq Analyses.....	29
2.3.3 lcWGS Analysis	30
2.4 Discussion	32
2.4.1 Identifying Genetic Structure Using Different Genetic Markers	34
2.4.2 Pleistocene Glaciations Shape the Population Genetic Structure.....	35
2.4.3 Role of Migratory Divides and Contemporary Barriers.....	36
2.4.4 Migratory Connectivity	36
2.4.5 Genomic Scans Provide Functional Insights	37
2.5 Conclusions	39
References	40
Chapter 3: Population Genetic Structure of North American Bank Swallows	57
Abstract	57
3.1 Introduction	58
3.2 Methods.....	62
3.2.1 Low Coverage Whole Genome Sequencing (lcWGS) Analyses.....	62
3.2.2 Identifying Bank Swallow Migratory Pathways	65
3.3 Results	66
3.3.1 Population Genetic Structure of Bank Swallows using lcWGS Data	66
3.3.2 Comparative Population Analysis	66
3.3.3 Migratory Pathways.....	67
3.4 Discussion.....	67
3.4.1 Insights into the Role of Glacial Refugia in Shaping Species Distribution and Colonization of Bank Swallows	68

3.4.2 Role of Contemporary Barriers and Migratory Pathways	70
3.5 Conclusions	72
References	73
Chapter 4: General Discussion.....	86
4.1 Aerial Insectivores.....	86
4.2 Population Genetic Structure of North American Barn Swallows and Bank Swallows	86
4.3 Migratory Connectivity Using Genomic Approaches.....	90
4.4 Usage of Different Markers in a Population Genetic Study.....	90
4.5 Conclusions and Future Directions	91
References	94
Appendices.....	96
Appendix 1: Supplementary Information for Chapter 2	96
Appendix 2: Supplementary Information for Chapter 3	107

LIST OF TABLES

Table 2.1:	Sampling locations and number of samples utilized in each analysis after filtering: microsatellite (Msat), restriction site associated DNA sequencing (RADseq) and low coverage whole genome sequencing (lcWGS).	46
Table 2.2:	Comparison of observed and (H_O) and expected (H_E) heterozygosity for each population using RADseq analysis.	47
Table 2.3:	Genes of interest from genome scans using lcWGS between each population comparison.	48
Table 3.1:	Sampling locations and the number of samples utilized in the analyses.	78

LIST OF FIGURES

Figure 1.1:	The annual population trends of breeding bird survey (BBS) of swallows, represented with a log scale from 1996 – 2010. Sand martin also known as bank swallow in the New World.	15
Figure 1.2:	Subspecies distribution and differences of plumage coloration of barn swallow (<i>Hirundo rustica</i>) subspecies complex. Dash arrows represent hypothesized colonization events (Scordato & Safran, 2014).	16
Figure 1.3:	Subspecies distribution of bank swallow (<i>Riparia riparia</i>) subspecies complex. Red stars indicate the sample locations used in the Schweizer et al. (2018) and black stars indicate samples used in the Pavlova et al. (2008). Hypothetical subspecies boundaries indicated in question marks (Schweizer et al., 2018).	17
Figure 2.1:	Estimated migration routes and wintering sites for barn swallows breeding in North America using archival light-level geolocators.	49
Figure 2.2:	Range map of barn swallows in North America and sample locations used in this study. Range data obtained from BirdLife International, and map was created using QGIS 3.2.	50
Figure 2.3:	Principal coordinate analysis (PCoA) of eight populations using six microsatellite markers.	51
Figure 2.4:	Principal Component Analysis (PCA) indicating (a) first three components with loose filtering options including 39 individuals and 850 SNPs (b) first three components for tight filtering option including 29 individuals and 604 SNPs using RADseq data for barn swallows.	52
Figure 2.5:	Principal component analysis for barn swallows including breeding, migratory and wintering populations using lcWGS indicating (a) PC 1 vs PC 2 (b) PC 3 vs PC 4. PC1, PC2, PC3 and PC4 explain 2.55%, 2.49%, 2.41% and 2.39% variance respectively. Triangles = migratory, circles = breeding, and squares = wintering populations. Shades of red indicate the western populations, shades of purple indicate the central populations, and shades of blue indicate the eastern populations.	53
Figure 2.6:	Principal component analysis for barn swallows using only breeding populations data from lcWGS.	54

Figure 2.7:	Pairwise F_{ST} comparison across the genome between three different clusters using lcWGS (a) west vs east (b) west vs central (c) east vs central.	55
Figure 2.8:	Site Frequency Spectrum (SFS) for each cluster using lcWGS data (a) SFS for the western, (b) eastern, and (c) central clusters.	56
Figure 3.1:	Range map of bank swallows in North America and sample locations used in this study. The range map was obtained from the Cornell Lab of Ornithology.	79
Figure 3.2:	Principal component analysis for breeding populations of bank swallows using lcWGS indicating (a) PC1 vs PC2 (b) PC1 vs PC3. PC1, PC2 and PC3 explain 9.39%, 4.92% and 4.84% variance respectively.	80
Figure 3.3:	Principal component analysis for bank swallows indicating three clusters using lcWGS.	81
Figure 3.4:	Genome scans of F_{ST} from the west (red) and central (purple) comparison. A and B indicate the high F_{ST} peaks. π and Tajima's D are included for those two regions of interest (A) scaffold JAIXNV010000025.1 (B) scaffold JAIXNV010000006.1.	82
Figure 3.5:	Genome scans of F_{ST} from the west (red) and east (blue) comparison. (A) scaffold JAIXNV010000025.1 indicates the high F_{ST} peak. π and Tajima's D are included for the region of interest.	83
Figure 3.6:	Genome scans of F_{ST} from the central (purple) and east (blue) comparison. (A) scaffold JAIXNV010000031.1 indicates the high F_{ST} peak. π and Tajima's D are included for the region of interest.	84
Figure 3.7:	Estimated migration routes for bank swallows breeding in North America using publicly available Motus data. Birds are from Alaska (purple), British Columbia (red) and Newfoundland (blue).	85

LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS

ANGSD	Analysis of next-generation Sequencing Data
BANS	Bank Swallow
BARS	Barn Swallow
°C	Degrees Celsius
DNA	Deoxyribonucleic acid
F _{ST}	Fixation index
GIS	Geographical information system
He	Expected heterozygosity
Ho	Observed heterozygosity
HWE	Hardy-Weinberg Equilibrium
INDELs	Insertions and deletions
K	Number of clusters/ancestors
lcWGS	Low coverage whole genome sequencing
LGM	Last glacial maximum
M13	Universal primer sequence tag
MgCl ₂	Magnesium chloride
mM	Millimolar
Msat	Microsatellite
mtDNA	Mitochondrial DNA
n	Sample size
NGS	Next generation sequencing
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PE	Paired-end
RADseq	Restriction site Associated DNA sequencing
SFS	Site frequency spectrum
SNP	Single nucleotide polymorphism
T1	Primer annealing temperature one
T2	Primer annealing temperature two
vcf	Variant call format
WGS	Whole genome sequencing
μM	Micromolar
π	pi/nucleotide diversity

Abbreviations for Populations

AL	Alabama
CA	California
COL	Colombia
CO	Colorado
SEAB	Cypress Hills, Southeast Alberta

MB	Manitoba
NB	New Brunswick
ND	North Dakota
ON	Ontario
QC	Quebec
SK	Saskatchewan
SBC	Southern British Columbia
TX	Texas
WA	Washington

Chapter 1: General Introduction

1.1 Population Genetics

1.1.1 Genetic Diversity

Genetic diversity in a population is a core factor in evolutionary biology. When gene flow is restricted and genetic diversity is low, the fitness of a population decreases due to limitations in the adaptive potential for changing ecological conditions (Frankham, 2003; Reed & Frankham, 2003). Changing environmental conditions such as habitat loss, climate change, introduced pathogens, or competition for resources can adversely affect isolated populations, particularly small ones (Frankham et al., 1999). Sudden environmental change may cause populations to experience reduced fitness, increased mortality, or decreased fecundity, leading to negative population growth in an unfamiliar environment (Bell & Gonzalez, 2009). Moreover, small, isolated populations with increased levels of homozygosity are susceptible to the negative consequences of genetic drift and inbreeding depression (Robinson et al., 2019). On the other hand, population connectivity facilitates gene flow and plays a vital role in maintaining genetic diversity. It allows for the exchange of genetic material between populations, which helps mitigate the effects of stochastic environmental changes by ensuring genetically diverse individuals within the population (Lacy, 1987).

1.1.2 Barriers to Gene Flow

Within species, some individuals can exhibit genetic distinctiveness due to barriers to gene flow (Ravinet et al., 2017). These barriers to gene flow can be diverse including geographical, behavioural, or climatic barriers (Adams & Burg, 2015). For instance, widespread glaciation during the Pleistocene caused species to reside solely in ice-

free refugium. The last series of glaciations started ~ 2.5 million years ago, and the last glacial maximum (LGM) ended ~12,000 years ago. Major climatic oscillations between cold glacial and warmer interglacial, habitat shifts, fragmentation and extinction events were observed in many species during the LGM. These stochastic and selective effects on genetic variation and architecture due to the effect of glacial cycles persist to this day (Hewitt, 2004). For instance, woolly mammoth (*Mammuthus primigenius*) experienced a significant decline in population size and genetic diversity during the LGM, likely due to habitat shifts and fragmentation of populations. This genetic bottleneck event resulted in a loss of adaptive variation, rendering the species vulnerable to extinction in the face of further environmental changes (Palkopoulou et al., 2015).

Glaciation events also played a significant role in shaping migratory pathways. The changes brought about by glaciation, such as the formation of ice sheets, melting of icesheets, and shifts in the landscape, have led to alterations in the availability of suitable habitats and the emergence of geographic barriers. Geographical barriers such as mountain ranges, water bodies, and habitat fragmentation can restrict migratory pathways and limit gene exchange among populations. For example, the Rocky Mountain range, spanning from British Columbia in Canada to New Mexico in the United States, acts as a prominent geographical barrier in North America, resulting in genetic differentiation between populations on either side of the mountain range (Hindley et al., 2018; Lait & Burg, 2013).

As a result of the formation of barriers that obstructed traditional migratory routes for birds, and they were forced to adapt and find alternative pathways to reach their breeding and wintering grounds. Some species/populations may have followed newly

exposed land routes, while others may have utilized coastal regions that remained ice-free.

1.2. Migration and Migratory Divides

Migration is a prevalent phenomenon observed throughout the animal kingdom, where many species undertake regular movements between different locations for breeding, foraging, or other purposes (Alerstam et al., 2003). Migration patterns in closely related species can vary even among closely related species and within a species with some populations being resident and others migratory. Additionally, within migratory species, there can be variation in the distance of their migratory journeys, with some birds exhibiting short-distance migration while others engage in long-distance. This diversity in migration patterns can be influenced by various factors, such as the availability of food, breeding grounds, and changes in weather patterns (Alerstam et al., 2003). Long-distance migrants undertake some of the most remarkable and challenging journeys. These birds travel vast distances, sometimes crossing entire oceans, mountain ranges or continents, to reach their breeding or feeding grounds. For instance, bird species such as hudsonian godwit (*Limosa haemastica*) breed in boreal Arctic North America and move to northern Argentina to winter (Senner, 2012).

Species breeding in northern temperate regions (e.g., Nearctic Neotropical migrants) travel a relatively long distance to Central and South America to make use of all available non-breeding habitats (Finch et al., 2017; Machado et al., 2018). Both of the subspecies of study species, barn swallows and bank swallows studied in this study are long distance migrants who breed in northern parts of North America and migrate to Central and South America to winter.

Despite the challenges and high consumption of energy during migration, individuals benefit from migratory behaviour. Migration reduces the competition for resources, and the tropics provide favourable conditions such as warm temperatures (Lank et al., 2003). Migration can also enhance gene flow. When we compare a species with both migratory and resident populations, migratory populations show higher genetic diversity compared to the resident populations (Delmore et al., 2020). Different migratory routes, high migratory connectivity, and different migration timing can act as barriers to gene flow (Delmore & Irwin, 2014).

Migratory divides are observed in populations that breed in close proximity but exhibit divergent migration routes to get to their overwintering grounds (Delmore et al., 2012). Studies have shown that migratory divides reduce gene flow resulting in genetic differentiation between populations. Migratory divides are widely observed in many migratory bird species. For example, Swainson's thrushes (*Catharus ustulatus*) in western North America exhibit a migratory divide at the Coast Mountains corresponding to inland populations and coastal populations which exhibit population genetic structure between western and eastern populations. Similarly, migratory European blackcaps (*Sylvia atricapilla*) west of the Ural Mountains migrate between Europe or Africa whereas populations east of the mountains migrate to the Indian subcontinent or Southern Asia (Berthold, 2001; Delmore et al., 2012).

Migratory divides enhance the population genetic structure as it reduces gene flow between populations. Differences in migratory behaviour can also cause divergence of morphological traits. For instance, long-distance migrants will show morphological adaptations to travel long distances, and adaptation to specific conditions that they face during the migration and so on. For an example longer wing length, pointed wings, storing fuel as energy are some of the morphological and physiological

adaptations. However, high local adaptation to specific conditions can cause extinction of entire populations due to stochastic events (Møller et al., 2011). On the other hand, gene flow reduces local adaptation but decreases susceptibility to environmental perturbations and pathogens or parasites. Although many North American bird species are migratory, aerial insectivores are group of great conservation concern because of large declines in their populations.

1.3 Study Species

Aerial insectivores are a group of birds that feed on flying insects and commonly includes flycatchers, swallows, nightjars, and swifts. North American aerial insectivores have experienced overall population declines averaging of 59% over the last ~60 years. Aerial insectivores are experiencing more severe declines compared to the other passerine species (Nebel et al., 2020; Rosenberg et al., 2019). A recent global assessment suggested that insect biomass was falling at a rate of 2.5% annually (Sánchez-Bayo & Wyckhuys, 2019). Climate change, anthropogenic modifications, heavy pesticide usage and habitat loss combined with other factors such as predation are interconnected reasons for the recent dramatic decline of aerial insectivores (Imlay & Leonard, 2019; Spiller & Dettmers, 2019). Rates of decline vary by species and region; northeastern North America experienced the highest declines of aerial insectivores (Michel et al., 2016; Nebel et al., 2020; Rosenberg et al., 2019). For this study, I chose two steep declining species of aerial insectivores; barn swallow (*Hirundo rustica*), and bank swallow (*Riparia riparia*). According to the annual breeding bird survey population index barn swallows show 68.9% decline from 1970 to 2019 whereas bank swallows show 98% loss in Canada for the same time period (Figure 1.1) (Imlay & Leonard, 2019; Sauer et al., 2017). Conservation efforts can be

expensive, making conservation prioritization necessary, especially for more vulnerable species. For instance, if we identify populations that are both vulnerable and genetically distinct, they should receive higher prioritization and require unique management practices in conservation planning (Evermann et al., 1989).

1.3.1 Barn Swallow (*Hirundo rustica*)

Barn swallows are small passerine birds (17 – 20 g); they are the most abundant and widely distributed swallow species in the world (Del Hoyo et al., 2014). Barn swallows breed throughout North America and Eurasia with six recognized subspecies present in barn swallows; five subspecies in the Old World and one subspecies in New World. Based on mitochondrial DNA analysis of four subspecies, two main clades were identified; the first clade includes *H. r. rustica* (distributed across Europe and central Asia), *H. r. savignii* (Egypt), and *H. r. transitiva* (eastern Mediterranean), and the second clade consists of American subspecies *H. r. erythrogaster* (Dor et al., 2010). The plumage coloration of the six barn swallow subspecies (Figure 1.2) also matches the phylogenetic structure of the subspecies complex (Del Hoyo et al., 2014; Dor et al., 2010). Phylogenetic studies of barn swallow subspecies indicate expansion from an African common ancestor. The North American subspecies (*H. r. erythrogaster*) is believed to have arisen from the Asian population crossing through the Bering Sea about 100,000 years ago. My thesis focuses on the North American populations of the barn swallow subspecies *H. r. erythrogaster*.

North American barn swallows spend their breeding season in most of North and Central America and overwinter in lowlands of southern Mexico and South America (Brown, 2020). Using light level geolocators, continental migratory divides for barn

swallows were identified. Eastern populations from New Brunswick and Ontario travel longer distances via East Coast and Caribbean islands to central South America, whereas western populations from Washington and Saskatchewan travel relatively shorter distances to Central America and northwestern South America via western routes (Hobson et al., 2015). As described earlier in this chapter, a migratory divide could act as a barrier to gene flow, therefore I used molecular tools to identify the genetic structure in barn swallows across the migratory divide.

1.3.2 Bank Swallow (*Riparia riparia*)

Bank swallows are also known as collared sand martins in the Old World, and they are considered to be a more primitive or basally branching group of swallows as they are burrow nesting species (Schweizer et al., 2018). They are widespread across most of Europe, Asia and North America. Studies suggest that pale sand martin (*Riparia diluta*) and collared sand martin or bank swallow (*Riparia riparia*) as it is known in North America, diverged sometime between late Pliocene and middle Pleistocene (Schweizer et al., 2018). The absence of well supported clades and haplotype networks using mitochondrial markers indicates relatively low genetic diversity and recent divergence in bank swallows (Pavlova et al., 2008). Three subspecies of bank swallows are recognized based on mitochondrial markers: *Riparia riparia riparia* breeds in North America, Eurasia, northwest Africa and the Mediterranean region. North American populations are also listed under the name *R. r. maximiliani* as they have a smaller body size (10 – 18 g) compared to the Eurasian population (Garrison & Turner, 2020;). The second and third subspecies *R. r. ijimae* and *R. r. shelleyi* breed in eastern part of the Asia and Egypt respectively (Figure 1.3) (Garrison & Turner, 2020; Schweizer et al., 2018).

I will be focusing on bank swallows breeding in North America (*R. r. riparia*) in this study. North American bank swallows breed throughout northern part of the North America from Alaska to north Yukon, northern Manitoba, northern Ontario, central Quebec, throughout Maritime provinces, and Newfoundland south to the upper part of the United States of America (Garrison & Turner, 2020). Studies of migratory patterns of the North American bank swallows are limited and most are restricted to populations breeding in Atlantic Canada and wintering in northern Argentina and southern Brazil (Imlay et al., 2020). Unlike barn swallows, only a few studies were performed to understand the genetic structure of the bank swallows in North America and all of those studies were limited to mitochondrial markers (Zink et al., 2006). Mitochondrial markers indicate a shallow history and low genetic diversity for this group, so I used lcWGS to identify the population genetic structure.

1.4 Molecular Markers

The use of genetic markers to understand population structure provides important insights about adaptive loci, and how landscape level and environmental factors can shape the population in different ways (Barbosa et al., 2021). In this study, I used microsatellite markers as well as single nucleotide polymorphic markers (SNPs) from restriction site-associated DNA sequencing (RAD-seq) and low coverage whole genome sequencing (lcWGS) to examine the population genetic structure of *Hirundo rustica* (barn swallows) () and *Riparia riparia* (bank swallows).

1.4.1 Microsatellite Markers

Microsatellite markers are short tandem repeats of DNA with one to six nucleotides in the core sequence. These markers are biparentally inherited, codominant alleles, providing both male and female histories (Jarne & Lagoda, 1996). Microsatellite loci

have a high level of genetic diversity and mutations arise due to polymerase slipping during DNA replication. The variation in the length and number of repeats are responsible for the polymorphism (Liu & Cordes, 2004). Microsatellites offer a cost-effective method, allowing for the inclusion of a large number of samples in the analyses, without requiring substantial amounts of high-quality DNA like next-generation sequencing methods. Although microsatellite markers have been used to identify population genetic structure, they do not provide as high coverage and high resolution for identifying genetic differences or the ability to find loci under selection as next-generation sequencing (NGS).

1.4.2 Next Generation Sequencing

1.4.2.1. Restriction Site Associated DNA Sequencing (RADseq)

With the advancement of sequencing technology, NGS facilitates increasing the breadth and depth of coverage for sequencing using different sequencing methods (Lou et al., 2021). RADseq is a type of next-generation sequencing method, where DNA is fragmented using restriction enzymes and resulting fragments are sequenced. In this study, I used three restriction enzymes to fragment the DNA (Abeyrama et al., 2023; Peterson et al., 2012). RAD sequencing usually achieves a high depth of coverage (> 20X), and it is a useful tool to study population genetic structure based on SNPs (Wright et al., 2019). Despite the ability of RADseq to examine genome-wide patterns of variation, it only samples a subset of the genome ~ 1%, thereby, missing a lot of variation, selection and divergence in the genome (Lou et al., 2021). On the other hand, WGS sequences the entire genome which is useful to detect the fine scale genetic structure, especially in the population level study.

1.4.2.2. Low Coverage Whole Genome Sequencing (lcWGS)

Whole genome sequencing increases the breadth of coverage, and studies that used WGS can examine genetic differences associated with phenotypes that RADseq may not detect. However, this process is very expensive and requires a lot of resources. Therefore, utilizing lcWGS to study entire genomes is becoming increasingly relevant. As the term implies, lcWGS has a low to medium depth of coverage (< 20X) compared to high depth WGS (>20X), and uses a statistical approach, such as likelihood estimations to identify genotypes (Lou et al., 2021).

1.5 Thesis Overview

This thesis is comprised of four chapters. The first chapter, the Introduction describes the basic population genetics, barriers to gene flow, molecular markers and the study species used in this study. The second chapter examines the population genetic structure of barn swallows across the migratory divide using microsatellite markers, SNPs using RADseq and lcWGS. The third chapter focuses on population genetics of bank swallows breeding across North America using SNPs data from lcWGS. The final chapter summarizes the results and discusses the importance and implications of this research. The last chapter discusses future directions and gaps that need to be filled to fully implement conservation plans and management practices to prevent the decline of these species.

References

- Abeyrama, D. K., Boyle, B., & Burg, T. M. (2023). Comparison of genotyping by sequencing procedures to determine population genetic structure. *Functional & Integrative Genomics*, 23(1), 9.
- Adams, R., A. & Burg, T. M. (2015). Influence of ecological and geological features on rangewide patterns of genetic structure in a widespread passerine. *Heredity*, 114(2), 143-154.
- Alerstam, T., Hedenström, A., & Åkesson, S. (2003). Long-distance migration: Evolution and determinants. *Oikos*, 103(2), 247-260.
- Barbosa, S., Andrews, K. R., Goldberg, A. R., Gour, D. S., Hohenlohe, P. A., Conway, C. J., & Waits, L. P. (2021). The role of neutral and adaptive genomic variation in population diversification and speciation in two ground squirrel species of conservation concern. *Molecular Ecology*, 30(19), 4673-4694.
- Bell, G., & Gonzalez, A. (2009). Evolutionary rescue can prevent extinction following environmental change. *Ecology Letters*, 12(9), 942-948.
- Berthold, P. (2001). *Bird Migration: A General Survey*. Oxford University Press.
- Brown, M. B. & Brown C. R. (2020). Barn swallow (*Hirundo rustica*). Cornell Lab of Ornithology.
- Del Hoyo J, & A, E. (2014). Barn swallow (*Hirundo rustica*). In *Handbook of the Birds of the World*. Lynx Edicions.
- Delmore, K. E., Fox, J. W., & Irwin, D. E. (2012). Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences*, 279(1747), 4582-4589.
- Delmore, K. E., & Irwin, D. E. (2014). Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, 17(10), 1211-1218.
- Delmore, K. E., Van Doren, B. M., Conway, G. J., Curk, T., Garrido-Garduño, T., Germain, R. R., Hasselmann, T., Hiemer, D., van der Jeugd, H. P., & Justen, H. (2020). Individual variability and versatility in an eco-evolutionary model of avian migration. *Proceedings of the Royal Society B*, 287(1938), 20201339.
- Dor, R., Safran, R. J., Sheldon, F. H., Winkler, D. W., & Lovette, I. J. (2010). Phylogeny of the genus *Hirundo* and the barn swallow subspecies complex. *Molecular Phylogenetics and Evolution*, 56(1), 409-418.
- Evermann, J. F., Heeney, J. L., McKeirnan, A. J., & O'Brien, S. J. (1989). Comparative features of a coronavirus isolated from a cheetah with feline infectious peritonitis. *Virus Research*, 13(1), 15-27.
- Finch, T., Butler, S. J., Franco, A. M., & Cresswell, W. (2017). Low migratory connectivity is common in long-distance migrant birds. *Journal of Animal Ecology*, 86(3), 662-673.

- Frankham, R. (2003). Genetics and conservation biology. *Comptes Rendus Biologies*, 326, 22-29.
- Frankham, R., Lees, K., Montgomery, M. E., England, P. R., Lowe, E. H., & Briscoe, D. A. (1999). Do population size bottlenecks reduce evolutionary potential? *Animal Conservation* (4), 255-260.
- Garrison, B. A. (1999). Bank Swallow (*Riparia riparia*). *The Birds of North America*, (414), 28.
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183-195.
- Hindley, J. A., Graham, B. A., & Burg, T. M. (2018). The influence of latitude, geographic distance, and habitat discontinuities on genetic variation in a high latitude montane species. *Scientific Reports*, 8(1), 1-11.
- Hobson, K. A., Kardynal, K. J., Van Wilgenburg, S. L., Albrecht, G., Salvadori, A., Cadman, M. D., Liechti, F., & Fox, J. W. (2015). A continent-wide migratory divide in North American breeding barn swallows (*Hirundo rustica*). *PloS One*, 10(6), e0129340.
- Imlay, T., Saldanha, S., & Taylor, P. (2020). The fall migratory movements of bank swallows, *Riparia riparia*: Fly-and-forage migration? *Avian Conservation and Ecology*, 15(1), 2.
- Imlay, T. L., & Leonard, M. L. (2019). A review of the threats to adult survival for swallows (family: *Hirundinidae*). *Bird Study*, 66(2), 251-263.
- Jarne, P., & Lagoda, P. J. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, 11(10), 424-429.
- Lacy, R. C. (1987). Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology*, 1(2), 143-158.
- Lait, L. A., & Burg, T. M. (2013). When east meets west: Population structure of a high-latitude resident species, the boreal chickadee (*Poecile hudsonicus*). *Heredity*, 111(4), 321-329.
- Lank, D. B., Butler, R. W., Ireland, J., & Ydenberg, R. C. (2003). Effects of predation danger on migration strategies of sandpipers. *Oikos*, 103(2), 303-319.
- Liu, Z., & Cordes, J. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, 238(1-4), 1-37.
- Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, 30(23), 5966-5993.
- Machado, A. P., Clément, L., Uva, V., Goudet, J., & Roulin, A. (2018). The rocky mountains as a dispersal barrier between barn owl (*Tyto alba*) populations in North America. *Journal of Biogeography*, 45(6), 1288-1300.

- Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, 18(4), 189-197.
- Michel, N. L., Smith, A. C., Clark, R. G., Morrissey, C. A., & Hobson, K. A. (2016). Differences in spatial synchrony and interspecific concordance inform guild-level population trends for aerial insectivorous birds. *Ecography*, 39(8), 774-786.
- Møller, A., Garamszegi, L., Peralta-Sánchez, J., & Soler, J. (2011). Migratory divides and their consequences for dispersal, population size and parasite–host interactions. *Journal of Evolutionary Biology*, 24(8), 1744-1755.
- Nebel, S., Casey, J., Cyr, M. A., Kardynal, K., Krebs, E., Purves, E., Bélisle, M., Brigham, R., Knight, E., & Morrissey, C. (2020). Falling through the policy cracks: Implementing a roadmap to conserve aerial insectivores in North America. *Avian Conservation and Ecology*, 15(1), 1-8.
- Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., Omrak, A., Vartanyan, S., Poinar, H., & Götherström, A. (2015). Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Current Biology*, 25(10), 1395-1400.
- Pavlova, A., Zink, R. M., Drovetski, S. V., & Rohwer, S. (2008). Pleistocene evolution of closely related sand martins *Riparia riparia* and *R. diluta*. *Molecular Phylogenetics and Evolution*, 48(1), 61-73.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & 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.
- Ravinet, M., Faria, R., Butlin, R., Galindo, J., Bierne, N., Rafajlović, M., Noor, M., Mehlig, B., & Westram, A. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450-1477.
- Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, 17(1), 230-237.
- Robinson, J. A., Räikkönen, J., Vucetich, L. M., Vucetich, J. A., Peterson, R. O., Lohmueller, K. E., & Wayne, R. K. (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Science Advances*, 5(5), eaau0757.
- Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., & Parr, M. (2019). Decline of the North American avifauna. *Science*, 366(6461), 120-124.
- Sánchez-Bayo, F., & Wyckhuys, K. A. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, 232, 8-27.
- Sauer, J. R., Pardieck, K. L., Ziolkowski Jr, D. J., Smith, A. C., Hudson, M.-A. R., Rodriguez, V., Berlanga, H., Niven, D. K., & Link, W. A. (2017). The first 50 years of the North American breeding bird survey. *The Condor: Ornithological Applications*, 119(3), 576-593.

- Schweizer, M., Liu, Y., Olsson, U., Shirihai, H., Huang, Q., Leader, P. J., Copete, J. L., Kirwan, G. M., Chen, G., & Svensson, L. (2018). Contrasting patterns of diversification in two sister species of martins (Aves: *Hirundinidae*): The sand martin *Riparia riparia* and the pale martin *R. diluta*. *Molecular Phylogenetics and Evolution*, *125*, 116-126.
- Scordato, E. S., & Safran, R. J. (2014). Geographic variation in sexual selection and implications for speciation in the barn swallow. *Avian Research*, *5*(8), 1-13.
- Senner, N. R. (2012). One species but two patterns: Populations of the Hudsonian Godwit (*Limosa haemastica*) differ in spring migration timing. *The Auk*, *129*(4), 670-682.
- Spiller, K. J., & Dettmers, R. (2019). Evidence for multiple drivers of aerial insectivore declines in North America. *The Condor*, *121*(2), duz010.
- Wright, B., Farquharson, K. A., McLennan, E. A., Belov, K., Hogg, C. J., & Grueber, C. E. (2019). From reference genomes to population genomics: Comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics*, *20*(1), 1-10.

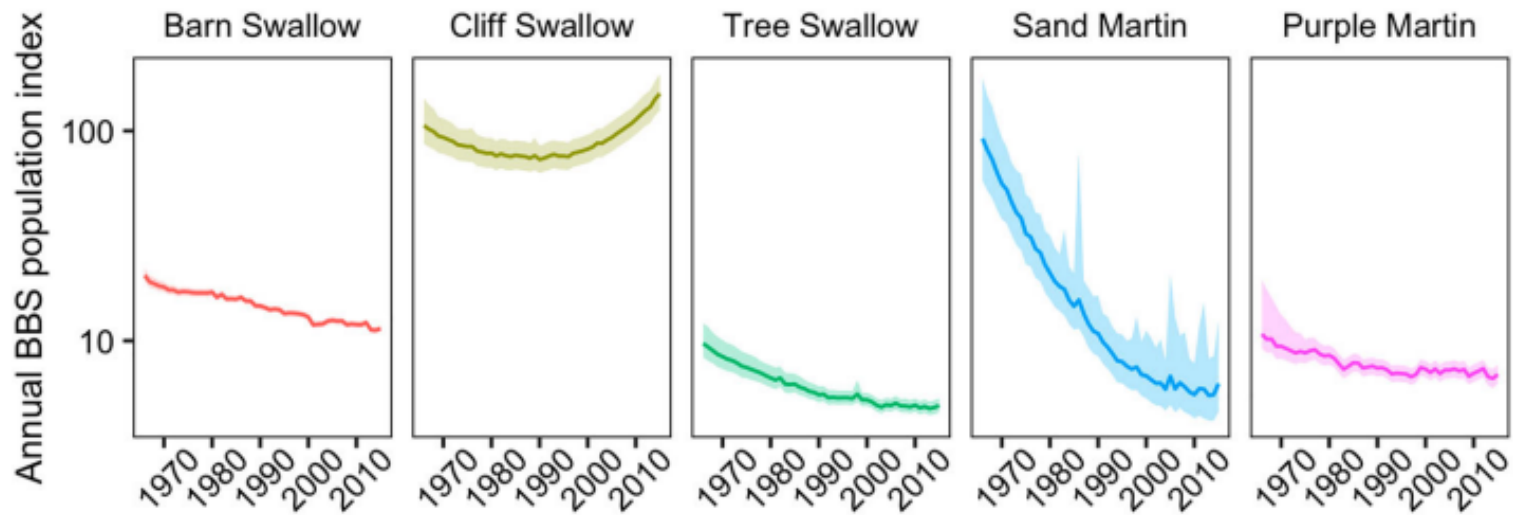


Figure 1.1: The annual population trends of breeding bird survey (BBS) of swallows, represented with a log scale from 1996 – 2010. Sand martin also known as bank swallow in the New World. (Imley & Leonard, 2019).

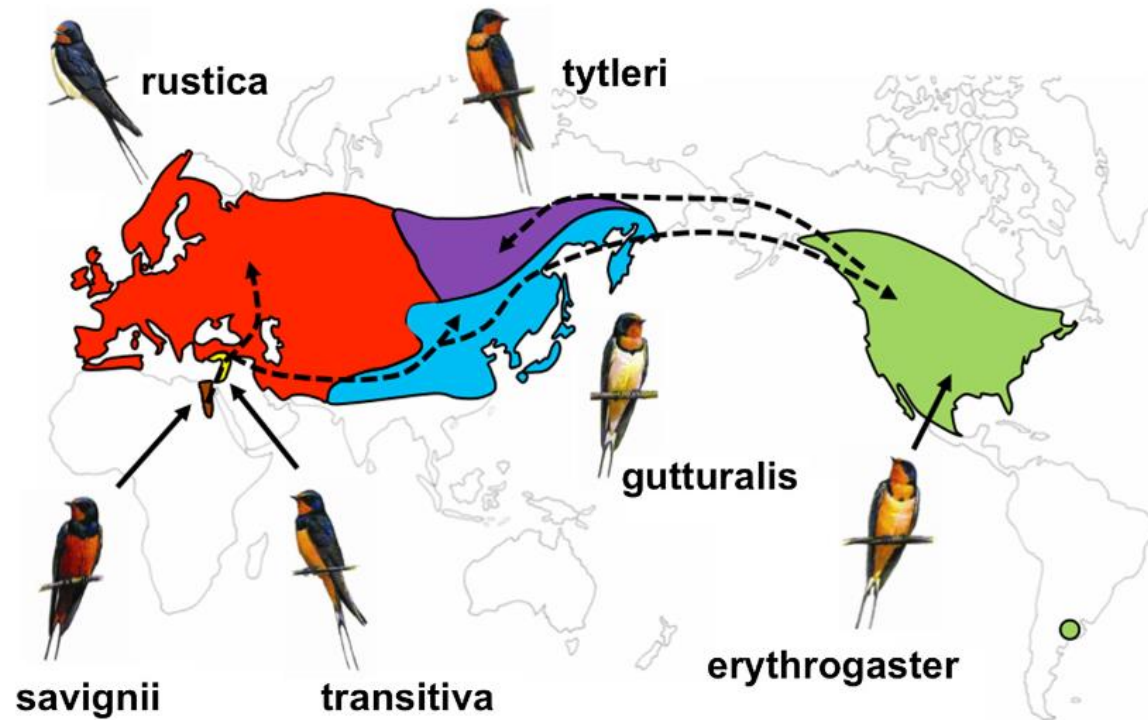


Figure 1.2: Subspecies distribution and differences of plumage coloration of barn swallow (*Hirundo rustica*) subspecies complex. Dash arrows represent hypothesized colonization events (Scordato & Safran, 2014).

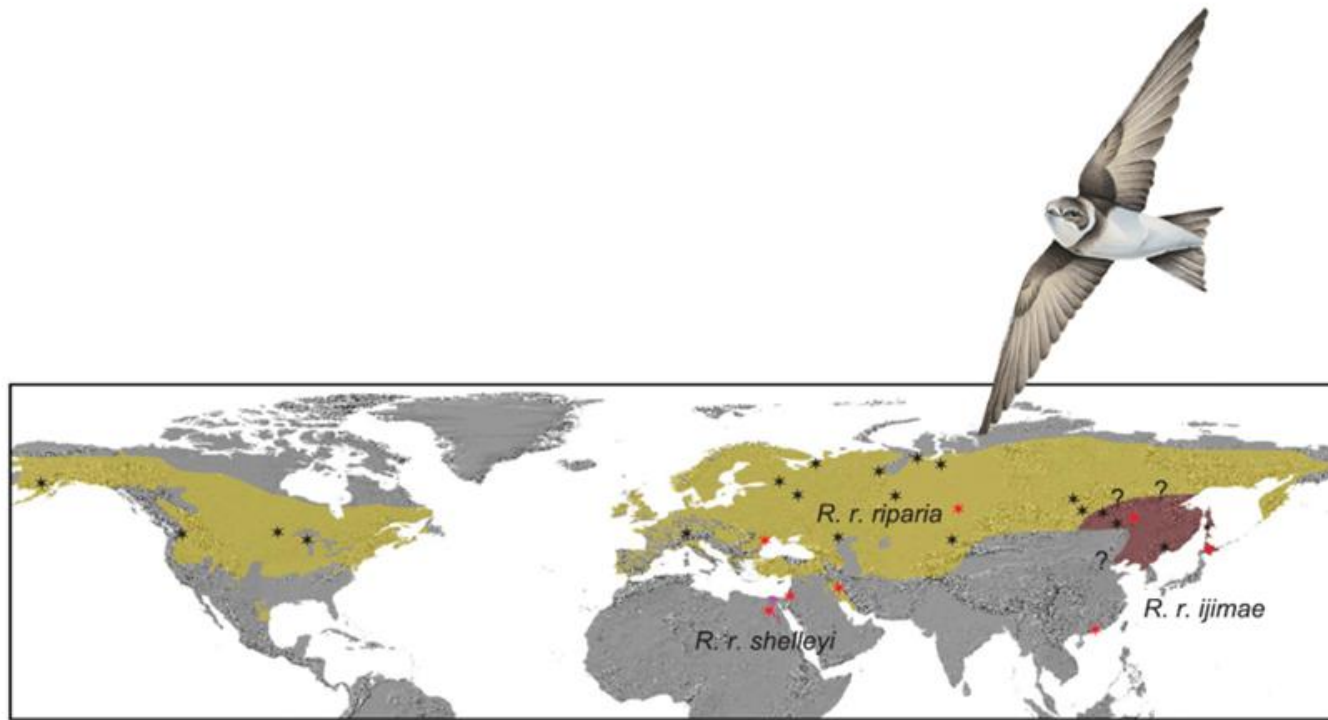


Figure 1.3: Subspecies distribution of bank swallow (*Riparia riparia*) subspecies complex. Red stars indicate the sample locations used in the Schweizer et al. (2018) and black stars indicate samples used in the Pavlova et al. (2008). Hypothetical subspecies boundaries indicated in question marks (Schweizer et al., 2018).

Chapter 2: Landscape Genetic Structure of Barn Swallows Breeding in North America

Abstract

North American aerial insectivores, including the barn swallow (*Hirundo rustica*), have experienced significant population declines of 59% over the past few decades, with the steepest declines observed in northeastern North America. To gain insights into the species' conservation needs, I examined the population genetic structure of barn swallows across the continent. I employed a genomic approach, including microsatellite analysis, restriction-site associated sequencing (RADseq), and low coverage whole genome sequencing data (lcWGS), from 13 sites across North America, encompassing breeding, migratory, and overwintering sites. The microsatellite and RADseq data suggest a certain degree of genetic exchange among populations, while the lcWGS data reveal the presence of three population clusters corresponding to western, central, and eastern North America. The most parsimonious explanation for these patterns is that Pleistocene glaciations, along with associated factors such as migratory pathways and contemporary barriers, play a vital role in shaping genetic differences between populations.

2.1 Introduction

Migration is a prevalent phenomenon observed across the animal kingdom, where many species undertake regular movements between different locations for breeding, foraging, or other ecological purposes (Alerstam et al., 2003). Migratory behaviour is commonly believed to promote gene flow in contrast to nonmigratory species (Ralston & Kirchman, 2013). However, changes in migratory behaviour such as migration patterns, routes, or timing can result in the development of population genetic structure, leading to genetic differentiation among populations (Ralston et al., 2019; Rissler et al., 2004). Migration patterns can vary among closely related species, and even within species, while some species display partial migration whereas other species display differential migration. For instance, the majority of red-breasted nuthatches (*Sitta canadensis*) are year-round residents, but some populations migrate southward to utilize better resources (Harrap and Quinn, 1995; Cannings et al., 1987). On the other hand, species like Wilson's warbler (*Cardellina pusilla*), barn swallows (*Hirundo rustica*) and American redstarts (*Setophaga ruticilla*) are long-distance migrants that exhibit a "leap-frog" migration pattern (Colbeck et al., 2008; Hobson et al., 2015; Kelly et al., 2002).

Differences in migratory behaviour also create migratory divides (Rissler et al., 2004). Migratory divides are observed in many taxa including fish species (e.g.: *Oncorhynchus nerka*), most of the terrestrial bird species (e.g.: *Sylvia atricapilla*, *Tachycineta bicolor*, *Hirundo rustica*), and some butterfly species (e.g.: *Danaus plexippus*) (Hobson et al., 2015; Knight et al., 2018; La Sorte & Fink, 2017; Merlin & Liedvogel, 2019; Turbek et al., 2022). Migratory divides act as barriers to gene flow and enhance the population genetic differences between populations (Delmore et al.,

2012). For instance, long-distance migrants will show morphological adaptations to travel long distances (e.g. longer wings) and adapt to specific conditions that they face during the migration. Biogeographic history such as isolation in distinct glacial refugia correspond to divergent migratory pathways in many birds (Colbeck, 2008; Boulet, & Gibbs, 2006). These diverse pathways are thought to restrict gene flow and maintain reproductive isolation via selection against hybrids (Ruegg, 2008; Irwin, 2005). Hybrids have intermediate migration routes or follow parental routes (Delmore & Irwin, 2014).

Among many North American migratory species, aerial insectivores are a group of great conservation concern as they have experienced an overall population decline of 59% over the last ~60 years (Rosenberg et al., 2019). Climate change, anthropogenic modifications, pesticide usage, decrease insect abundance, and habitat loss are interconnected hypotheses explaining the recent dramatic decline of aerial insectivores (Nebel et al., 2010; Nicole, 2008; Spiller et al., 2019). Implementing conservation plans for small migratory birds is challenging as they spend their annual cycle in different parts of the world. Furthermore, difficulties in tracking small animals and lack of genetic information across migratory divides, particularly the genetic data for migratory and overwintering populations, limit our understanding of how migratory divides restrict gene flow and shape population genetic structure. A combination of migratory, environmental, and genetic data are required to facilitate the development of effective conservation and management plans for declining migratory species.

The barn swallow (*Hirundo rustica*) is an example of a species that exhibits a migratory divide, whereby different populations display distinct migration patterns and occupy geographically separated wintering areas. There are six recognized

subspecies of barn swallows, five subspecies distributed throughout the Old World and only a single subspecies (*H. r. erythrogaster*) present in the New World (i.e., North America) (Dor et al., 2010). Given the large geographic range of barn swallows, they exhibit shallow genetic divergence across subspecies (Dor et al., 2010; Zink et al., 2006). The colonization events after the Last Glacial Maximum (LGM) which occurred ~10,000 years ago in North America and population bottleneck events which occurred ~7,700 years ago when humans started to build structures in North America before the subspecies differentiation and subsequent growth could also be contributing factors (Smith et al., 2018).

In North America, continental-wide migratory divides both latitudinally and longitudinally for barn swallows have been identified using light-level geolocators and stable isotopes (García-Pérez et al., 2014; Hobson et al., 2015). Eastern populations from New Brunswick (NB) and Ontario (ON) travel long distances via the East Coast and Caribbean islands to reach central South America, whereas western populations from Washington (WA) and Saskatchewan (SK) travel relatively short distances to Central America and northwestern South America through western routes (Hobson et al., 2015) (Figure 2.1). Apart from the differences in migratory pathways, barn swallows also exhibit regional variation in decline, with northeastern North America experiencing the highest probability of decline among aerial insectivores, including barn swallows, compared to other parts of North America (Michel et al., 2016; Nebel et al., 2010). The high-risk factors, such as oceanic crossings, associated with eastern flyways might influence the high decline of northeastern populations. Populations with longer migration distances face stronger time-selection pressures when arriving at breeding grounds. The different breeding

times can act as a pre-zygotic barrier to reproductive isolation thus leading to population genetic differences (Bearhop et al., 2005).

In this study, I used microsatellite markers and next-generation sequencing to better understand the genetic structure across the migratory divide of barn swallows in North America. Lower-resolution microsatellite markers and restriction site-associated DNA sequencing (RADseq) were used as preliminary data to investigate population genetic structure, while higher resolution low coverage whole genome sequencing (lcWGS) was utilized to study the fine-scale genetic structure of barn swallow population. In addition to the breeding populations, migratory and overwintering populations were included in the study, where data on such populations were available. I investigated the following questions: 1) Are there population genetic differences across the migratory divide? If so, 2) which populations exhibit low genetic diversity and which loci are under selection? This study also provides insights into the usage of different genetic markers in population assignment studies for migratory species.

2.2 Methods

2.2.1 Sample Collection and DNA Extraction

I obtained feather/blood samples from bird banding stations. Birds were caught using mist nets and either a blood or feather sample was collected. Approximately 50 μ l of blood was collected from the brachial vein or a single tail feather was plucked (Owen, 2011). Then birds were released at the capture sites. The collected samples were stored in ethanol at ambient temperature until they could be returned to the laboratory, where they were stored at -20°C. Frozen tissue samples or skin clips were collected from museums (Beaty Biodiversity Museum, Royal Saskatchewan Museum

and Denver Museum). The sampling period ranged from 1999 to 2022 (see Figure 2.2 for sampling sites and Table 2.1 for detailed sampling information).

DNA extraction for 176 samples (Figure 2.2 and Table 2.1) was carried out using Chelex DNA extraction for microsatellite analysis (Walsh et al., 2013). A modified salting-out extraction procedure was used for $n = 101$ samples (Aljanabi & Martinez, 1997). A subset of these $n = 55$ salting out DNA extractions ($n = 55$) from Washington (WA), North Dakota (ND), Saskatchewan (SK), Ontario (ON), Alabama (AL) and New Brunswick (NB) were used for triple digest restriction site-associated DNA sequencing (RADseq). The remaining, 46 samples were used to perform a low-coverage whole genome sequencing (lcWGS). These individual samples from Washington (WA), California (CA), British Columbia (BC), Saskatchewan (SK), Colorado (CO), Colombia (COL), Manitoba (MB), New Brunswick (NB) and Quebec (QC). The lcWGS was performed to capture the fine-scale genetic structure and address any biases associated with the lower breadth of coverage in RADseq method.

2.2.2 Microsatellite Analyses

A subset of six samples from different geographical locations was used for the initial screening of microsatellite loci before performing the large-scale screening. Initial screening was performed for 27 microsatellite loci to assess variation and DNA amplification. From the 27 loci, five were used to screen all 176 samples: Titgata 39 (Wang et al., 2005), Tall 11 (Makarewich et al., 2009), Ase 18 (Richardson et al., 2000), Hir 20 and Hir 15 (Tsyusk et al., 2007). DNA was amplified in 10 μ l PCR reactions containing 1x FroggABio or GoFlexi buffer, 0.25 mM dNTP, 0.5 μ M forward primer, 1 μ M reverse primer, 0.05 μ M fluorescently labelled M13 tag and 0.5 units of FroggABio or GoFlexi 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 Hir 20, Ase 18, Hir 15 and Titgata 39 primers, the annealing temperatures were T₁=45 °C and T₂=47 °C, for Tal 11 T₁=48 °C and T₂=50 °C. The PCR products were mixed with loading dye, denatured at 94 °C and run on a 6% acrylamide gel on a Li-Cor 4300 DNA analyzer. A set of three positive controls were used for each locus to ensure consistent sizing. Gels were scored independently by two people to ensure reliable scoring. Samples that amplified poorly across multiple loci after two or three attempts were removed from downstream analyses.

Genepop version 4.7 was used to identify the deviations from Hardy-Weinberg equilibrium (HWE). Microsatellite genotypes were analyzed using STRUCTURE v.2.3.4 with 100K burnin followed by 200K runs, the admixture model with *loc priors* and correlated allele frequencies to identify the optimal number of genetic clusters (K). Sampling locations were used as *loc priors* (Pritchard et al., 2000; Raymond & Rousset, 1995). Hierarchical STRUCTURE analysis was performed for 10 iterations from K=1 to K=8. Arlequin 3.5 was used to perform an analysis of molecular variance (AMOVA), and pairwise F_{ST} analysis (Excoffier & Lischer, 2010).

2.2.3 RADseq Analysis (Restriction Site Associated DNA Sequencing)

2.2.3.1 Sequencing and Preprocessing of Data

Three restriction enzymes, *PstI* (6-base pair cutter), *NsiI* (6-base-pair cutter) and *MspI* (4-base-pair cutter), were used to create the reduced representation sequencing library (Peterson et al., 2012). Library preparation for RADseq was performed at the University of Laval following Abed et al. (2019). Samples were sequenced at

G nome Qu bec on an Illumina NovaSeq 6000 S4 PE100 with paired-end sequencing.

Raw reads were demultiplexed based on unique barcodes using the “sabre pe” option available in the Sabre/1.00 program (Andrews, 2010). Before proceeding with the analyses, data quality was evaluated, and sequences were trimmed. Fastqc/0.11.9 was used to generate quality reports. Adapter and barcode sequences were removed from reads using Cutadapt 1.16 and the low-quality sequences trimmed at each end using q30,30 option in the Cutadapt to remove both leading and trailing low-quality bases (Phred scores < 30) (Martin, 2011). Sequences containing less than 100,000 reads were removed. After filtering, chromosomal level assembled barn swallow (*Hirundo rustica*) reference genome from the Vertebrate Genomes Project (GenBank assembly number: GCF_015227805.1) was used to align the sequences using Burrows-Wheeler alignment (bwa) (Li et al., 2009). Custom scripts were used to create SAM files (Li et al., 2009).

Genotyping and variant calling were performed using Stacks/2.3e, a program specifically designed for reduced representation sequencing (Catchen et al., 2013). The resulting variant call format (VCF) file was filtered using VCFtools/0.1.16 (Danecek et al., 2011). SNPs with a Phred score below 30 were filtered out as well as loci with a minor allele count less than three. A filtering scheme was applied and removed SNPs with more than 20% missing data and individuals with more than 60% missing data. The filtered data set contained 850 SNPs and 39 individuals. An additional iterative filtering scheme allowing up to 10% missingness for SNPs and 20% missingness for individuals was used for PCA to ensure accurate SNPs calling. The second data set contained 29 individuals and 604 SNPs. Additionally, I used random a SNP per locus to check the effect of linked SNPs. Since I did not notice any

differences between random SNP per locus and all SNPs, I included results from all SNPs in this study.

2.2.3.2 Population Genetic Analyses

Mean observed and expected heterozygosities were calculated for the entire dataset prior to downstream analyses, using the population step in the Stacks pipeline.

Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) was 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 & Hochberg, 1995). Populations containing less than three individuals were excluded from pairwise F_{ST} calculations to reduce false positives.

The filtered VCF file was used to estimate population genetic structure using principal component analysis (PCA) as implemented by the package Adegenet 2.1.5 in R studio. I used both the tight (604 SNPs, 29 individuals) and less conservative (850 SNPs, 39 individuals) data sets to generate PCA. Based on the eigenvalues which represent the percentage of variance, the first three principal components (x, y and z) were used to plot both data sets.

2.2.4 Low Coverage Whole Genome Sequencing (lcWGS) Analyses

2.2.4.1 lcWGS Sequencing and Mapping

Salting out extractions were sent to the Génome Québec for Shotgun Library preparation step and Illumina NovaSeq PE 150 lcWGS by targeting genomic coverage of ~5X per sample.

After obtaining the sequences, adaptors were removed and quality-trimmed reads using the q30,30 option in cutadapt version 1.16 to eliminate the reads with Phred score below 30 (Martin, 2011). The sequences were then aligned to the *Hirundo*

rustica (barn swallow) reference genome obtained from the Vertebrate Genomes Project (GenBank assembly number: GCF_015227805.1) using bwa alignment tool and sorted using Samtools v1.10 (Li et al., 2009). The PCR duplicates were identified and marked using Picard (Picard, 2018). Furthermore, overlapping regions of the sequences were clipped and insertions and deletions (INDELs) were realigned. Before using sequences for downstream analyses, I validated and checked the read depth of the final binary alignment files to ensure the reliability of the sequences.

2.2.4.2 Genotype Likelihood Estimation

Genotype likelihood data were obtained using ANGSD v0.933 (Analysis of Next Generation Sequencing Data), a program specially designed to analyse lcWGS data by considering low to moderate coverage associated with lcWGS (Korneliussen et al., 2014). I used SAMtools genotype likelihood model to incorporate the uncertainty of true genotypes, as low sequencing coverage approach prohibits genotype calling (Li, 2011; Li et al., 2009). Using ANGSD, I removed low-quality reads (-uniqueOnly 1 -remove_bads 1, -minMapQ 30, -minQ 30), unmapped reads (-only_proper_pairs) and confined to the sites with a P value less than or equal to $1e-6$ (-SNP_pval $1e-6$). Furthermore, I used (-minMaf 0.05) to retain the sites with minor allele frequency above 0.05.

2.2.4.3 Population Structure

The lcWGS dataset was used to perform multivariate analyses and evaluate the population genetic structure of barn swallows. First, principal component analysis (PCA) was performed using PCAngsd which uses genotype likelihoods to build the covariant matrix (Meisner & Albrechtsen, 2018). The resulting covariant matrix was visualized using the RStudio ggplot package. I also performed PCA analysis only for breeding birds to reduce potential confounding factors such as exact location of

migratory and wintering birds. The admixture analysis using NGSadmix was performed for $K = 2 - 4$ to identify the ancestry proportions of each individual (Skotte et al., 2013). The tolerance for considering a site as missing was set to 0.05, the convergence tolerance was set to 0.0001 and a minimum number of informative individuals was set to 20 (~50% of total number of individuals) (Korneliussen et al., 2014). The optimal Evanno K value was identified and visualized admixture plots using CLUMPAK (Kopelman et al., 2015).

2.2.4.4. Comparative Population Analyses

To examine the population-level genetic differences, allele frequencies were calculated using ANGSD based on genotype likelihood data across all populations. Based on the clusters obtained by PCA and admixture plots, individuals were grouped into three different populations, western, central, and eastern, and allele frequencies at all SNPs were calculated using the “- doSaf” option (Nielsen et al., 2012). I performed folded site frequency spectrum using the barn swallow reference genome as the outgroup to polarize the sample allele frequencies and estimate the derived allele frequency. After filtering to remove low quality sites, extreme outliers in sequencing depth, and spurious alignments; the realSFS command was used to generate the folded site frequency spectrum for each population.

To understand the differentiation between west, east and central population clusters and to detect loci under selection, I calculated the fixation index (F_{ST}) using ANGSD. Windowed estimation of F_{ST} was performed using 10 kb windows across the genome with a 1 kb step and visualized through Manhattan plots using R package. The highest F_{ST} values (>0.08) across the genome were selected from the annotated reference genome to look at the genes in those areas.

2.3 Results

2.3.1 Microsatellite Analyses

A total of 176 samples were screened for five polymorphic microsatellite loci (Table 2.1). Genepop did not detect any major deviations from Hardy-Weinberg equilibrium or linkage disequilibrium. The STRUCTURE clustering analysis with all five loci and eight populations did not show any clear population structure. The optimal K value was K=1 and the analyses for K=2 to K=8 did not show any obvious patterns. Furthermore, none of the pairwise F_{ST} values were significant following correction for multiple tests.

The first and second axes of PCoA contain 23.6% and 13.45% variance respectively. The PCoA indicates populations clustering together though the Ontario population showed weak separation from the rest of the group on the coordinate 2 (Figure 2.3).

2.3.2 RADseq Analyses

The “stacks” step in the Stacks pipeline yielded 11,113,762 paired end reads, with a mean coverage of 13.3x per sample. I did not notice any significant difference between expected or observed heterozygosity in any populations ($P > 0.05$) (Table 2.2). From the initial set of 262,632 SNPs and 55 individuals, I obtained 850 SNPs and 39 individuals for downstream analyses. For the principal component analysis, after applying an additional iterative filtering step to handle missing data for individuals, the data set included 29 individuals and 604 SNPs.

Pairwise F_{ST} and associated P values indicated significant differences between ON and SK data $F_{ST} = 0.04148$ ($P = 0.0272$) as well as ON and WA populations $F_{ST} = 0.04432$ ($P = 0.0181$). SK and WA populations do not show significant difference (P

> 0.05). NB, ND, and AL populations were excluded from pairwise F_{ST} comparisons due to sample size ($n < 3$).

The principal component analysis using less stringent filtering (with 20% missingness for SNPs and 60% missingness for individuals) identified two main genetic clusters corresponding to eastern (ON, AL) and western (SK, WA, ND) North America based on the variation explained by the first three components (PC1 9.64%, and PC2 5.39% and PC3 4.7%). I also observed samples from NB separated from the two main clusters (Figure 2.4a). The tight filtering option (10% missingness for SNPs and 20% missingness for individuals) resulted in the exclusion of the entire NB population. Additionally, one individual from ND and two individuals from AL were removed due to high missingness.

The PCA with tight filtering identified two genetic clusters corresponding to ON, AL and the rest of the populations (SK, WA, ND). The first three PCs explained the highest variance (PC1 7.4%, PC2 6.1% and PC3 5.5%); therefore, I focused on plots using the first three axes (Figure 2.4b).

The ancestry matrix using STRUCTURE software did not detect any biologically meaningful genetic clusters at any of the K values.

2.3.3 lcWGS Analysis

The results from lcWGS data show a much clearer resolution of population structure compared to microsatellite and RADseq analyses. I obtained an average of 4.96x coverage per sample and a total number of 16,310,859 sites after the filtering.

2.3.3.1 Population Structure of Barn Swallows

Principal component analysis using the PCAngsd with genotype likelihoods demonstrated two distinct clusters corresponding to the populations west and east of

the Rocky Mountains, as indicated by the first two principal components. PC1 explains 2.55% of the variance, whereas PC2 explains 2.49% of the variance (Figure 2.5a). The genetic variation within western cluster seems to be high compared to the eastern cluster. When considering PC3 (2.41%) and PC4 (2.39%), three distinct clusters were observed corresponding to the west (SBC, WA, and CA), central (SK, MB, migratory populations from SK, MB, CO and wintering population from COL) and eastern (QC and NB) North America (Figure 2.5b). Furthermore, I observed one migratory individual from Colorado intermediate to the western and central clusters. When focusing solely on the analysis of breeding birds, the population structure becomes even clearer, revealing the same three distinct clusters: west (SBC, WA and CA), central (MB, SK) and east (QC, NB) (Figure 2.6).

Using the Evanno method, $K=3$ was identified as the optimal number of clusters in the NGSadmixture analysis. The three clusters predominantly correspond to the west, central and east populations.

2.3.3.2 Comparative Population Analyses

The windowed F_{ST} values between each population comparison identified a few gene regions with comparatively higher F_{ST} values (0.08-0.15) (Figure 2.7). Regions associated with high F_{ST} values and a low π indicate signs of selection. ADCY AP1 (Adenylate Cyclase-Activating Polypeptide 1) and NAV3 genes had higher F_{ST} values in all three comparisons (Table 2.3). ADCY AP1 gene is associated with migratory behaviour in birds and mammals (Mueller et al., 2011). CLOCK (Circadian Locomotor Output Cycles Kaput) gene and NPAS2 (Neuronal PAS Domain Protein 2) are two other important genes identified. Both genes are associated with circadian regulation and circadian behaviour.

In addition, higher F_{ST} values were observed around GYG1, NAE1, NBAS, and LOCI 120754374.

The site frequency spectrum (SFS) for each population gives insights of past demographic history. SFS of the central population indicates a more stationary phase, while western and eastern populations have a smaller proportion of derived alleles compared to the reference genome, indicating low genetic diversity (Figure 2.8).

2.4 Discussion

This study using low-resolution RADseq data shows two broad genetic clusters of barn swallow populations corresponding to the eastern and western regions of North America. With the high-resolution lcWGS data, I was able to identify further substructure within the western cluster, revealing two genetic clusters corresponding to central and western parts of North America. This finding is supported by available migratory data using geolocators and isotope analysis (García-Pérez & Hobson, 2014; Hobson et al., 2015). Historical events, along with the migratory flyways and contemporary barriers are likely playing important roles in shaping population genetic differences of barn swallows. The results of this study will aid in the development of conservation plans and prioritization strategies for the declining barn swallow populations in North America.

Barn swallows are one of the most extensively studied swallow species in terms of behaviour, ecology and genetic component of sexual selection (Møller, 2001; Scordato & Safran, 2014; Scordato et al., 2017; Soler et al., 1998). However, this is the first study to utilize different genetic markers to understand the population genomic structure of North American barn swallows (*H. r. erythrogaster*).

Phylogenetic studies suggest that North American barn swallows originated from

Asian populations that crossed the Bering Strait (Scordato & Safran, 2014). The low genetic differentiation between subspecies of barn swallows supports their recent and rapid expansion (Winkler, 2017). Furthermore, the absence of genetic structure in the neutral markers used in the microsatellite analyses also support high levels of historical gene flow.

The PCA results from RADseq analysis identify two main clusters corresponding to west (WA, SK, ND) and east (ON, AL). The NB population is also separated in PCA (Figure 2.4 a). Separation of the NB population from the rest of the eastern population can be due to restricted gene flow between two eastern populations (NB and ON) by migratory differences and/or any other landscape features. Because migratory pathways from Hobson et al. (2015) indicate that individuals from NB remain along the East Coast, whereas individuals from ON use inland pathways. However, due to the high missingness of the sequences of NB samples and the small sample size, I did not make any strong inferences for this separation. To eliminate the possible effects of missing data, I used a second more stringent filtering scheme in addition to the main filtering scheme. The second filtering scheme for PCA analysis using RADseq filtered entire NB population and a few individuals from ND and AL, but this tight filtering approach significantly enhances our ability to discern potential population structure by utilizing a more refined and accurate dataset. High-quality sequencing or computational intense imputation programs might be useful to draw conclusions about these populations. Additionally, comparatively low observed heterozygosity in the NB population in the variant sites is an important consideration in terms of conservation efforts, particularly because the northeastern population is experiencing a more pronounced decline compared to other regions (Michel et al., 2016, Nebel et al., 2010).

The lcWGS analyses display fine scale genetic structure in barn swallows. In addition to the west and east clusters, I noticed an additional structure within the western population i.e., populations west to the Rocky Mountains; SBC, WA and CA cluster separately from the populations breeding in central regions; SK and MB of North America. Within the eastern cluster unlike RADseq analysis, I did not observe isolation of NB population from the other eastern population (QC) in the PCA. If there is any intra-population structure present within the eastern populations, the inclusion of ON samples for lcWGS would be useful in future studies.

2.4.1 Identifying Genetic Structure Using Different Genetic Markers

Various markers used in this study have differing levels of resolution to identify the genetic structure and can aid in understanding the effect of selection and genetic drift. For instance, microsatellite markers are typically neutral markers that have a low breadth of coverage and therefore lower resolution compared to next-generation sequencing. In this study, results from microsatellite markers do not show much genetic structure, as a few markers throughout the genome were insufficient to capture the fine-scale genetic structure within a subspecies (Fola et al., 2020). The next-generation sequencing approach that I used, RADseq and lcWGS, both detected finer scale genetic structure. Even though, RADseq approach plays an important role in the fields of genomics, conservation, and evolution by facilitating the examination of genomes, incomplete genome coverage limits the ability to detect important signals and can miss signatures of selection and adaptive divergence (Lowry et al., 2017). Therefore, I also used lcWGS to get better resolution. My results demonstrate that the resolution of genetic analysis can vary depending on the markers employed.

2.4.2 Pleistocene Glaciations Shape the Population Genetic Structure

The most likely explanation for the observed splits of west, central and east populations in the North American barn swallow is the Pleistocene glaciations. The genetic structure observed in many other North American species provides evidence of past glaciation events and expansions from ice free refugia south of the ice sheets, along the coasts and Beringia (Hewitt, 2004). Unlike many species of birds, barn swallows are less likely to be associated with boreal forests. Given their habitats and diets, they are mostly associated with open habitats and feed on flying insects. Historically barn swallows nested in caves and cliffs and by the early 1800s had begun to adapt to manmade structures and rarely use natural nest sites today (Zink et al., 2006). Considering refugia for insect species, possible refugia for barn swallows during LGM could include Beringia, southern or Pacific Northwest refugia in the west and southeast or Atlantic coastal refugia in the east (Pedersen et al., 2016; Shafer et al., 2010). It is most likely that they survived in three different refugia, two in the west/central and one in the east, corresponding to their flyways and based on my results. Prolonged separation followed by the secondary contact shaped the intraspecific divergence that I identified among three populations. Many other bird studies have reported Beringia and southern refugia as possible refugia during the LGM (Burg et al., 2006; Shafer et al., 2010). Furthermore, song sparrows (*Melospiza melodia*), American redstarts (*Setophaga ruticilla*) studies provide examples for Atlantic coastal refugia (Colbeck et al., 2008; Zink et al., 1993).

The Pleistocene glaciations also influenced range shifting and shaping the migratory pathways (Thorup et al., 2021). Birds and other species were forced to adapt and find alternative pathways to reach their breeding and wintering grounds. Some

species/populations may have followed newly exposed land routes, while others may have utilized coastal regions that remained ice-free.

2.4.3 Role of Migratory Divides and Contemporary Barriers

Previous studies explored genetic differences in other barn swallow subspecies breeding in Asia (*H. r. rustica* and *H. r. gutturalis*) and revealed that the migratory divide maintains subspecies boundaries by preventing gene flow between populations (Turbek et al., 2022). East-west migratory divide is also reported for barn swallows in Europe and Africa (Ambrosini, 2009). Similar genetic differences based on migratory patterns are found in other species of birds, butterflies, some amphibians and fish (Liedvogel et al., 2011). The diverse migratory patterns observed within species might have been influenced by the inhospitable geographic features, and habitats.

My results also provide evidence that geographical barriers, such as the Rocky Mountains, Appalachian Mountains and Great Lakes play important roles in maintaining the population structure of barn swallows, mainly by influencing their migratory routes. Many other species provide examples of differences around these geographical barriers in terms of migration/flyways or population genetics (Colbeck et al., 2008; Hindley et al., 2018; Knight et al., 2018; Lait & Burg, 2013).

2.4.4 Migratory Connectivity

When I include the migratory and wintering populations in the analysis, it becomes evident that migratory individuals from SK and MB, as well as wintering individuals from Colombia cluster together, indicating a high migratory connectivity within the central population. However, the western cluster indicates an increased level of genetic diversity, as indicated by the results from PCA analysis.

It is important to note that, I included both migratory and overwintering individuals in order to gain insights into the genetic connectivity between migratory and breeding populations and potential gene flow across the migratory divide. However, during the sampling of individuals in the migratory season, I was unable to determine their exact breeding population. For an example, in PCA analysis (Figure 2.5), I observed one migratory individual from Colorado cluster with the western population. That individual could belong to either the central or western population. Given that Colorado is in the southeast Rocky Mountains, it is also possible that the collected migratory sample from Colorado is bird breeding to the west of the Rocky Mountains. The possibility of gene flow between the western and central populations is supported by examining the migratory pathways outlined in Hobson et al. (2015) and considering the wintering grounds of the central and western populations. In addition to the major migratory pathways, it is possible that long distance migrants may utilize alternate routes to reach their wintering grounds. For instance, Knight et al. (2018) showed that tree swallows exhibit intermixing, yet they identified three distinct major flyways. Similarly, in my study, I observed intermixing between populations in the admixture analysis, but I observed a strong separation of three (west, central, and east) clusters.

2.4.5 Genomic Scans Provide Functional Insights

The genome-wide scans using F_{ST} analysis and lcWGS data identified several gene regions under selection. From the identified genes (Table 2.3), CLOCK gene, ADCYAP1, NPAS2 regions are associated with the circadian clock and migratory behaviour (Mueller et al., 2011). Previous barn swallow studies identified low levels of polymorphism in the CLOCK gene, combined with related differences in migratory traits have a negative effect on late breeding (Caprioli et al., 2012). Studies

of great tits (*Parus major*), and blue tits (*Cyanistes caeruleus*) further support CLOCK gene variation associated with breeding phenological variation for seasonal timing of reproduction (Liedvogel et al., 2009). The CLOCK gene is mainly responsible for circadian oscillator mechanisms (Caprioli et al., 2012) and polymorphism within its polyglutamine-rich region (Poly-Q) stems from its prevalence across both inter- and intra-population contexts. This genetic variability holds the potential to influence the timing of seasonal activities, shaping phenological responses. Noteworthy is the identification of a latitudinal gradient in CLOCK allele size at the inter-population level, evident in select migratory bird and fish species, wherein longer alleles become more frequent as one moves from south to north. Concurrently, at the intra-population level, the CLOCK polymorphism may contribute to an individual's adaptability to changes in photoperiod (Bazzi et al., 2016). Studies of blackcaps (*Sylvia atricapilla*) indicate a range of variations in migratory patterns, and they suggest the ADCYAP1 gene is a candidate gene which controls the expression of different migratory behaviour (Mueller et al., 2011). ADCY AP1 gene encodes neuropeptide known as PACAP (pituitary adenylate cyclase-activating polypeptide) which is important to foster physiological and behavioural shifts associated to migration (Mueller et al., 2011). Many studies suggest that microsatellite polymorphism associated with variation in migratory restlessness and migratory distance (Bazzi et al., 2016; Peterson et al., 2013; Mueller et al., 2011). NAV3 (Neuron Navigator 3) is recognized as a gene required for multiple aspects of fibrogenesis, including actin polymerization linked to cell migration and sustained activation of the mechanosensitive transcription (Raza et al., 2021). However, most of the available literature identifies NAV3 gene function in mammals including humans and mice (Maes et al., 2002), not birds.

The GYG1 gene is involved in glucose metabolism, glycogen synthesis and it was identified as being under positive selection for diverse species including hummingbirds, sunbirds, and lorikeets (Osipova, 2022). The NAE1 gene is required during cellular stress caused by infections to protect against cell death (Muffels et al., 2023). The NBAS gene codes for a protein which forms a subunit of the NZR complex and plays a role in the membrane transport in cells (Claeys et al., 2021). I did not find specific functions for NAE1, NBAS and LOCI 120754374 genes in birds. The higher F_{ST} values associated with my data near these regions suggest possible candidate genes responsible for the migration differences in barn swallows.

2.5 Conclusions

In conclusion, this study found evidence of population genetic structure in barn swallows breeding on either side of migratory divides, and an additional genetic cluster identified in the central region. My findings provide insights into different migratory traits, particularly effects associated with different migratory routes that act as barriers to gene flow. The strong population differentiation observed in the eastern population is another major finding of this study and that would be very useful in prioritizing conservation efforts for this declining barn swallow population especially given large decline in the region.

References

- Abed, A., Légaré, G., Pomerleau, S., St-Cyr, J., Boyle, B., & Belzile, F. J. (2019). Genotyping-by-sequencing on the ion torrent platform in barley. *Barley: Methods in Molecular Biology*, 1900, 233-252.
- Alerstam, T., Hedenström, A., & Åkesson, S. (2003). Long-distance migration: Evolution and determinants. *Oikos*, 103(2), 247-260.
- Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25(22), 4692-4693.
- Bazzi, G., Galimberti, A., Hays, Q. R., Bruni, I., Cecere, J. G., Gianfranceschi, L., ... & Rubolini, D. (2016). Adcyap1 polymorphism covaries with breeding latitude in a Nearctic migratory songbird, the Wilson's warbler (*Cardellina pusilla*). *Ecology and Evolution*, 6(10), 3226-3239.
- Bearhop, S., Fiedler, W., Furness, R. W., Votier, S. C., Waldron, S., Newton, J., Bowen, G. J., Berthold, P., & Farnsworth, K. (2005). Assortative mating as a mechanism for rapid evolution of a migratory divide. *Science*, 310(5747), 502-504.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300.
- Boulet, M., & Gibbs, H. L. (2006). Lineage origin and expansion of a Neotropical migrant songbird after recent glaciation events. *Molecular Ecology*, 15(9), 2505-2525.
- Cannings, R. A., Cannings, R. J., & Cannings S. G. (1987). Birds of the Okanagan Valley, British Columbia. Royal British Columbia Museum, Canada.
- Caprioli, M., Ambrosini, R., Boncoraglio, G., Gatti, E., Romano, A., Romano, M., Rubolini, D., Gianfranceschi, L., & Saino, N. (2012). Clock gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PloS One*, 7(4), e35140.
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124-3140.
- Claeys, L., Storoni, S., Eekhoff, M., Elting, M., Wisse, L., Pals, G., Bravenboer, N., Mageri, A., & Micha, D. (2021). Collagen transport and related pathways in osteogenesis imperfecta. *Human Genetics*, 140(8), 1121-1141.
- Colbeck, G. J., Gibbs, H. L., Marra, P. P., Hobson, K., & Webster, M. S. (2008). Phylogeography of a widespread North American migratory songbird (*Setophaga ruticilla*). *Journal of Heredity*, 99(5), 453-463.
- Delmore, K. E., & Irwin, D. E. (2014). Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, 17(10), 1211-1218.

- Dor, R., Safran, R. J., Sheldon, F. H., Winkler, D. W., & Lovette, I. J. (2010). Phylogeny of the genus *Hirundo* and the barn swallow subspecies complex. *Molecular Phylogenetics and Evolution*, 56(1), 409-418.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611-2620.
- Excoffier, L., & Lischer, H. E. (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.
- Fola, A. A., Kattenberg, E., Razook, Z., Lautu-Gumal, D., Lee, S., Mehra, S., Bahlo, M., Kazura, J., Robinson, L. J., & Laman, M. (2020). SNP barcodes provide higher resolution than microsatellite markers to measure *Plasmodium vivax* population genetics. *Malaria Journal*, 19(1), 1-15.
- García-Pérez, B., & Hobson, K. A. (2014). A multi-isotope ($\delta^{2}\text{H}$, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$) approach to establishing migratory connectivity of barn swallow (*Hirundo rustica*). *Ecosphere*, 5(2), 1-12.
- García-Pérez, B., Hobson, K. A., Albrecht, G., Cadman, M. D., & Salvadori, A. (2014). Influence of climate on annual survival of barn swallows (*Hirundo rustica*) breeding in North America. *The Auk: Ornithological Advances*, 131(3), 351-362.
- Harrap, S., and D. Quinn. (1995). Chickadees, tits, nuthatches and treecreepers. Princeton University Press, Princeton.
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183-195.
- Hindley, J. A., Graham, B. A., & Burg, T. M. (2018). The influence of latitude, geographic distance, and habitat discontinuities on genetic variation in a high latitude montane species. *Scientific Reports*, 8(1), 1-11.
- Hobson, K. A., Kardynal, K. J., Van Wilgenburg, S. L., Albrecht, G., Salvadori, A., Cadman, M. D., Liechti, F., & Fox, J. W. (2015). A continent-wide migratory divide in North American breeding barn swallows (*Hirundo rustica*). *PloSone*, 10(6), e0129340.
- Irwin, D. E., Irwin, J. H., Greenberg, R., & Marra, P. (2005). Siberian migratory divides: The role of seasonal migration in speciation. *Birds of Two Worlds: The Ecology and Evolution of Migration*, 27, 40.
- Kelly, J. F., Atudorei, V., Sharp, Z. D., & Finch, D. M. (2002). Insights into Wilson's Warbler migration from analyses of hydrogen stable-isotope ratios. *Oecologia*, 130, 216-221.
- Knight, S. M., Bradley, D. W., Clark, R. G., Gow, E. A., Bélisle, M., Berzins, L. L., Blake, T., Bridge, E. S., Burke, L., & Dawson, R. D. (2018). Constructing and evaluating a continent-wide migratory songbird network across the annual cycle. *Ecological Monographs*, 88(3), 445-460.

- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across k. *Molecular Ecology Resources*, *15*(5), 1179-1191.
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). Angsd: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 1-13.
- Lait, L. A., & Burg, T. M. (2013). When east meets west: Population structure of a high-latitude resident species, the boreal chickadee (*Poecile hudsonicus*). *Heredity*, *111*(4), 321-329.
- La Sorte, F. A., & Fink, D. (2017). Migration distance, ecological barriers and en-route variation in the migratory behaviour of terrestrial bird populations. *Global Ecology and Biogeography*, *26*(2), 216-227.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987-2993.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and samtools. *Bioinformatics*, *25*(16), 2078-2079.
- Liedvogel, M., Åkesson, S., & Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology & Evolution*, *26*(11), 561-569.
- Liedvogel, M., Szulkin, M., Knowles, S. C., Wood, M. J., & Sheldon, B. C. (2009). Phenotypic correlates of clock gene variation in a wild blue tit population: Evidence for a role in seasonal timing of reproduction. *Molecular Ecology*, *18*(11), 2444-2456.
- Lou, R. N., Jacobs, A., Wilder, A. P., & 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., & 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.
- Maes, T., Barceló, A., & Buesa, C. (2002). Neuron navigator: A human gene family with homology to unc-53, a cell guidance gene from *Caenorhabditis elegans*. *Genomics*, *80*(1), 21-30.
- Makarewich, C. A., Stenzler, L. M., Ferretti, V., Winkler, D. W., & Lovette, I. J. (2009). Isolation and characterization of microsatellite markers from three species of swallows in the genus *Tachycineta*: *T. albilinea*, *T. bicolor* and *T. leucorrhoa*. *Molecular Ecology Resources*, *9*(2), 631-635.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 10-12.
- Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, *210*(2), 719-731.

- Merlin, C., & Liedvogel, M. (2019). The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. *Journal of Experimental Biology*, 222(Suppl_1), jeb191890.
- Michel, N. L., Smith, A. C., Clark, R. G., Morrissey, C. A., & Hobson, K. A. (2016). Differences in spatial synchrony and interspecific concordance inform guild-level population trends for aerial insectivorous birds. *Ecography*, 39(8), 774-786.
- Møller, A., Garamszegi, L., Peralta-Sánchez, J., & Soler, J. (2011). Migratory divides and their consequences for dispersal, population size and parasite–host interactions. *Journal of Evolutionary Biology*, 24(8), 1744-1755.
- Møller, A. P. (2001). The effect of dairy farming on barn swallow *Hirundo rustica* abundance, distribution and reproduction. *Journal of Applied Ecology*, 38(2), 378-389.
- Mueller, J. C., Pulido, F., & Kempenaers, B. (2011). Identification of a gene associated with avian migratory behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 278(1719), 2848-2856.
- Muffels, I. J., Schene, I. F., Rehmann, H., Massink, M. P., van der Wal, M. M., Bauder, C., Labeur, M., Armando, N. G., Lequin, M. H., & Houben, M. L. (2023). Bi-allelic variants in *nae1* cause intellectual disability, ischiopubic hypoplasia, stress-mediated lymphopenia and neurodegeneration. *The American Journal of Human Genetics*, 110(1), 146-160.
- Nebel, S., Mills, A., McCracken, J., & Taylor, P. (2010). Declines of aerial insectivores in North America follow a geographic gradient. *Avian Conservation and Ecology*, 5(2), 1.
- Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS ONE*, 7(7), e37558.
- Osipova, P. D. E. (2022). *Genomic adaptations to a high sugar diet in birds*. Doctoral dissertation, der Technischen Universität Dresden.
- Owen, J. C. (2011). Collecting, processing, and storing avian blood: A review. *Journal of Field Ornithology*, 82(4), 339-354.
- Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L., Beaudoin, A. B., Zutter, C., & Larsen, N. K. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature*, 537(7618), 45-49.
- Peterson, M. P., Abolins-Abols, M., Atwell, J. W., Rice, R. J., Milá, B., & 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.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.

- Ralston, J., & 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*, 543-555.
- Ralston, J., Lorenc, L., Montes, M., DeLuca, W. V., Kirchman, J. J., Woodworth, B. K., Mackenzie, S. A., Newman, A., Cooke, H. A., & Freeman, N. E. (2019). Length polymorphisms at two candidate genes explain variation of migratory behaviors in blackpoll warblers (*Setophaga striata*). *Ecology and Evolution*, *9*(15), 8840-8855.
- Raymond, M., & Rousset, F. (1995). Genepop (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, *86*(3), 248-249.
- Raza, S., Jokl, E., Pritchett, J., Martin, K., Su, K., Simpson, K., Birchall, L., Mullan, A. F., Athwal, V. S., & Doherty, D. T. (2021). Sox9 is required for kidney fibrosis and activates nav3 to drive renal myofibroblast function. *Science Signaling*, *14*(672), eabb4282.
- Richardson, D., Jury, F., Dawson, D., Salgueiro, P., Komdeur, J., & Burke, T. (2000). Fifty seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in *Sylviidae* species and their cross-species amplification in other passerine birds. *Molecular Ecology*, *9*(12), 2226-2231.
- Rissler, L. J., Wilbur, H. M., & Taylor, D. R. (2004). The influence of ecology and genetics on behavioral variation in salamander populations across the eastern continental divide. *The American Naturalist*, *164*(2), 201-213.
- Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., & Parr, M. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120-124.
- Ruegg, K. (2008). Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution*, *62*(2), 452-466.
- Schneider, N. A. (2008). Effects of climate change on avian life history and fitness. *Introductory Research Essay (Department of Ecology, SLU)*, *4*, 1-17.
- Scordato, E. S., & Safran, R. J. (2014). Geographic variation in sexual selection and implications for speciation in the barn swallow. *Avian Research*, *5*(8), 1-13.
- Scordato, E. S., Wilkins, M. R., Semenov, G., Rubtsov, A. S., Kane, N. C., & Safran, R. J. (2017). Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Molecular Ecology*, *26*(20), 5676-5691.
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693-702.
- Smith, C. C., Flaxman, S. M., Scordato, E. S., Kane, N. C., Hund, A. K., Sheta, B. M., & Safran, R. J. (2018). Demographic inference in barn swallows using whole-genome data shows signal for bottleneck and subspecies differentiation during the Holocene. *Molecular Ecology*, *27*(21), 4200-4212.

- Soler, J. J., Cuervo, J. J., Møller, A. P., & De Lope, F. (1998). Nest building is a sexually selected behaviour in the barn swallow. *Animal Behaviour*, 56(6), 1435-1442.
- Spiller, K. J., & Dettmers, R. (2019). Evidence for multiple drivers of aerial insectivore declines in North America. *The Condor*, 121(2), duz010.
- Tsyusko, O. V., Peters, M. B., Hagen, C., Tuberville, T. D., Mousseau, T. A., Møller, A. P., & Glenn, T. C. (2007). Microsatellite markers isolated from barn swallows (*Hirundo rustica*). *Molecular Ecology Notes*, 7(5), 833-835.
- Turbek, S. P., Schield, D. R., Scordato, E. S., Contina, A., Da, X. W., Liu, Y., Liu, Y., Pagani-Núñez, E., Ren, Q. M., & Smith, C. C. (2022). A migratory divide spanning two continents is associated with genomic and ecological divergence. *Evolution*, 76(4), 722-736.
- Walsh, P. S., Metzger, D. A., & Higuchi, R. (2013). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, 54(3), 134-139.
- Wang, M. T., Hsu, Y. C., Yao, C. T., & Li, S. H. (2005). Isolation and characterization of 12 tetranucleotide repeat microsatellite loci from the green-backed tit (*Parus monticolus*). *Molecular Ecology Notes*, 5(2), 439-442.
- Winkler, D.W., Gandoy, F.A., Areta, J.I., Iliff, M.J., Rakhimberdiev, E., Kardynal, K.J. & Hobson, K.A., 2017. Long-distance range expansion and rapid adjustment of migration in a newly established population of barn swallows breeding in Argentina. *Current Biology*, 27(7), 1080-1084.
- Zink, R. M., Pavlova, A., Rohwer, S., & Drovetski, S. V. (2006). Barn swallows before barns: Population histories and intercontinental colonization. *Proceedings of the Royal Society B: Biological Sciences*, 273(1591), 1245-1251.
- Zink, R. M., & Dittmann, D. L. (1993). Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution*, 47(3), 717-729.

Table 2.1: Sampling locations and number of samples utilized in each analysis after filtering: microsatellite (Msat), restriction site associated DNA sequencing (RADseq) and low coverage whole genome sequencing (lcWGS).

Population ID	Number of Samples			Geographical Cluster	Stage of Annual Cycle
	Msat	RADseq	lcWGS		
Washington (WA)	0	4	5	West	Breeding
Southern British Columbia (SBC)	12	0	4	West	Breeding
California (CA)	0	0	2	West	Breeding
North Dakota (ND)	0	2	0	Central	Breeding
Saskatchewan (SK)	9	7	5	Central	Breeding
	0	0	3	Central	Migration
Manitoba (MB)	20	0	1	Central	Breeding
	20	0	7	Central	Migration
Colorado (CO)	0	0	4	Central	Migration
Texas (TX)	4	0	0	Central	Breeding
Colombia (COL)	12	0	5	Central	Wintering
Alabama (AL)	0	3	0	East	Breeding
Ontario (ON)	85	17	0	East	Breeding
Quebec (QC)	7	0	5	East	Breeding
New Brunswick (NB)	7	6	5	East	Breeding
Total	176	39	46		

Table 2.2: Comparison of observed and (H_O) and expected (H_E) heterozygosity for each population using RADseq analysis.

Population	Number of Individuals	All sites			Variable sites only		
		Number of positions	H_E	H_O	Number of SNPs	H_E	H_O
AL	3	28,721,904	0.00052	0.00077	32,025	0.06605	0.09828
NB	6	17,002,302	0.00027	0.00025	9,957	0.03561	0.03373
ON	17	41,452,582	0.0014	0.00101	250,245	0.22083	0.16015
SK	7	41,452,441	0.0009	0.00063	118,311	0.1424	0.09987
ND	2	39,804,376	0.00003	0.00005	2,703	0.01646	0.02485
WA	4	25,819,530	0.0004	0.00039	24,078	0.05149	0.04936

Table 2.3: Genes of interest from genome scans using lcWGS between each population comparison.

West vs. East			
Scaffold	Chromosome	Candidate Gene	Function
NC_053488.1	Z		
NC_053450.1	1	ADCYAP1	Encodes PACAP neuropeptide which could promote physiological and behavioural shifts related to migration
NC_053451.1	2	NPAS2	Circadian behaviour
NC_053452.1	3	NBAS	NBAS subunit of NRZ tethering complex, gene encodes a protein with two leucine zipper domains
NC_053453.1	4	NAV3	Neuron navigator 3 isoform X7, axon
NC_053454.1	5	CLOCK	Clock circadian regulator
NC_053455.1	6	LOC120754374	Uncharacterized
West vs. Central			
Scaffold	Chromosome	Candidate Gene	Function
NC_053450.1	1	ADCYAP1	Expression of migratory behaviour
NC_053451.1	2	NPAS2	Circadian behaviour
NC_053453.1	4	NAV3	Neuron navigator 3 isoform X7
NC_053455.1	6	LOC120754374	Uncharacterized
NC_053460.1	11	NAE1	NEDD8-activating enzyme E1 regulatory subunit isoform and responsible for signal-transducing properties
East vs. Central			
Scaffold	Chromosome	Candidate Gene	Function
NC_053488.1	Z		
NC_053450.1	1	ADCYAP1	Expression of migratory behaviour
NC_053453.1	4	NAV3	Neuron navigator 3 isoform X7
NC_053454.1	5	CLOCK	Clock circadian regulator
NC_053459.1	10	GYG1	Catalyzes the formation of a short glucose polymer from uridine

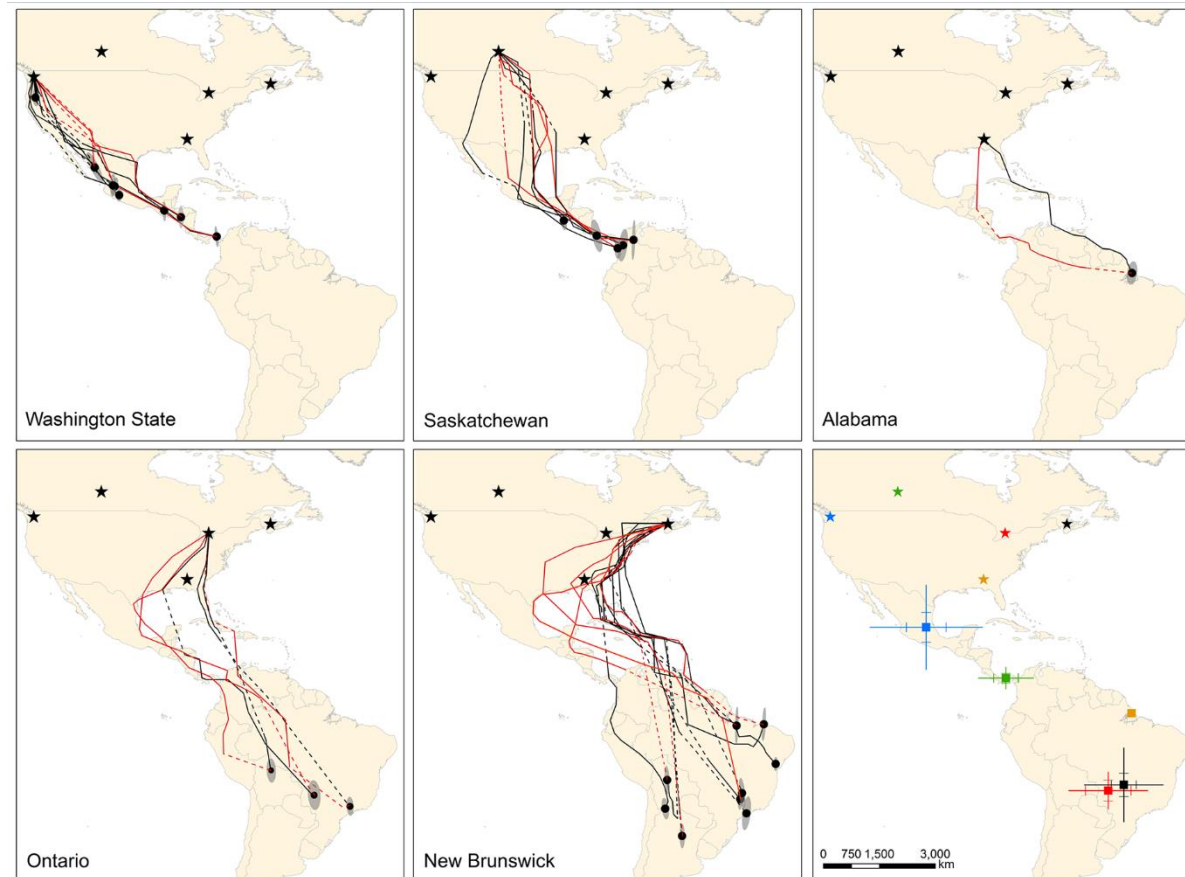


Figure 2.1: Estimated migration routes and wintering sites for barn swallows breeding in North America using archival light-level geolocators (Hobson et al., 2015).

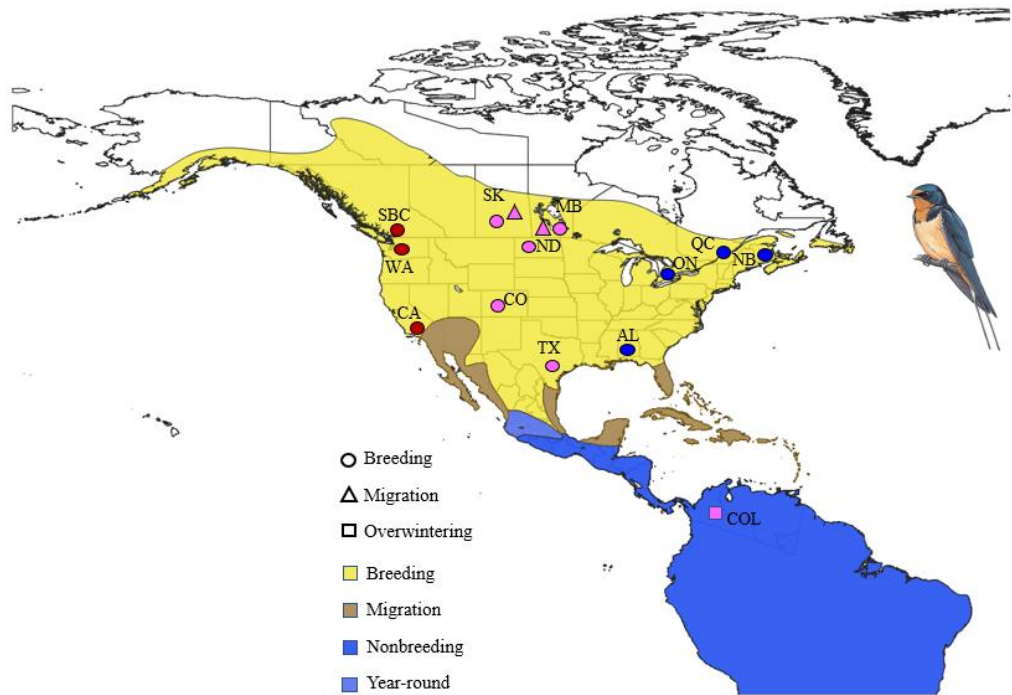


Figure 2.2: Range map of barn swallows in North America and sample locations used in this study. Range data obtained from BirdLife International, and map created using QGIS 3.2.

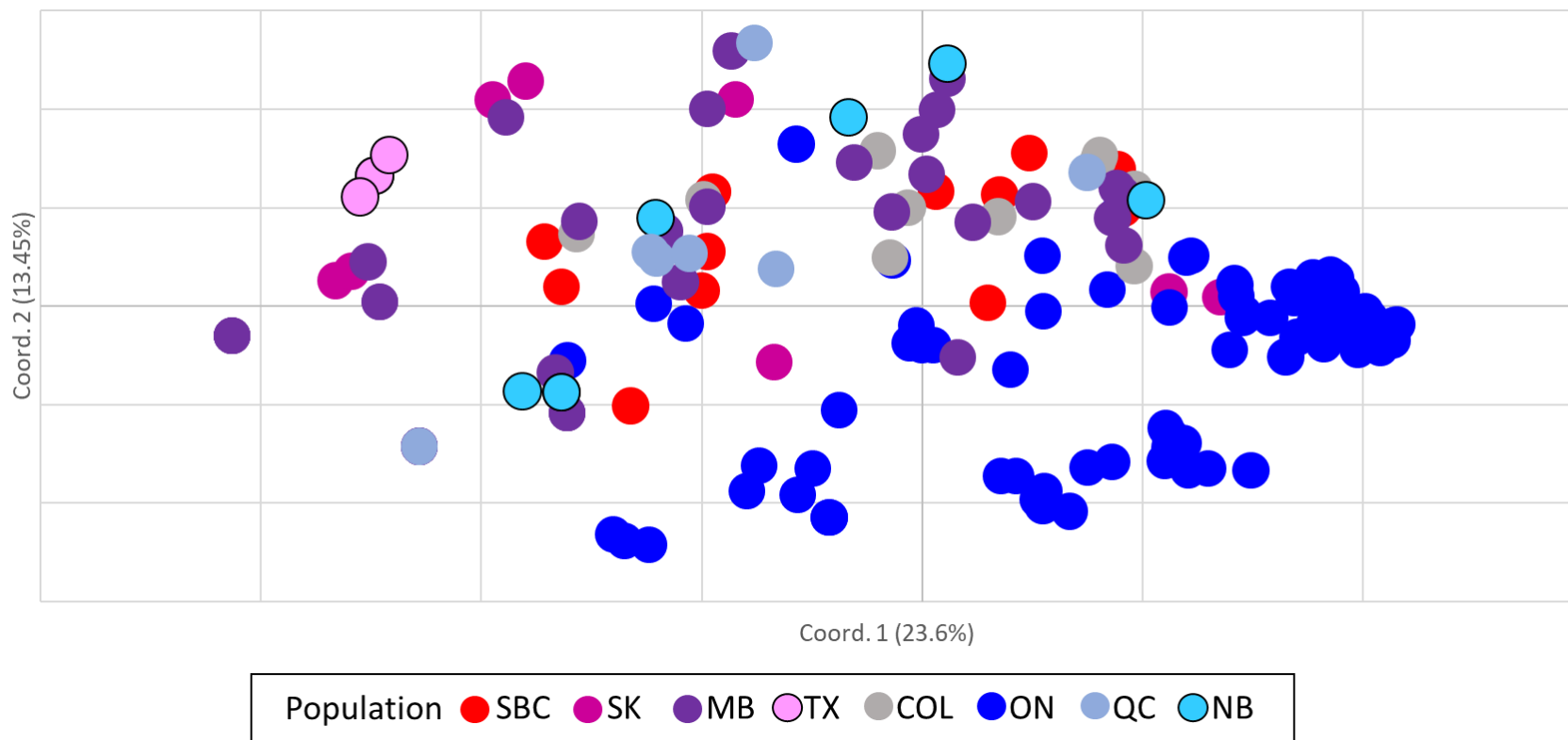


Figure 2.3: Principal coordinate analysis (PCoA) of eight populations using six microsatellite markers.

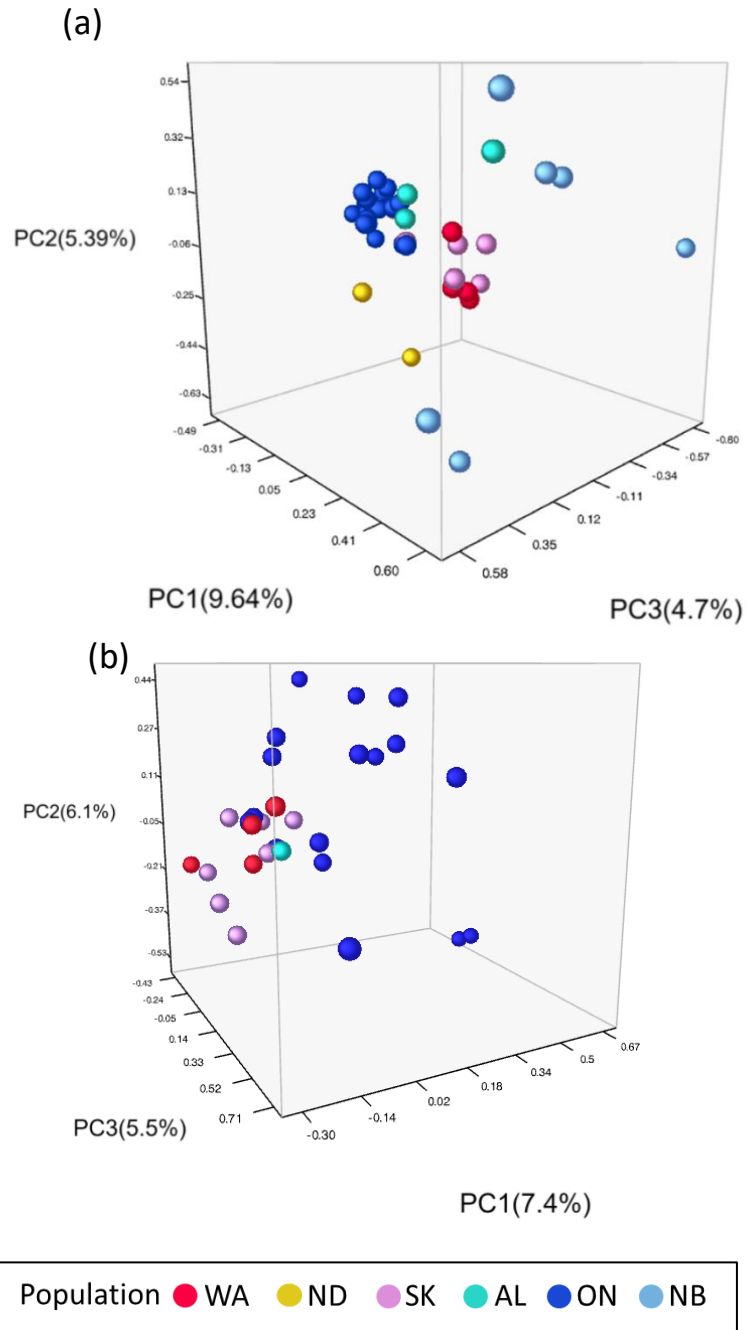


Figure 2.4: Principal Component Analysis (PCA) indicating (a) first three components with loose filtering options including 39 individuals and 850 SNPs (b) first three components for tight filtering option including 29 individuals and 604 SNPs using RADseq data for barn swallows.

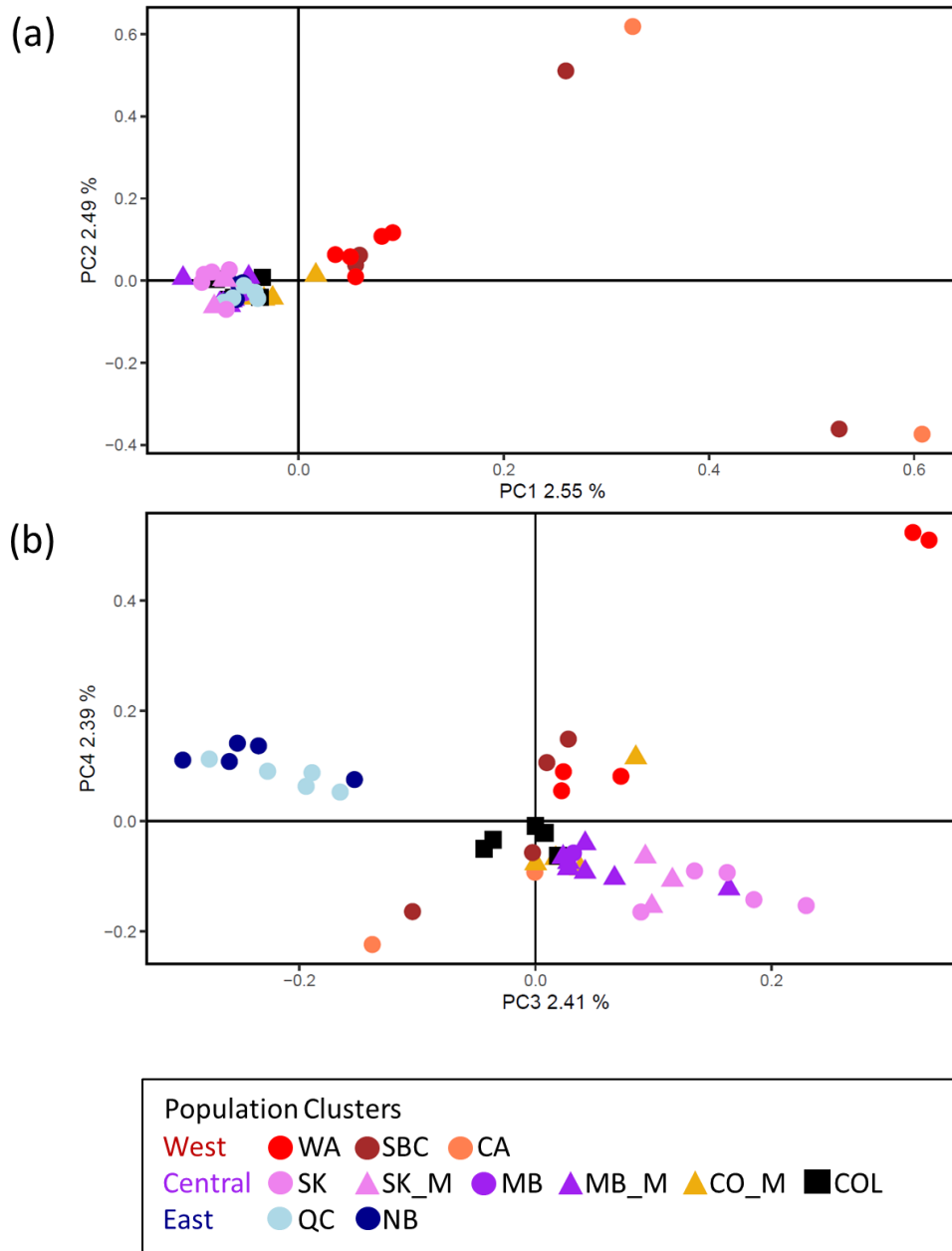


Figure 2.5: Principal component analysis for barn swallows including breeding, migratory and wintering populations using lcWGS indicating (a) PC 1 vs PC 2 (b) PC 3 vs PC 4. PC1, PC2, PC3 and PC4 explain 2.55%, 2.49%, 2.41% and 2.39% variance respectively. Triangles = migratory, circles = breeding, and squares = wintering populations. Shades of red indicate the western populations, shades of purple indicate the central populations, and shades of blue indicate the eastern populations.

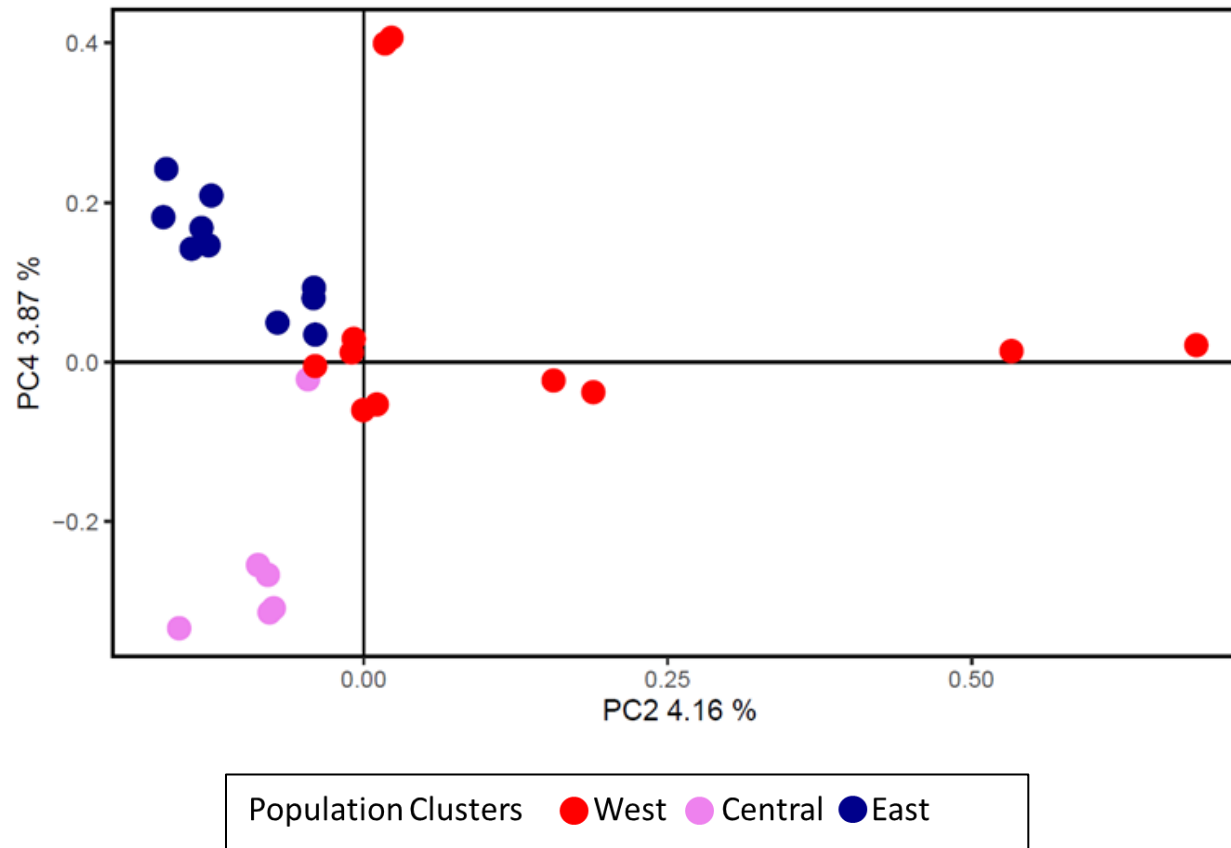


Figure 2.6: Principal component analysis for barn swallows using only breeding populations data from lcWGS.

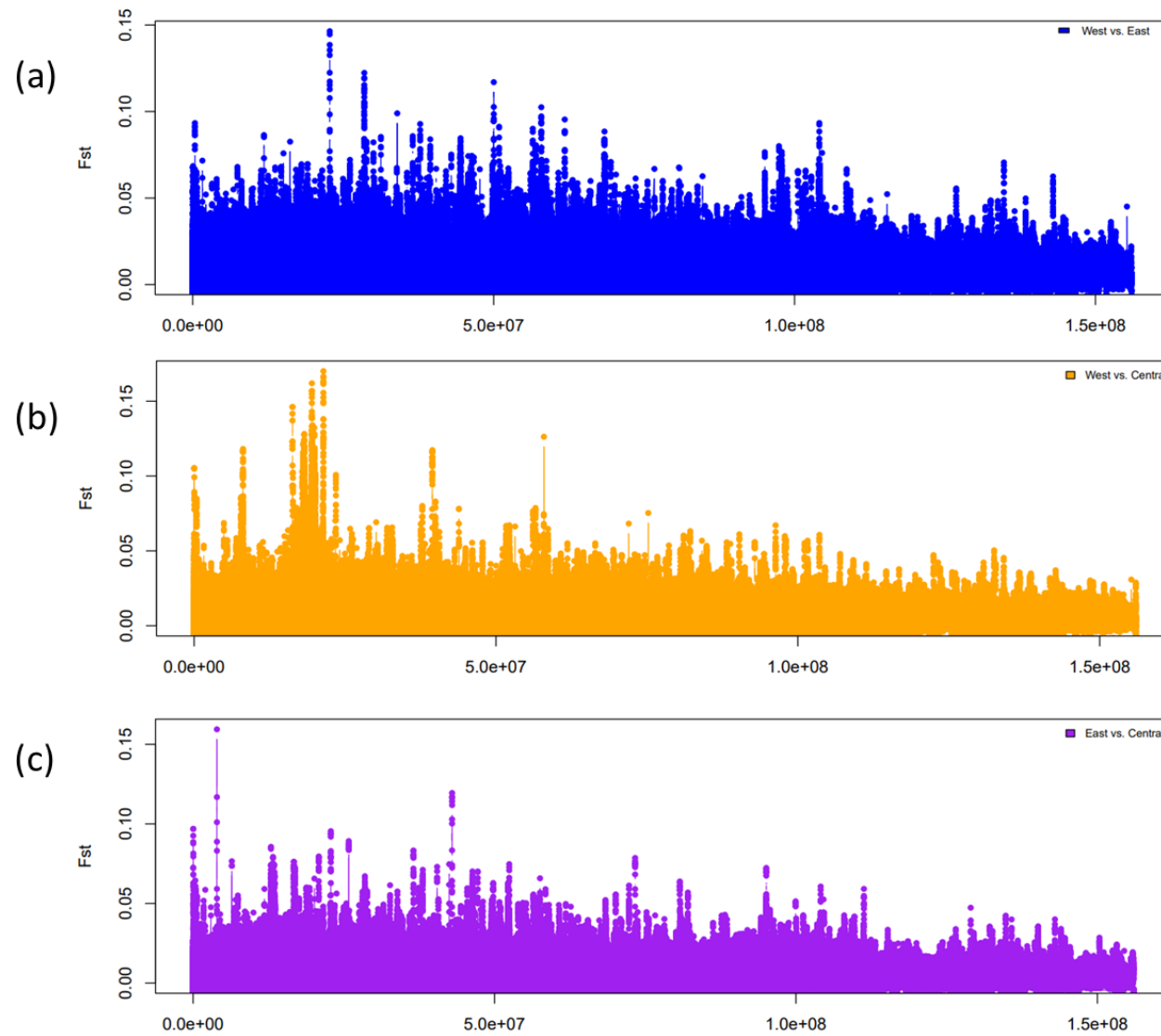


Figure 2.7: Pairwise F_{ST} comparison across the genome between three different clusters using lcWGS (a) west vs east (b) west vs central (c) east vs central.

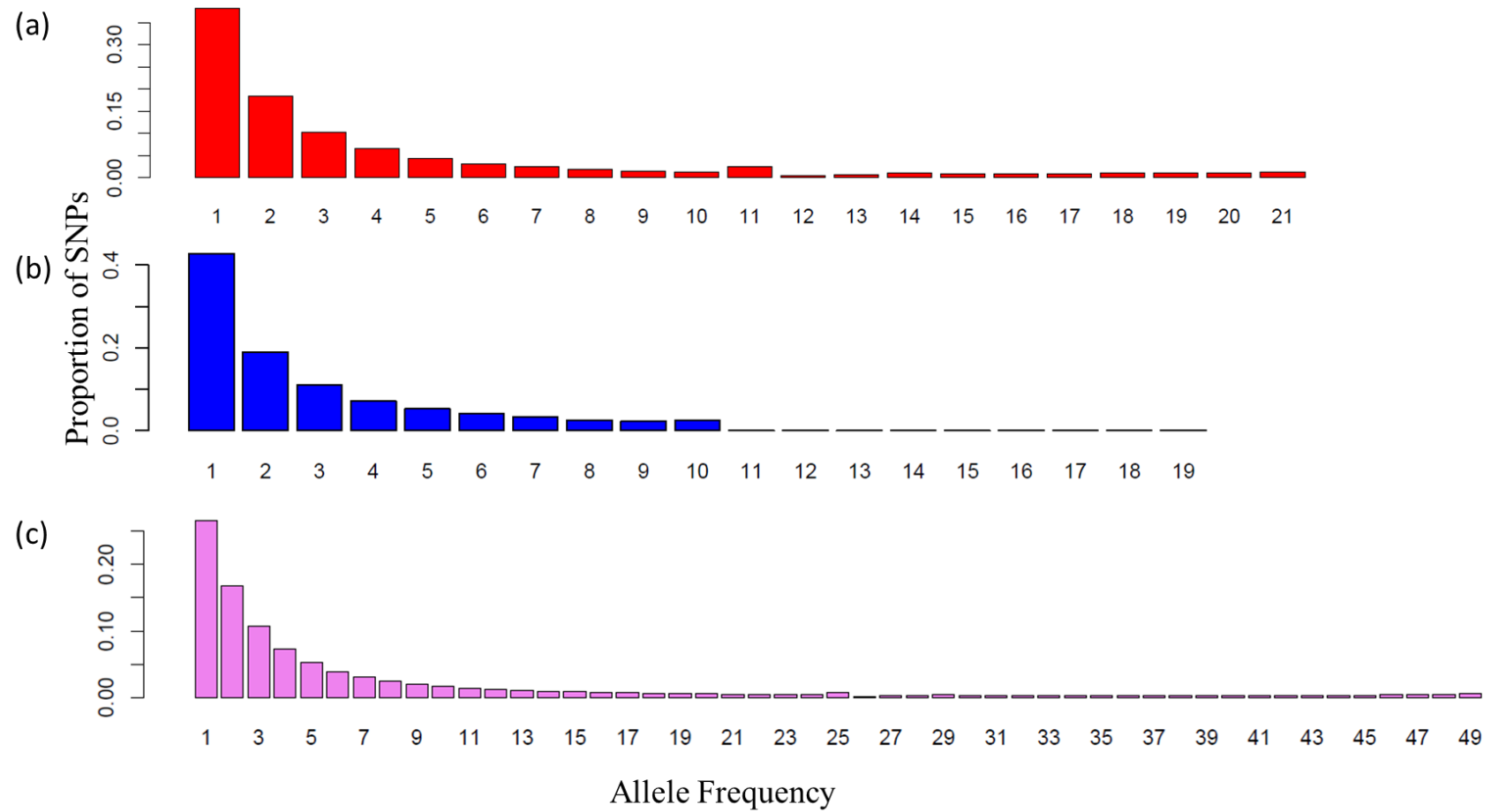


Figure 2.8: Site Frequency Spectrum (SFS) for each cluster using lcWGS data (a) SFS for the western, (b) eastern, and (c) central clusters.

Chapter 3: Population Genetic Structure of North American Bank Swallows

Abstract

Bank swallows (*Riparia riparia*) have emerged as one of the most affected aerial insectivores in North America with a decline of 10% annually and numbers have plummeted since 1970. The available studies on bank swallows have focused primarily on specific regions, providing only localized insights into their population structure and genetic diversity. To fill this knowledge gap and gain insights into effective conservation practices, my study employed six sampling sites spanning the bank swallow's breeding range and integrated the advanced whole genome sequencing, to understand the population genetic structure across the continent. My findings revealed the presence of three distinct genetic clusters in bank swallows across North America, aligning with the western, central, and eastern regions. These genetic differentiations could be attributed to historical isolation resulting from the last ice age, as well as restricted gene flow influenced by geographical barriers and migratory pathways.

3.1 Introduction

Conservation efforts for migratory species, especially those undertaking lengthy journeys and spending their annual cycle in different geographic locations, pose unique challenges. Implementing conservation plans for long distant migrant species requires a comprehensive understanding of the multiple threats they encounter throughout their entire annual cycle, spanning breeding, migration, and wintering periods (Faaborg et al., 2010; Lovette et al., 2004).

Using population genomic approaches, it is possible to identify different evolutionary units within migratory bird populations. This allows for more accurate conservation strategies specifically tailored to the unique genetic traits and ecological requirements of these populations (Funk et al., 2012). It is essential to consider the evolutionary significant units and their genetic diversity to address these challenges and effectively prioritize conservation actions. Population genetics uses genetic data to understand how populations are connected across their range. Within species, some populations can be genetically distinct from one another due to barriers to gene flow (Ravinet et al., 2017). These barriers encompass a wide range of factors, including historical events, geographical constraints, behavioural variation, and climatic conditions (Adams & Burg, 2015). Historical events such as the last glacial maximum (LGM) are one of the well-documented barriers that contribute to the population genetic structure in many North American species. Widespread glaciations during the Pleistocene (~ 2.5 million years ago) caused species to reside solely in ice-free refugia (Hewitt, 2004). Major climatic oscillations between cold glacial and warmer interglacial, habitat shifts, fragmentation and extinction events were observed in many species during the glaciation events. For instance, the three-toed woodpecker

(*Picoides tridactylus*) and the winter wren (*Troglodytes troglodytes*) demonstrate the genetic differences between North America and Eurasia corresponding to the LGM, indicating the impact of the LGM on their genetic divergence (Zink et al., 2002). Whereas other species like barn swallows (*Hirundo rustica*) provide examples for secondary colonization of Eurasia from North America via the Bering land bridge during the LGM (Drovetski et al., 2004; Zink et al., 2006;). These stochastic and selective effects on genetic variation and architecture due to the LGM persist to this day (Hewitt, 2004).

While historic events such as the last glacial maximum play a major role in the population genetic structure, contemporary landscape factors continue to restrict gene flow between populations. Mountain ranges, large bodies of water, microclimatic variations, anthropogenic modifications, and habitat differences can all influence the genetic architecture of present-day populations. The Cascade/Coast Mountain range in the west and the Rocky Mountains in the interior, for instance, have been shown to contribute to genetic differentiation among populations in a diverse range of species (Burg et al., 2006; Green et al., 1996; Hindley et al., 2018; Lait & Burg, 2013). These contemporary barriers, along with glacial refugia, significantly altered species-specific dispersal and migratory patterns. Different migratory pathways within species also can be influenced by geographical barriers as a result, continental-wide migratory divides have been observed in many species corresponding to barriers. For instance, tree swallows (*Tachycineta bicolor*) in North America show three distinct migratory pathways that represent west, central and eastern regions, with the west and central pathways are mostly separated by the Rocky Mountains (Knight et al., 2018). Furthermore, migratory divides can act as barriers to gene flow (Delmore et al., 2012; Delmore & Irwin, 2014). Therefore, historical events combined with contemporary

barriers and migratory routes influence the population genetic structure of migratory species.

The decline of Neotropical Nearctic migrant birds, known for their remarkable long-distance journeys, has become an alarming concern (Rosenberg et al., 2019). Among these migratory species, aerial insectivores have experienced the most significant population decline, averaging 59% over the last 60 years. Bank swallows (*Riparia riparia*) have emerged as one of the most affected aerial insectivores in North America with a decline of 10% annually and numbers have plummeted since 1970. Canadian bank swallow populations have declined 98% over the last 40 years and are categorized as threatened by COSEWIC (Committee on the Status of Endangered Wildlife in Canada) (Nebel et al., 2020; Nicole, 2008; Sauer et al., 2017).

Bank swallows (*Riparia riparia*) are of particular interest as a species of concern due to their rapid decline in North America and documented low genetic differentiation using one mitochondrial marker and one nuclear marker (Pavlova et al., 2008). Bank swallows, belonging to the swallow family (Hirundinidae), are small migratory aerial insectivores that nest in burrows along vertical banks and bluffs. Based on the annual breeding bird survey population index from 1970 to 2010, bank swallows have experienced a steep decline, particularly in northeastern North America (Imlay et al., 2020; Sauer et al., 2017). In fact, northeastern North America or Atlantic Canada reported a staggering 98% decline in bank swallow populations between 1970 and 2010, leading to their designation as a threatened species under Canada's Species at Risk Act and an endangered in Nova Scotia Endangered Species Act (COSEWIC, 2013; Sauer et al., 2017).

Bank swallows have a wide distribution across Europe, Asia, and North America. Studies suggest that the pale sand martin (*Riparia diluta*) and collared sand martin or bank swallow (*Riparia riparia*) as it is known in North America, diverged sometime between late Pliocene and middle Pleistocene (Schweizer et al., 2018; Pavlova, 2008). The absence of well supported clades and haplotype networks using mitochondrial markers suggest relatively low genetic diversity and recent divergence in bank swallows (Pavlova et al., 2008). *R. diluta* is distributed in the Middle East regions and three to four subspecies of *Riparia riparia* have been recognized: *Riparia riparia riparia* breeds in North America, Eurasia, northwest Africa and the Mediterranean region. North American populations are also referred to as *R. r. maximiliani* and are slightly smaller in body size compared to their Eurasian counterparts (Arny, 1952). The second and third subspecies *R. r. ijimae* and *R. r. shelleyi* breed in eastern parts of Asia and Egypt respectively (Del Hoyo J & Elliott A, 2014; Garrison & Turner, 2020).

North American bank swallows breed across the northern regions of North America, ranging from Alaska to the northeast of the Northwest Territories and into the central United States (Figure 3.1) (Garrison & Turner, 2020). The studies of migratory patterns of the North American bank swallows are limited to populations in Atlantic Canada. These breeding populations winter between northern Argentina and southern Brazil (Imlay et al., 2020).

In contrast to many other swallow species, only a few studies have examined the genetic structure of bank swallows breeding in North America, and these studies have predominantly utilized mitochondrial markers (Pavlova et al., 2008). Since mitochondrial markers indicate a shallow evolutionary history and low genetic diversity within this group, I employed low-coverage whole-genome sequencing to

investigate the population genetic structure of bank swallows breeding across North America. I aim to answer the following questions: 1) Are there genetically distinct populations of bank swallows in North America? If so, 2) What factors act as barriers to gene flow? My study will identify genetically distinct populations and ultimately define evolutionary significant units for conservation prioritization for rapidly declining bank swallow populations.

3.2 Methods

3.2.1 Low Coverage Whole Genome Sequencing (lcWGS) Analyses

3.2.1.1 Sample Collection and DNA Extraction

I obtained feather samples and/or blood samples from bird banding stations and tissue samples from museums during the breeding season. Birds were caught using mist nets and up to 50 µl of blood was collected using brachial vein puncture or one tail feather was obtained (Owen, 2011). Birds were released at the sites where they were caught. Samples were stored in ethanol at ambient temperature until they could be returned to the lab where they were stored at -20°C. Frozen tissue samples or skin clips were collected from the museums (Royal Saskatchewan Museum, Beaty Biodiversity Museum and Burke Museum).

I performed modified salting out DNA extractions for 21 samples (New Brunswick (n=1), Ontario (n=6), Alberta (n=3), Saskatchewan (n=4), British Columbia (n =1) and Washington (n=6) (Table 3.1 and Figure 3.1) and samples were quantified using a Nanodrop spectrophotometer (Aljanabi & Martinez, 1997). Samples were sent to Genome Quebec and Novogene California for paired-end Illumina NovaSeq 6000 S4 PE 150 low-coverage whole genome sequencing.

3.2.1.2 Processing of Low-coverage Whole Genome Sequencing (lcWGS) Data

I removed sequence adaptors and quality trimmed reads using the q30,30 option in cutadapt version 1.16 to remove the reads with a Phred score less than 30 (Martin, 2011). The sequences were then aligned to the scaffold-level assembled bank swallow reference genome (GCA_020917445.1) (Tang et al., 2022). The bwa alignment tool was used to align sequences to the reference genome and sorted using Samtools v1.10 (Li et al., 2009). The PCR duplicates were identified and marked using Picard (Picard, 2018). Overlapping regions of the sequences were clipped and the sequences around insertions and deletions (INDELs) were realigned. Before using sequences for downstream analyses, I validated and checked the depth of the final binary alignment files to ensure the reliability of the sequences.

3.2.1.3 Genotype Likelihood Estimation

Genotype likelihood data were obtained using ANGSD v0.933 (Analysis of Next Generation Sequencing Data), a program specially designed to analyze lcWGS data by considering low to moderate coverage associated with lcWGS (Korneliussen et al., 2014). I used SAMtools genotype likelihood model to incorporate the uncertainty of true genotypes, as low sequencing coverage approach restricts genotype calling (Li, 2011; Li et al., 2009). Using ANGSD, I removed low-quality reads (-uniqueOnly 1 -remove_bads 1, -minMapQ 30, -minQ 30), and unmapped reads (-only_proper_pairs) obtained a P value $\leq 10^{-6}$ (-SNP_pval 1e-6). Furthermore, I used (-minMaf 0.05) to retain the sites with minor allele frequency above 0.05.

3.2.1.4 Population Genetic Structure

For the evaluation of population genetic structure in bank swallows, I used the filtered lcWGS dataset to perform multivariate analyses. First, I conducted principal

component analysis (PCA) using PCAngsd which uses genotype likelihoods to build the covariant matrix (Meisner & Albrechtsen, 2018). The resulting covariant matrix was visualized using the RStudio ggplot package. To determine the ancestry proportions of each individual, I performed admixture analysis using NGSadmix for $K = 2 - 4$ (Skotte et al., 2013). The tolerance for considering a site as missing was set to 0.05, the convergence tolerance was set to 0.0001 and a minimum number of informative individuals was set to 10 (~50% of total number of individuals) (Korneliussen et al., 2014). The optimal Evanno K value was identified and admixture plots were visualized using CLUMPAK (Kopelman et al., 2015).

3.2.1.5 Genetic Differentiation Between Populations

To examine the population-level genetic differences, allele frequencies were calculated using ANGSD based on genotype likelihood data across all populations. Based on the clusters obtained by PCA and admixture plots, individuals were grouped into three different populations (west (BC, WA), central (SEAB, SK), and east (ON, NB)) and allele frequencies at all SNPs were calculated using the `-doSaf` option (Nielsen et al., 2012). I performed folded site frequency spectrum (SFS) using the assembled bank swallow reference genome as the outgroup to polarize the allele frequencies and estimate the derived allele frequency. After filtering to remove low quality sites, extreme outliers in sequencing depth, and spurious alignments; the `realSFS` command was used to generate the folded site frequency spectrum for each population.

To examine the differentiation between west, east and central population clusters and to detect loci under selection, I calculated the fixation index (F_{ST}) using ANGSD. Windowed estimation of F_{ST} was performed using 10 kb windows across the genome

with a 1 kb step and visualized through Manhattan plots using R package. However, due to the absence of fully annotated reference genome for bank swallows, I was unable to detect which genes are under selection.

3.2.1.6 Genetic Diversity and Neutrality Statistics Within Populations

Thetas (θ) were calculated using allele frequencies for each population of interest (west, central and east) to estimate nucleotide diversity (π) and Tajima's D statistics using ANGSD. Both π and Tajima's D were calculated for each site in non-overlapping 10 kb windows across the genome, with a 1 kb step. This involved estimating the SFS in the realSFS tool. Subsequently, I plotted the values of π and Tajima's D against the regions with high F_{ST} values to identify potential signs of selection. Regions exhibiting high F_{ST} values and low π suggest the possibility of selection acting upon those specific gene regions. A positive Tajima's D indicates balancing selection or a population contraction, while a negative Tajima's D suggests population expansion or positive selection.

3.2.2 Identifying Bank Swallow Migratory Pathways

I utilized publicly available Motus migratory data from 2018 - 2023 to gain insights of the migratory pathways of bank swallows across North America. No refined filters were applied to individual tracks, and publicly accessible animations from tags of the "Bank Swallow Migratory Connectivity", "Ontario Bank Swallows", "Swallows" and "Swallow Roosting Ecology" projects were employed to identify initial migratory flyways. I utilized maps from Motus receiving stations located throughout North and Central America to estimate the main bank swallow flyways. I only included major flyways as I did not perform any filtering on the publicly available dataset.

3.3 Results

I obtained an average per individual genome-wide coverage of $\sim 6.56 X$ (ranging from $2.8X - 8.8X$) for 21 individual bank swallows. A total of 19,799,442 sites and an average coverage of $\sim 3.8X$ per individual were retained after the filtering.

3.3.1 Population Genetic Structure of Bank Swallows using lcWGS Data

Principal component analysis indicates three separate clusters (Figure 3.2) which correspond to the western, central, and eastern regions of North America. PC1 and PC2 resulted in the highest variance, 9.39% and 4.92% respectively. The western cluster comprises samples from Washington and British Columbia. Samples from Saskatchewan and southeast Alberta comprise the central cluster, and samples from New Brunswick and Ontario form the eastern cluster (Figure 3.3). None of the clusters overlap in the PCA, and high separation was observed between WA and ON populations. PC3 explained 4.84% of variance, however it only separated 2 SK individuals from the other samples. It is interesting to notice the samples from southeast Alberta (Cypress Hills) cluster together very closely indicating low genetic diversity between individuals.

3.3.2 Comparative Population Analysis

Windowed F_{ST} values between each population comparison revealed a few regions with comparatively higher F_{ST} values ($0.20 - 0.25$) (Figures 3.4, 3.5 and 3.6).

Furthermore, I observed a low π ($\pi < 0.15$) associated with those high F_{ST} regions indicating signs of selection. The high F_{ST} and low π were observed in the scaffold JAIXNV010000025.1 for both west and central, and west and east comparison (Figure 3.4 and Figure 3.5). I observed F_{ST} peaks for the central and west comparison in scaffold JAIXNV010000006.1 (Figure 3.4). For the east and central comparison,

scaffold JAIXNV010000031.1 also contains a peak with high F_{ST} value and low π (Figure 3.6).

Tajima's D values had a broad range; however, I observed a negative mean value for all three populations within the same windows (Figures 3.4, 3.5 and 3.6) indicating possible population expansion or selective sweep.

3.3.3 Migratory Pathways

Using maps from the Motus wildlife tracking system, I identified interconnected migratory networks for bank swallows, however, three main pathways can be clearly identified. Birds from the Yukon ($n = 8$) use the pathways east of the Rocky Mountains through Alberta and Saskatchewan towards Central America, whereas birds from the west of the Rocky Mountains ($n = 6$) remain west of the Rocky Mountains, separating the migratory routes of populations west and east of the Rocky Mountains. Populations from the northeastern region ($n = 7$) primarily migrate towards Central America along the eastern seaboard of North America (Figure 3.7). In addition to the major flyways, I also observed individuals from two individuals from Ontario migrate to the west of the Rocky Mountains. However, I only included major flyways.

3.4 Discussion

Using low-coverage whole genome sequencing data, I was able to detect three distinct genetic clusters for bank swallows corresponding to western, central, and eastern regions of North America (Figure 3.3). Historical barriers, along with contemporary barriers and migratory pathways are likely playing important roles in creating these genetically distinct populations.

My analyses revealed significant genetic differences between North American populations. In particular, WA and ON were quite distinct in the PCA analysis (Figure 3.2). Additionally, the central populations from SEAB and SK formed a unique cluster in the PCA, distinct from both the eastern and western clusters. Due to limited sample sizes for NB and SBC populations, drawing definitive conclusions for those populations is limited. However, taking into account their geographical proximity and landscape structure, I assigned the NB individual to the eastern cluster and the SBC individual to the western cluster. Genomic scans for each population comparison indicate possible selection in a few gene regions (Figures 3.4, 3.5 and 3.6). It is interesting to observe the selection of the same region in both the west and central comparison and the west and east comparison. However, some of the corresponding to the regions with high F_{ST} in both comparisons match while other regions in the same contig differed. The matching regions might be because of shared selective pressure between the populations. On the other hand, the differing regions within the same contig might be specific genetic differences that are unique to each population comparison. However, the absence of an annotated reference genome for bank swallows limits our ability to detect specific gene regions of selection which are driving these patterns.

3.4.1 Insights into the Role of Glacial Refugia in Shaping Species Distribution and Colonization of Bank Swallows

The most likely explanation for the patterns I observed for bank swallows is that they resided in multiple glacial refugia during the Pleistocene and remained separate following recolonization. Unlike many bird species that are closely associated with boreal forests, bank swallows have a different ecological niche, primarily linked to water bodies and feeding on flying insects. Considering the diets of bank swallows, it

is likely that the glacial refugia for insect species played a significant role in providing suitable environments to support bank swallow refugia during the LGM. Bank swallows exclusively feed on insects like grasshoppers, beetles, flies, bees, and wasps. The grasshoppers (*Melanoplus* spp.) are believed to have resided in multiple southern refugia in Montana and Idaho (Knowles, 2001), and mitochondrial and microsatellite markers suggest the existence of Beringian and south and Pacific Northwest refugia for beetle species (*Dendroctonus rufipennis*) (Maroja et al., 2007). In eastern North America, tiger beetles (*Cicindella dorsalis*) have been reported to have Atlantic coastal refugia and saw-combed fishfly (*Nigronia serricornis*) indicates the presence of six geographically structured clades in the eastern region, suggesting northward movement from southeast refugium (Heilveil & Berlocher, 2006; Vogler et al., 1993). Some studies on a wide range of species including insects and birds also provide evidence for the Appalachian Mountains and southern USA as potential refugia (Jones et al., 2006; Liu et al., 2006; Soltis et al., 2006).

Similar congruence is found in bird studies, for example Beringia and southern refugia have been reported in many bird studies (Burg et al., 2006; Shafer et al., 2010). Furthermore Colbeck et al., (2008) and Lait & Burg (2013) reported Atlantic coastal refugia for the widespread North American redstarts (*Setophaga ruticilla*), and boreal chickadees (*Poecile hudsonicus*) respectively.

Based on the factors discussed and my findings, I propose that bank swallows recolonized their current range from two refugia in the western region, possibly Beringia and the Pacific Northwest or southern refugia as well as one in eastern North America. The existence of two refugia in the west explains the divergence between the western and central populations, which are separated by the Rocky Mountains. If both the western and central populations had originated from a single refugium, I

would not observe a clear separation between the two clusters. Obtaining high-resolution lcWGS data from the Yukon region, increasing the sample size and sampling coverage in future studies will contribute to more robust conclusions.

3.4.2 Role of Contemporary Barriers and Migratory Pathways

Glaciation events also played a significant role in shaping their migratory pathways. The changes resulting from glaciation, such as the formation of ice sheets, the melting of glaciers, and shifts in the landscape, have led to alterations in the availability of suitable habitats and the emergence of geographic barriers. As a consequence, birds and other species were forced to adapt and find alternative pathways to reach their breeding and wintering grounds. The Rocky Mountains serve as a significant barrier to gene flow between bank swallow populations, contributing to their genetic differentiation. The Rocky Mountain range is recognized as a suture zone for numerous North American species (Brelsford & Irwin, 2009; Flockhart & Wiebe, 2009). Unlike other species of swallows in North America, bank swallows are not known to breed west of the Coast/Cascade Mountains (Figure 3.1), possibly due to the challenging breeding conditions associated with a high elevation and precipitation (Garrison, 1999). Migratory patterns further support this, as birds from the Yukon region tend to migrate eastward, avoiding flying over the Rocky Mountains and instead move towards central North America. Similar patterns are observed in the migratory pathways of tree swallows (*Tachycineta bicolor*), where birds from the Yukon and central North America follow the Mississippi River while the western population remains west of the Rocky Mountains (Knight et al., 2018). Although the exact breeding grounds for different western populations are yet to be discovered, it is possible that distinct migratory routes and breeding times contribute to maintaining their separation, even if they mix on the wintering grounds.

When considering the separation of the eastern population from the other populations, it is crucial to take into account not only the eastern glacial refugium but also the migratory routes along contemporary barriers such as the Appalachian Mountains and the Mississippi River. Migratory data show that the majority of the bank swallows from the eastern population migrate along the East Coast, indicating distinct pathways (Imlay et al., 2020). Other than the clear separation of the eastern population, I also observed one sample from NB separate from the ON populations though my conclusions are limited due to having data from only one individual from New Brunswick. The presence of different migratory pathways and geographical barriers like the Mississippi River and the Appalachian Mountains in the east could be potential factors preventing gene flow between the NB and ON populations. A variety of amphibians, reptiles and mammal species exhibit a clear division between the eastern and western sides of the Appalachian Mountains (Avisé et al., 1979; Burbrink, 2002; Church et al., 2003; Soltis et al., 2006; Walker et al., 1997). However, the Appalachian Mountain discontinuity is not prominent in birds.

It is important to note that although some populations have relatively small sample sizes, my conclusions are primarily based on the three main population clusters, each comprised of seven samples. The sample size of seven individuals is unlikely to pose a problem for the analyses I conducted. This is because I employed a large number of SNPs distributed across the genome and average depth of coverage of 4.8X. Studies have demonstrated that even with sample sizes as low as four individuals per population, accurate analyses can still be achieved when a high number of SNPs are utilized (Willing et al., 2012). Furthermore, some studies show that five individuals per population with a coverage above 1X can obtain reliable results even in high migration scenario (Lou et al., 2021). To further investigate the population genetic

structure and migration routes of North American bank swallows, it is necessary to include samples from the isolated breeding population in Mexico, the central part of the United States, Yukon, and obtain additional samples from New Brunswick, in addition to migratory samples.

3.5 Conclusions

This is the first study of declining North American bank swallows that utilized whole genome sequencing data to understand the population genetic structure across the continent. My analyses revealed the presence of three distinct genetic clusters, representing the western, central, and eastern regions of North America. I propose that historical and contemporary barriers, in addition to migratory pathways, play a crucial role in limiting gene flow between these populations. Furthermore, an important outcome of this study is the identification of genetically distinct declining populations of North American bank swallows, highlighting the need to consider them as separate management units for conservation purposes.

References

- Adams, R. V., & Burg, T. M. (2015). Influence of ecological and geological features on range wide patterns of genetic structure in a widespread passerine. *Heredity*, *114*(2), 143-154.
- Aguillon, S. M., Walsh, J., & Lovette, I. (2020). Extensive hybridization reveals multiple coloration genes underlying a complex plumage phenotype. *Proceedings of the Royal Society B*, *288*(1943), 20201805
- Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*(22), 4692-4693.
- Arny, S. A. (1952). Taxonomic status of the bank swallow of North America. *The Condor*, *54*(6), 356-357.
- Awise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C., & Lansman, R. A. (1979). Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences*, *76*(12), 6694-6698.
- Brunner, P. C., Douglas, M. R., Osinov, A., Wilson, C. C., & Bernatchez, L. (2001). Holarctic phylogeography of Arctic charr (*Salvelinus alpinus l.*) inferred from mitochondrial DNA sequences. *Evolution*, *55*(3), 573-586.
- Burbrink, F. T. (2002). Phylogeographic analysis of the cornsnake (*Elaphe guttata*) complex as inferred from maximum likelihood and Bayesian analyses. *Molecular Phylogenetics and Evolution*, *25*(3), 465-476.
- Burg, T. M., Gaston, A. J., Winker, K., & Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, *15*(9), 2409-2419.
- Church, S. A., Kraus, J. M., Mitchell, J. C., Church, D. R., & Taylor, D. R. (2003). Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. *Evolution*, *57*(2), 372-383.
- Colbeck, G. J., Gibbs, H. L., Marra, P. P., Hobson, K., & Webster, M. S. (2008). Phylogeography of a widespread North American migratory songbird (*Setophaga ruticilla*). *Journal of Heredity*, *99*(5), 453-463.
- Del Hoyo J, & A, E. (2014). Barn swallow (*Hirundo rustica*). In *Handbook of the Birds of the World*. Lynx Edicions.
- Delmore, K. E., Fox, J. W., & Irwin, D. E. (2012). Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1747), 4582-4589.
- Delmore, K. E., & Irwin, D. E. (2014). Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, *17*(10), 1211-1218.

- Drovetski, S. V., Zink, R. M., Rohwer, S., Fadeev, I. V., Nesterov, E. V., Karagodin, I., Koblik, E. A., & Red'kin, Y. A. (2004). Complex biogeographic history of a Holarctic passerine. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1538), 545-551.
- Eklblom, R., Brechlin, B., Persson, J., Smeds, L., Johansson, M., Magnusson, J., Flagstad, Ø., & Ellegren, H. (2018). Genome sequencing and conservation genomics in the Scandinavian wolverine population. *Conservation Biology*, 32(6), 1301-1312.
- Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux Jr, S. A., Heglund, P., Hobson, K. A., Jahn, A. E., & Johnson, D. H. (2010). Conserving migratory land birds in the new world: Do we know enough? *Ecological Applications*, 20(2), 398-418.
- Forcina, G., Tang, Q., Cros, E., Guerrini, M., Rheindt, F. E., & Barbanera, F. (2021). Genome-wide markers redeem the lost identity of a heavily managed gamebird. *Proceedings of the Royal Society B*, 288(1947), 20210285.
- Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, 27(9), 489-496.
- Garrison, B. A. (1999). Bank swallow (*Riparia riparia*). In *The Birds of North America* (pp. 28).
- Garrison, B. A., & Turner, A. (2020). Bank swallow (*Riparia riparia*). In S. M. Billerman (Ed.), *Birds of the World* (Version 1.0 ed.). Cornell Lab of Ornithology.
- Green, D. M., Sharbel, T. F., Kearsley, J., & Kaiser, H. (1996). Postglacial range fluctuation, genetic subdivision and speciation in the western North American spotted frog complex, *Rana pretiosa*. *Evolution*, 50(1), 374-390.
- Heilveil, J., & Berlocher, S. (2006). Phylogeography of postglacial range expansion in *Nigronia serricornis say* (Megaloptera: Corydalidae). *Molecular Ecology*, 15(6), 1627-1641.
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183-195.
- Hickerson, M. J., & Ross, J. R. (2001). Post-glacial population history and genetic structure of the northern clingfish (*Gobbiopsis maeandricus*), revealed from mtDNA analysis. *Marine Biology*, 138, 407-419.
- Hindley, J. A., Graham, B. A., & Burg, T. M. (2018). The influence of latitude, geographic distance, and habitat discontinuities on genetic variation in a high latitude montane species. *Scientific Reports*, 8(1), 1-11.
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62-82.
- Holliday, J. A., Hallerman, E. M., & Haak, D. C. (2019) Genotyping and sequencing technologies in population genetics and genomics. In O. P. Rajora (Ed.),

Population genomics: *Concepts, Approaches and Applications* (pp. 83– 125). Springer Nature Switzerland AG.

- Imlay, T., Saldanha, S., & Taylor, P. (2020). The fall migratory movements of bank swallows, *Riparia riparia*: Fly-and-forage migration? *Avian Conservation and Ecology*, *15*(1), 2.
- Jones, M. T., Voss, S. R., Ptacek, M. B., Weisrock, D. W., & Tonkyn, D. W. (2006). River drainages and phylogeography: An evolutionary significant lineage of shovel-nosed salamander (*Desmognathus marmoratus*) in the southern Appalachians. *Molecular Phylogenetics and Evolution*, *38*(1), 280-287.
- Knight, S. M., Bradley, D. W., Clark, R. G., Gow, E. A., Bélisle, M., Berzins, L. L., Blake, T., Bridge, E. S., Burke, L., & Dawson, R. D. (2018). Constructing and evaluating a continent-wide migratory songbird network across the annual cycle. *Ecological Monographs*, *88*(3), 445-460.
- Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, *10*(3), 691-701.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across k. *Molecular Ecology Resources*, *15*(5), 1179-1191.
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). Angsd: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 1-13.
- Lait, L. A., & Burg, T. M. (2013). When east meets west: Population structure of a high-latitude resident species, the boreal chickadee (*Poecile hudsonicus*). *Heredity*, *111*(4), 321-329.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987-2993.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Subgroup, G. P. D. P. (2009). The sequence alignment/map format and Samtools. *Bioinformatics*, *25*(16), 2078-2079.
- Liu, F. G. R., Moler, P. E., & Miyamoto, M. M. (2006). Phylogeography of the salamander genus *Pseudobranchius* in the southeastern United States. *Molecular Phylogenetics and Evolution*, *39*(1), 149-159.
- Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, *30*(23), 5966-5993.
- Lovette, I. J., Clegg, S. M., & Smith, T. B. (2004). Limited utility of mtDNA markers for determining connectivity among breeding and overwintering locations in three neotropical migrant birds. *Conservation Biology*, *18*(1), 156-166.
- Maroja, L. S., Bogdanowicz, S. M., Wallin, K. F., Raffa, K. F., & Harrison, R. G. (2007). Phylogeography of spruce beetles (*Dendroctonus rufipennis*

- kirby*)(Curculionidae: Scolytinae) in North America. *Molecular Ecology*, 16(12), 2560-2573.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, 17(1), 10-12.
- Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, 210(2), 719-731.
- Nebel, S., Casey, J., Cyr, M. A., Kardynal, K., Krebs, E., Purves, E., Bélisle, M., Brigham, R., Knight, E., & Morrissey, C. (2020). Falling through the policy cracks: Implementing a roadmap to conserve aerial insectivores in North America. *Avian Conservation and Ecology*, 15(1), 23.
- Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS One* 7, e37558 (2012).
- Owen, J. C. (2011). Collecting, processing, and storing avian blood: A review. *Journal of Field Ornithology*, 82(4), 339-354.
- Ravinet, M., Faria, R., Butlin, R., Galindo, J., Bierne, N., Rafajlović, M., Noor, M., Mehlig, B., & Westram, A. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450-1477.
- Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., & Parr, M. (2019). Decline of the North American avifauna. *Science*, 366(6461), 120-124.
- Sauer, J. R., Pardieck, K. L., Ziolkowski Jr, D. J., Smith, A. C., Hudson, M.-A. R., Rodriguez, V., Berlanga, H., Niven, D. K., & Link, W. A. (2017). The first 50 years of the North American breeding bird survey. *The Condor: Ornithological Applications*, 119(3), 576-593.
- Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, 15(11), 749-763.
- Schneider, N.A. (2008). Effects of climate change on avian life history and fitness. *Introductory Research Essay (Department of Ecology, SLU)*, 4, 1.
- Shafer, A. B., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, 19(21), 4589-4621.
- Sheldon, F. H., & Winkler, D. W. (1993). Intergeneric phylogenetic relationships of swallows estimated by DNA-DNA hybridization. *The Auk*, 110(4), 798-824.
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, 195(3), 693-702.
- Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, 15(14), 4261-4293.

- Steane, D. A., Potts, B. M., McLean, E., Collins, L., Prober, S. M., Stock, W. D., Vaillancourt, R. E., & Byrne, M. (2015). Genome-wide scans reveal cryptic population structure in a dry-adapted eucalypt. *Tree Genetics & Genomes*, *11*, 1-14.
- Tang, Q., Burri, R., Liu, Y., Suh, A., Sundev, G., Heckel, G., & Schweizer, M. (2022). Seasonal migration patterns and the maintenance of evolutionary diversity in a cryptic bird radiation. *Molecular Ecology*, *31*(2), 632-645.
- Vogler, A. P., Desalle, R., Assmann, T., Knisley, C. B., & Schultz, T. D. (1993). Molecular population genetics of the endangered tiger beetle *Cicindela dorsalis* (Coleoptera: Cicindelidae). *Annals of the Entomological Society of America*, *86*(2), 142-152.
- Walker, D., Nelson, W. S., Buhlmann, K. A., & Avise, J. C. (1997). Mitochondrial DNA phylogeography and subspecies issues in the monotypic freshwater turtle *Sternotherus odoratus*. *Copeia*, *1*, 16-21.
- Willing, E. M., Dreyer, C., & 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.
- Zink, R. M., Pavlova, A., Rohwer, S., & Drovetski, S. V. (2006). Barn swallows before barns: Population histories and intercontinental colonization. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1591), 1245-1251.
- Zink, R. M., Rohwer, S., Drovetski, S., Blackwell-Rago, R. C., & Farrell, S. L. (2002). Holarctic phylogeography and species limits of three-toed woodpeckers. *The Condor*, *104*(1), 167-170.

Table 3.1: Sampling locations and the number of samples utilized in the analyses.

Population ID	Number of Samples Utilized for lcWGS	Genetic Cluster
Washington (WA)	6	West
Southern British Columbia (SBC)	1	West
Cypress Hills, Southeast Alberta (SEAB)	3	Central
Saskatchewan (SK)	4	Central
Ontario (ON)	6	East
New Brunswick (NB)	1	East
Total	21	

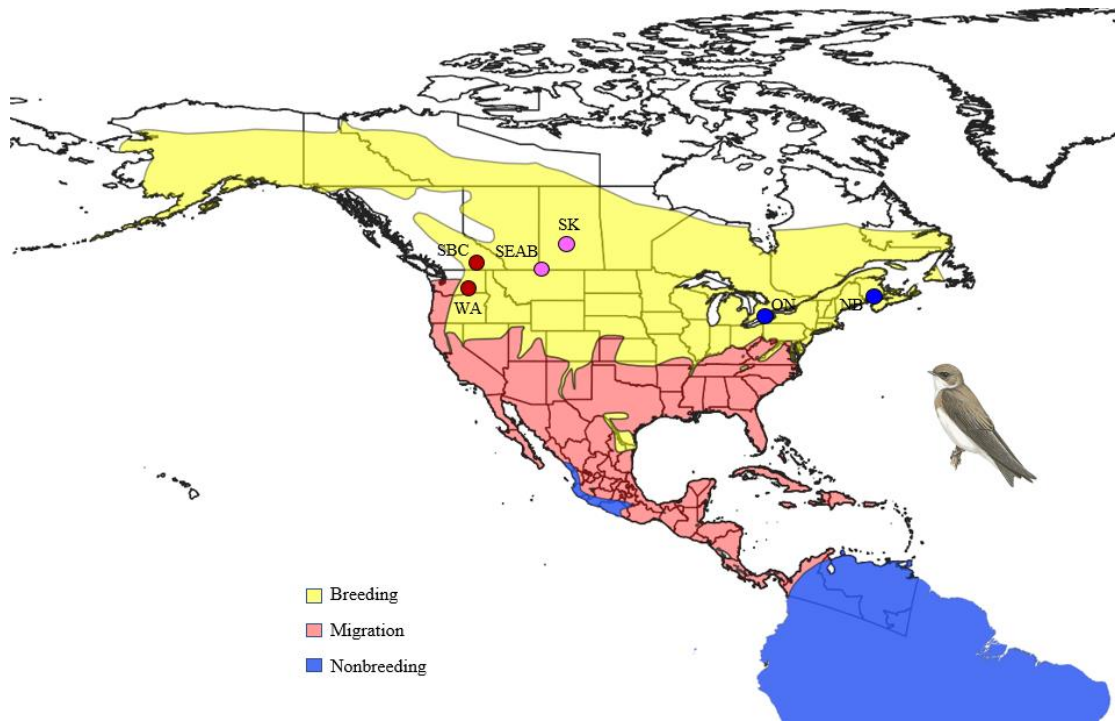


Figure 3.1: Range map of bank swallows in North America and sample locations used in this study. All samples were collected from breeding birds. Range data obtained from BirdLife International, and map was created using QGIS 3.2.

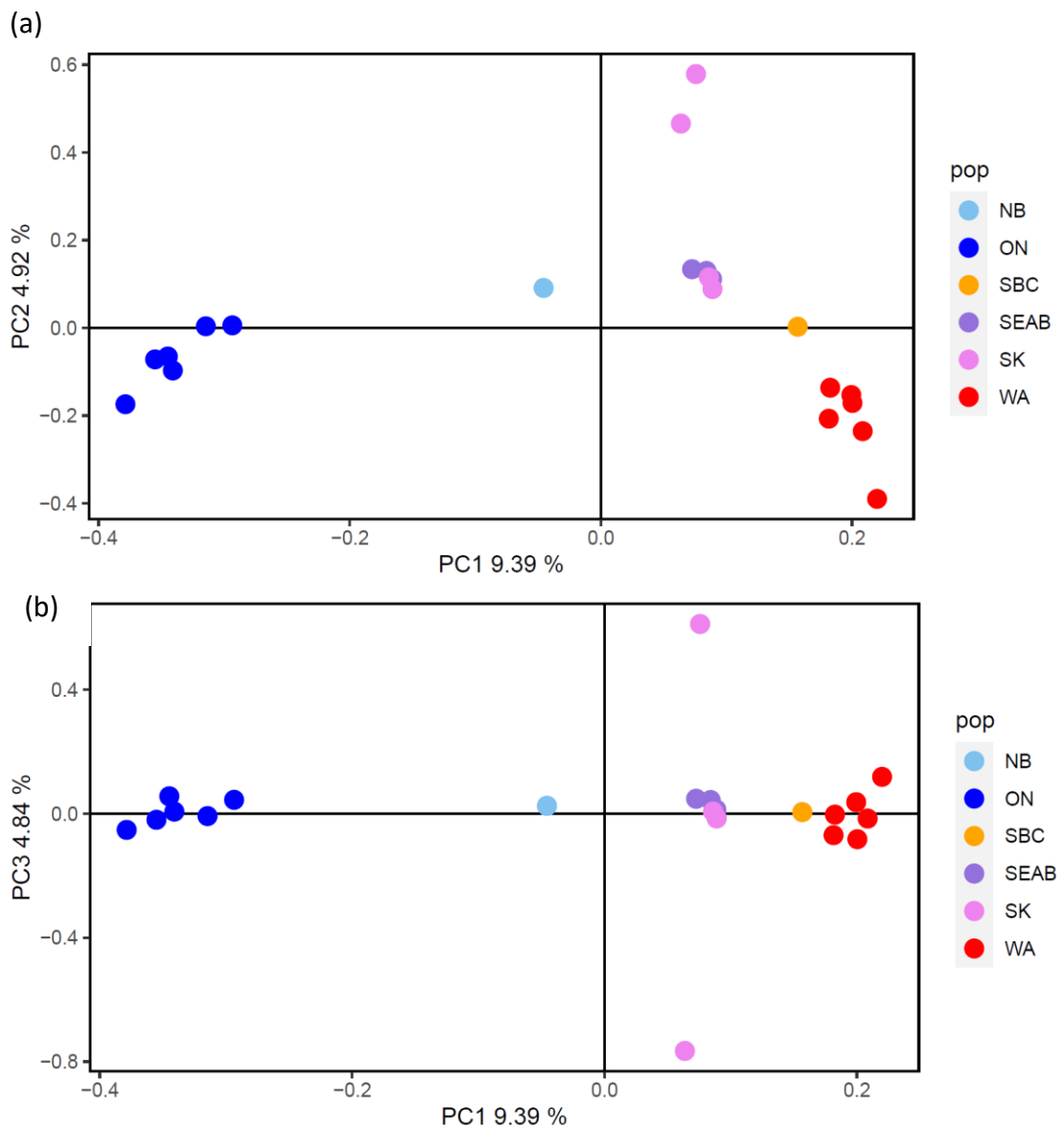


Figure 3.2: Principal component analysis for breeding populations of bank swallows using lcWGS indicating (a) PC1 vs PC2 (b) PC1 vs PC3. PC1, PC2 and PC3 explain 9.39%, 4.92% and 4.84% variance respectively.

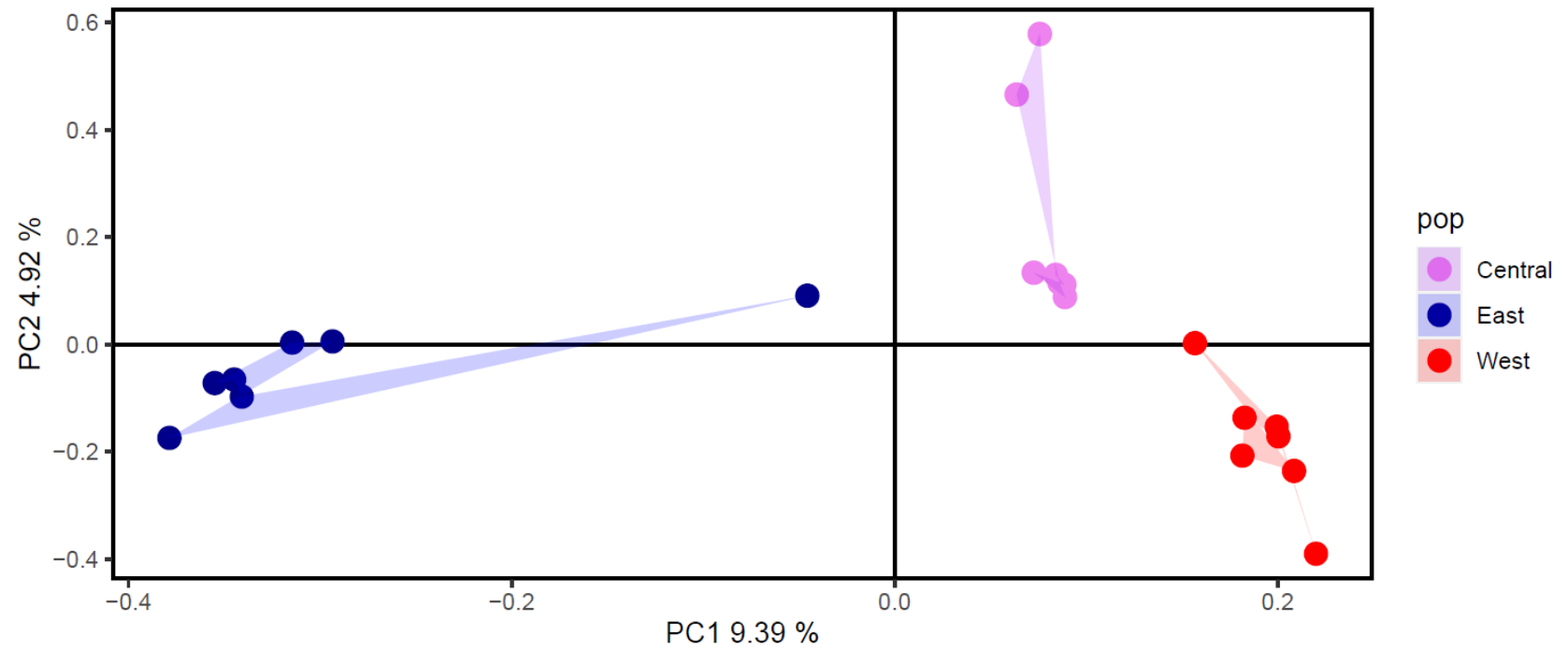


Figure 3.3: Principal component analysis for bank swallows indicating three clusters using lcWGS.

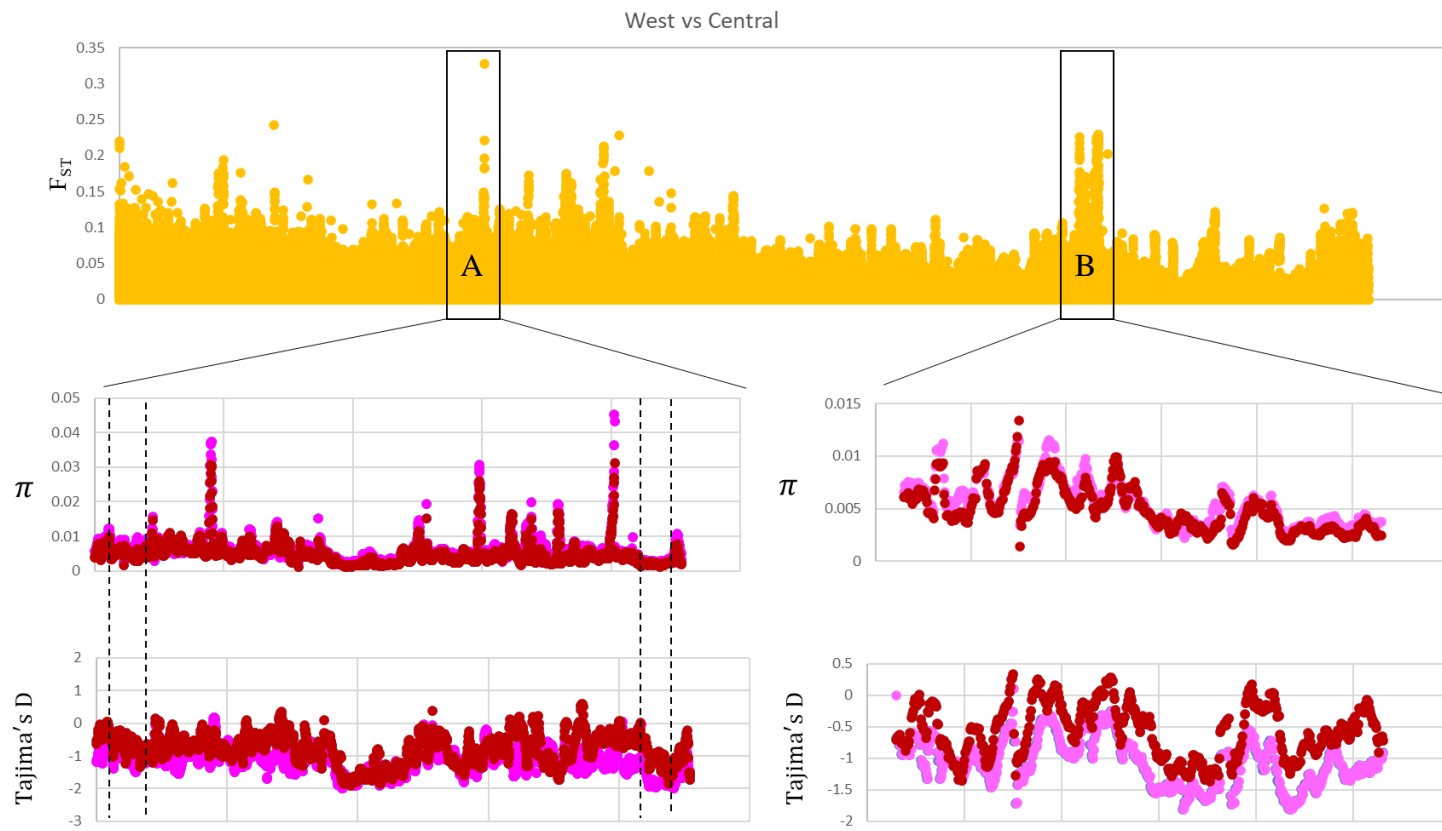


Figure 3.4: Genome scans of F_{ST} from the west (red) and central (purple) comparison. A and B indicate the high F_{ST} peaks. π and Tajima's D are included for those two regions of interest (A) scaffold JAIXNV010000025.1 (B) scaffold JAIXNV010000006.1.

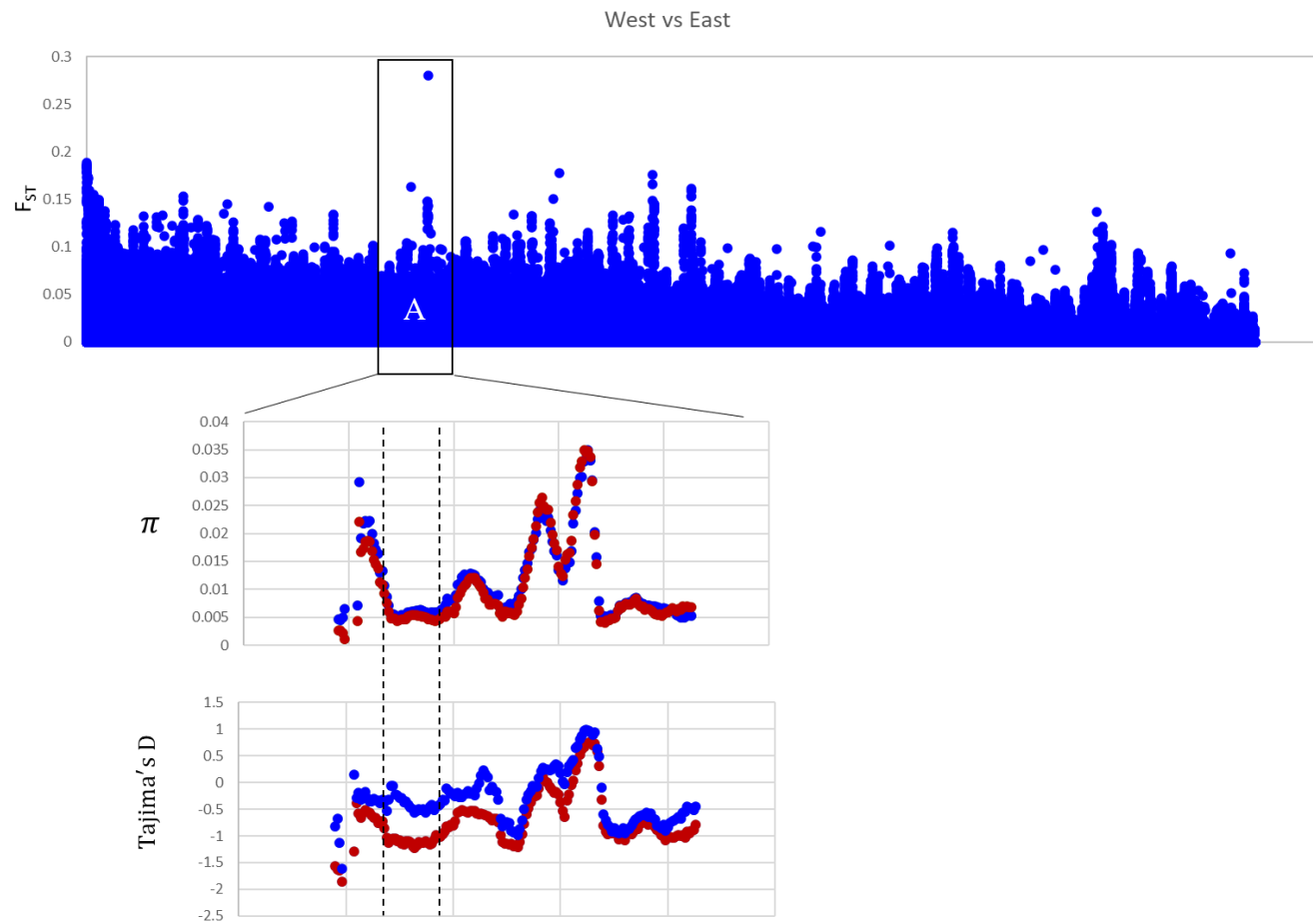


Figure 3.5: Genome scans of F_{ST} from the west (red) and east (blue) comparison. (A) scaffold JAIXNV01000025.1 indicates the high F_{ST} peak. π and Tajima's D are included for the region of interest.

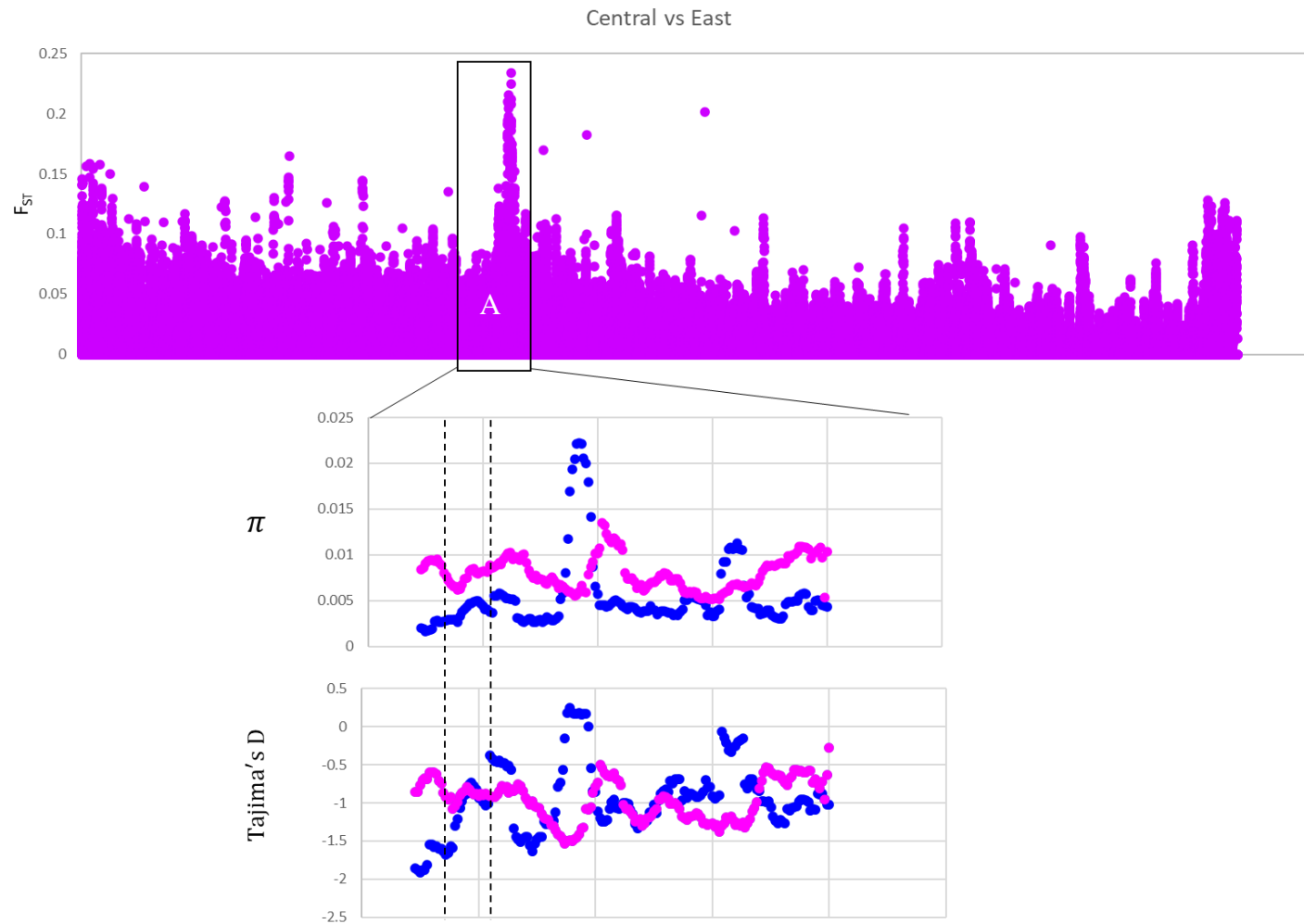


Figure 3.6: Genome scans of F_{ST} from the central (purple) and east (blue) comparison. (A) scaffold JAIXNV01000031.1 indicates the high F_{ST} peak. π and Tajima's D are included for the region of interest.



Figure 3.7: Estimated migration routes for bank swallows breeding in North America using publicly available Motus data website. Different colours indicate multiple projects and dashed lines indicate hypothesized routes.

Chapter 4: General Discussion

4.1 Aerial Insectivores

The main objective of this study was to investigate the population genetic structure of two declining Nearctic-Neotropical aerial insectivores. The levels of population decline in North American aerial insectivores vary across species and regions. Some species, such as tree swallows and barn swallows, exhibit relatively moderate declines compared to species like bank swallows (Imley & Leonard, 2019; Link & Sauer, 2002; Sauer et al., 2017). In this study, I specifically investigated the population genetic structure of two species: barn swallows, which are widely studied and experiencing a moderate decline, and bank swallows, which are less studied and undergoing a steep decline. Both species are long-distance migrants and have a wide distribution. By analyzing microsatellite data and RADseq data, I observed some ongoing gene flow in barn swallows. However, when using high resolution lcWGS data, we identified three distinct population clusters for both species. My study also highlights the importance of conducting comparative analyses using different genetic approaches (e.g., microsatellite analysis, RADseq, and lcWGS) to provide valuable insights into the strengths and limitations of each method.

4.2 Population Genetic Structure of North American Barn Swallows and Bank Swallows

I observed similar patterns of genetic structure for both species. The analyses of population structure for barn swallows revealed three distinct genomic clusters that correspond to different regions of North America: the western cluster representing samples from WA, SBC, and CA, the central cluster including samples from SK, MB, ND, and CO, and the eastern cluster consisting of samples from NB, ON, and QC. It should be noted that I have omitted the sampling sites used solely for microsatellite

analysis as no structure was observed using those markers. Similar patterns (west, central and east clusters) were observed for bank swallows; however, the coverage of sampling sites was more limited for this species. Added to that, the range is not as expansive as barn swallows across North America, and their breeding range does not extend as far south (Garrison, 1999). For bank swallows, the western cluster includes populations from SBC and WA, the central cluster comprises samples from SEAB and SK, and the eastern cluster contains samples from NB and ON. For future studies, it would be valuable to include samples from Alaska and Yukon and more samples from NB to further enhance our understanding of the overall population genetic structure of bank swallows.

Population genetic differences can arise as a result of barriers to gene flow, which can be either historical or contemporary barriers (Harris & Reed, 2002; Ravinet et al., 2017). It is crucial to understand population structure especially in species experiencing declines, in order to develop effective conservation strategies.

Populations that display genetic distinctiveness, in conjunction with ecological and behavioural differences, may be treated as separate conservation units and prioritized accordingly for conservation efforts.

The three distinct populations observed in both barn swallows and bank swallows could be a result of survival in three refugia during the Pleistocene glaciations.

Isolation in different refugia restricted gene flow, contributing to genetic differentiation between populations (Burg et al., 2006; Shafer et al., 2010). Based on my analysis of the genetic structure, I have identified three distinct and well-defined clusters for bank swallows, indicating clear genetic isolation. However, for barn swallows, the divergence between the west and central clusters is lower in some of the analyses, suggesting a more recent separation. It is also possible central and

western populations of barn swallows survived in a similar refugia in terms of habitats, while the western and central populations of bank swallows survived in more diverse refugia. During the Pleistocene, there were multiple glaciation events, but I have not been able to determine the exact timing of the population splits. The variation in the timing of separation for these populations could have resulted from the different impacts of these glaciation events. Possible refugia for both species could be Beringia and the southern refugia in the Pacific Northwest in the west and southeastern and Atlantic coastal North America in the east.

Glaciation events also caused range shifts and shaped the migratory pathways (Thorup et al., 2021). For example, the changes resulting from glaciation, such as the formation and melting of ice sheets, and shifts in the landscape, have led to alterations in the availability of suitable habitats and the emergence of geographic barriers resulting in obstructed previous migratory routes. As a consequence, birds and other species were forced to adapt and find alternative pathways to move between their breeding and wintering grounds. Some species/populations may have followed newly exposed land routes, while others may have utilized coastal regions that remained ice-free. Geographical features such as the Rocky Mountains, Appalachian Mountains, Great Lakes, and the Mississippi River also further influenced migratory pathways and played a role in shaping the genomic differences between populations of both study species and other species. While long-distance migrants could easily fly over these barriers, it is likely that most birds avoid migrating over them. For example, the migratory network of tree swallows depicted in Knight et al., (2018) demonstrates that central populations migrate along the Mississippi River Valley. Therefore, the population genetic structure of both barn swallows and bank swallows is strongly

influenced by glaciations and related factors, including migratory routes that are shaped by existing barriers.

The utilization of distinct migratory routes by different populations within the same species can result in migratory divides that prevent the further mixing of populations. A continental-wide migratory divide was identified for barn swallows across North America using geolocators (Hobson et al., 2015). Similarly, bank swallows breeding in Atlantic Canada, exhibit comparable migratory patterns to barn swallows, with wintering grounds located between northern Argentina and southern Brazil (Imlay et al., 2020). However, the lack of comprehensive studies focused on the continental-wide migratory pathways of bank swallows are limited and it hampers our ability to predict how these migratory routes influence the genetic structure of these species. To address this, I utilized publicly available geocator data to gain preliminary insights into the migratory patterns of bank swallows. These data revealed similar patterns between bank swallow populations west of the Rocky Mountains and those along the east coast to those of barn swallows. However, I observed that birds from the Yukon region and Alaska do not cross over the Rocky Mountains but instead fly to the east of the mountain range with the central group. The genetic clusters I identified align well with the distinct migratory pathways observed in both species, providing further support for the connection between migration patterns and genetic structure. Similar patterns which linked migratory divides with the genetic structure were found for Swainson's thrush (*Catharus ustulatus*) and European blackcaps (*Sylvia atricapilla*) (Berthold, 2001; Delmore et al., 2012; Delmore et al., 2020).

4.3 Migratory Connectivity Using Genomic Approaches

In addition to the breeding samples, I also included migratory and overwintering individuals in my study to gain insights into the migratory connectivity. In my analyses, I incorporated migratory samples of barn swallows from SK, MB, and CO, along with overwintering samples from COL. Interestingly, I found that all except one of these migratory and overwintering samples clustered together with the central breeding population, which was further supported by the migratory pathways analyzed using geolocators.

Interestingly, I observed that one migratory individual of barn swallows from CO clustered with the western population in the PCA (Figure 2.5 a and b). This migratory individual could suggest some degree of mixing between western and central populations during migration. To gain a better understanding of migratory pathways across North America using genomics, future studies should consider including more migratory and overwintering samples from both the western and eastern clusters.

I could not explore the migratory connectivity or pathways for bank swallows using genomic data due to lack of migratory samples. It would be valuable to have broader coverage and include migratory and overwintering samples in future studies to shed light on these aspects for bank swallows.

4.4 Usage of Different Markers in a Population Genetic Study

In this study, I employed three different genetic approaches. Traditional techniques such as microsatellite genotyping and mitochondrial sequencing offer valuable insights into natural populations (Allendorf, 2017). The cost of these methods is also low compared to next-generation sequencing and I can use samples with less or lower quality DNA, allowing for the incorporation of a larger sample size. However, their

limited number of genetic markers hampers our ability to achieve comprehensive resolution, particularly in population-level studies. To overcome this limitation, I utilized advanced next-generation sequencing techniques to explore the fine-scale population genetic structure across the entire genome (Holliday et al., 2019; Lou et al., 2021).

In my study, I utilized microsatellite analysis, RADseq analysis, and lcWGS analysis for barn swallows. Microsatellite analysis allowed for the utilization of a large number of samples, but with lower resolution. Additionally, next-generation sequencing (NGS) methods offered higher resolution but were constrained by a limited number of samples due to their cost and the requirement for high-quality DNA. In comparison, RADseq provided higher resolution but a lower number of samples, while lcWGS offered the highest resolution but also had a lower number of samples. However, due to limited sampling coverage and sample size, I focused solely on lcWGS for bank swallows, as it is a more sophisticated method to gain insights into the fine-scale population genetic structure with a smaller number of samples.

4.5 Conclusions and Future Directions

This study provides valuable insights into the population genetic structure of two declining Nearctic Neotropical aerial insectivores. Overall, I observed similar genetic clusters west of the Rockies, central and eastern North America for both barn swallows and bank swallows. While three distinct and well-defined clusters were identified for bank swallows, indicating clear genetic isolation, barn swallows exhibited less differentiation between the west and central clusters in some analyses. The eastern population of barn swallows displayed a high genetic differentiation

compared to the other populations and is experiencing a rapid decline, indicating the need for separate conservation considerations. The presence of three distinct genetic clusters in bank swallows indicates the importance of considering separate conservation units as well as proper prioritization in conservation plans.

To further enhance our understanding of barn swallows, it will be important to include samples from ON for lcWGS in future studies, as I observed the separation of ON samples from NB samples in the RADseq analysis. For lcWGS I was unable to obtain enough high-quality DNA in the samples from ON. Despite having both samples from ON and NB included in the RADseq analyses, the NB samples exhibited a significant amount of missing data. As a result, I refrained from drawing any conclusions based on the NB samples of the RADseq analyses. Additionally, incorporating migratory samples from western and eastern clusters would provide valuable insights into their continental-wide migratory connectivity using genomic approaches.

For bank swallows, increasing the sampling coverage and sample size for populations in SBC, SEAB, and NB is crucial for unravelling the fine-scale population-specific genetic structure. Similarly, incorporating migratory and overwintering samples would greatly contribute to understanding the migratory connectivity of bank swallows. Given the limited research on the migratory and ecological aspects of bank swallows, integrating genetic data with other relevant factors is necessary for effective conservation planning. My study is the first to perform whole genome sequencing for the continental-wide population of North American bank swallows. Therefore, conducting more genetic studies, particularly those focused on developing a fully annotated reference genome for bank swallows, would be invaluable in gaining insights into the specific gene regions that are under selection.

My study shows conducting comparative analyses using different genetic approaches (e.g., microsatellite analysis, RADseq, and lcWGS) can provide valuable insights into the strengths and limitations of each method. This contributes to the refinement of the study design and the selection of appropriate techniques for future population genetic studies in declining aerial insectivores and other species.

Investigating the interplay between historical factors (e.g., Pleistocene glaciations, past refugia) and contemporary factors (e.g., geographic features, migratory routes) in shaping the population genetic structure of aerial insectivores is a promising avenue for future research. Integrating historical demographic modelling along with landscape genetics, and spatial analysis techniques would enhance our understanding of the complex interactions between genetic variation and environmental factors.

Furthermore, connecting the observed genetic structures with ecological and behavioural factors is crucial. Identifying genetically distinct populations and understanding their specific ecological requirements will contribute to the development of targeted conservation strategies. This may involve delineating separate conservation units, prioritizing management efforts, and considering the impacts of migratory pathways and barriers in effective conservation planning.

References

- Berthold, P. (2001). *Bird Migration: A General Survey*. Oxford University Press.
- Burg, T. M., Gaston, A. J., Winker, K., & Friesen, V. L. (2006). Effects of Pleistocene glaciations on population structure of North American chestnut-backed chickadees. *Molecular Ecology*, *15*(9), 2409-2419.
- Delmore, K. E., Fox, J. W., & Irwin, D. E. (2012). Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1747), 4582-4589.
- Delmore, K. E., Van Doren, B. M., Conway, G. J., Curk, T., Garrido-Garduño, T., Germain, R. R., Hasselmann, T., Hiemer, D., van der Jeugd, H. P., & Justen, H. (2020). Individual variability and versatility in an eco-evolutionary model of avian migration. *Proceedings of the Royal Society B*, *287*(1938), 20201339.
- Garrison, B. A. (1999). Bank swallow (*Riparia riparia*). *The Birds of North America* 414, 28.
- Harris, R. J., & Reed, J. M. (2002). Behavioral barriers to non-migratory movements of birds. *Annales Zoologici Fennici* 39: 275-290.
- Hobson, K. A., Kardynal, K. J., Van Wilgenburg, S. L., Albrecht, G., Salvadori, A., Cadman, M. D., Liechti, F., & Fox, J. W. (2015). A continent-wide migratory divide in North American breeding barn swallows (*Hirundo rustica*). *PloS One*, *10*(6), e0129340.
- Holliday, J. A., Hallerman, E. M., & Haak, D. C. (2019) Genotyping and sequencing technologies in population genetics and genomics. In O. P. Rajora (Ed.), *Population genomics: Concepts, Approaches and Applications* (pp. 83– 125). Springer Nature Switzerland AG.
- Imlay, T., Saldanha, S., & Taylor, P. (2020). The fall migratory movements of bank swallows, *Riparia riparia*: Fly-and-forage migration? *Avian Conservation and Ecology*, *15*(1), 2.
- Imlay, T. L., & Leonard, M. L. (2019). A review of the threats to adult survival for swallows (family: *Hirundinidae*). *Bird Study*, *66*(2), 251-263.
- Knight, S. M., Bradley, D. W., Clark, R. G., Gow, E. A., Bélisle, M., Berzins, L. L., Blake, T., Bridge, E. S., Burke, L., & Dawson, R. D. (2018). Constructing and evaluating a continent-wide migratory songbird network across the annual cycle. *Ecological Monographs*, *88*(3), 445-460.
- Link, W. A., & Sauer, J. R. (2002). A hierarchical analysis of population change with application to cerulean warblers. *Ecology*, *83*(10), 2832-2840.
- Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, *30*(23), 5966-5993.
- Ravinet, M., Faria, R., Butlin, R., Galindo, J., Bierne, N., Rafajlović, M., Noor, M., Mehlig, B., & Westram, A. (2017). Interpreting the genomic landscape of

speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450-1477.

Sauer, J. R., Pardieck, K. L., Ziolkowski Jr, D. J., Smith, A. C., Hudson, M.-A. R., Rodriguez, V., Berlanga, H., Niven, D. K., & Link, W. A. (2017). The first 50 years of the North American breeding bird survey. *The Condor: Ornithological Applications*, 119(3), 576-593.

Shafer, A. B., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, 19(21), 4589-4621.

Thorup, K., Pedersen, L., Da Fonseca, R. R., Naimi, B., Nogués-Bravo, D., Krapp, M., & Rahbek, C. (2021). Response of an Afro-Palearctic bird migrant to glaciation cycles. *Proceedings of the National Academy of Sciences*, 118(52), e2023836118.

Appendices

Appendix 1: Supplementary Information for Chapter 2

Landscape Genetic Structure of Barn swallows (*Hirundo rustica*) Breeding in North America

Appendix 1.1: Individual sample details for barn swallows (BARS) included in this study: An asterisk (*) next to the sample ID indicates that the sample was used for two types of analyses.

Location	Species	Sample ID	Band/Museum ID	Latitude	Longitude
Samples used for Microsatellite Analyses					
Delta, BC	BARS	BC001	1840-78415	49.10099	-123.17142
Delta, BC	BARS	BC002	1840-78416	49.10099	-123.17142
Delta, BC	BARS	BC003	1840-78427	49.10099	-123.17142
Delta, BC	BARS	BC004	1840-78428	49.10099	-123.17142
Delta, BC	BARS	BC005	1840-78429	49.10099	-123.17142
Delta, BC	BARS	BC006	1840-78439	49.10099	-123.17142
Delta, BC	BARS	BC007	1840-78440	49.10099	-123.17142
Delta, BC	BARS	BC008	1840-78441	49.10099	-123.17142
Delta, BC	BARS	BC009	1840-78497	49.10099	-123.17142
Delta, BC	BARS	BC010	1840-78498	49.10099	-123.17142
Delta, BC	BARS	BC011	Unbanded	49.10099	-123.17142
Delta, BC	BARS	BC012	Unbanded	49.10099	-123.17142
Prince Albert National Park, SK	BARS	SK001	2610-78403	53.585064	-106.049769
Prince Albert National Park, SK	BARS	SK002	2610-78405	53.585064	-106.049769
Prince Albert National Park, SK	BARS	SK003	2610-78407	53.585064	-106.049769
Prince Albert National Park, SK	BARS	SK004	2610-78410	53.585064	-106.049769
Prince Albert National Park, SK	BARS	SK005	2610-78478	53.585064	-106.049769
Prince Albert National Park, SK	BARS	SK006	2610-78860	53.585064	-106.049769
Saskatoon, SK	BARS	SK043	2621-69100	52.146973	-106.647034
Saskatoon, SK	BARS	SK044	2621-69094	52.146973	-106.647034
Saskatoon, SK	BARS	SK045	2621-69095	52.146973	-106.647034
Delta Marsh, MB	BARS	MB001*	14988	50.17148	-97.13222
Delta Marsh, MB	BARS	MB002	14996	50.17148	-97.13222
Delta Marsh, MB	BARS	MB003	14997	50.17148	-97.13222
Delta Marsh, MB	BARS	MB004	14998	50.17148	-97.13222

Delta Marsh, MB	BARS	MB005	14999	50.17148	-97.13222
Delta Marsh, MB	BARS	MB006	15357	50.17148	-97.13222
Delta Marsh, MB	BARS	MB007	15361	50.17148	-97.13222
Delta Marsh, MB	BARS	MB008	15364	50.17148	-97.13222
Delta Marsh, MB	BARS	MB009	15366	50.17148	-97.13222
Delta Marsh, MB	BARS	MB010	15368	50.17148	-97.13222
Delta Marsh, MB	BARS	MB011	15370	50.17148	-97.13222
Delta Marsh, MB	BARS	MB012	15371	50.17148	-97.13222
Delta Marsh, MB	BARS	MB013	15372	50.17148	-97.13222
Delta Marsh, MB	BARS	MB014	15373	50.17148	-97.13222
Delta Marsh, MB	BARS	MB015	15374	50.17148	-97.13222
Delta Marsh, MB	BARS	MB016	15375	50.17148	-97.13222
Delta Marsh, MB	BARS	MB017	15376	50.17148	-97.13222
Delta Marsh, MB	BARS	MB018	15377	50.17148	-97.13222
Delta Marsh, MB	BARS	MB019	15378	50.17148	-97.13222
Delta Marsh, MB	BARS	MB020	15379	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB021*	2721-82291	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB022*	2721-82294	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB023*	2721-82304	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB024*	2721-82305	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB025*	2721-82306	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB026*	2721-82307	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB027*	2721-82308	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB028	2721-82309	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB029	2721-82310	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB030	2721-82311	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB031	2721-82312	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB032	2721-82313	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB033	2721-82314	50.17148	-97.13222

Oak Hammock Marsh, MB	BARS	MB034	2721-82315	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB035	2721-82317	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB036	2721-82318	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB037	2721-82319	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB038	2721-82320	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB039	2721-82326	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB040	2721-82328	50.17148	-97.13222
Centroid, TX	BARS	TX001	11BGP558	31.93859	-99.90181
Centroid, TX	BARS	TX002	11BGP559	31.93859	-99.90181
Centroid, TX	BARS	TX003	11BGP560	31.93859	-99.90181
Centroid, TX	BARS	TX004	11BGP561	31.93859	-99.90181
Colombia	BARS	COL001	BB09001	4.57086	-74.29733
Colombia	BARS	COL002	BB09002	4.57086	-74.29733
Colombia	BARS	COL003	BB09003	4.57086	-74.29733
Colombia	BARS	COL004	BB09004	4.57086	-74.29733
Colombia	BARS	COL005*	BB09005	4.57086	-74.29733
Colombia	BARS	COL006*	BB09006	4.57086	-74.29733
Colombia	BARS	COL007*	BB09007	4.57086	-74.29733
Colombia	BARS	COL008	BB09008	4.57086	-74.29733
Colombia	BARS	COL009*	BB09009	4.57086	-74.29733
Colombia	BARS	COL010*	BB09010	4.57086	-74.29733
Colombia	BARS	COL011	BB09011	4.57086	-74.29733
Colombia	BARS	COL012	BB09012	4.57086	-74.29733
Wellington, ON	BARS	ON_001	2631-83814	43.556278	-80.18375
Wellington, ON	BARS	ON_002*	2631-83821	43.556278	-80.18375
Wellington, ON	BARS	ON_003	2631-83831	43.556278	-80.18375
Wellington, ON	BARS	ON_004*	2631-83852	43.556278	-80.18375
Wellington, ON	BARS	ON_005*	2631-83846	43.556278	-80.18375
Wellington, ON	BARS	ON_006	2631-83838	43.556278	-80.18375

Wellington, ON	BARS	ON_007*	2631-83865	43.556278	-80.18375
Wellington, ON	BARS	ON_008	2631-83875	43.556278	-80.18375
Wellington, ON	BARS	ON_009	2631-91299	43.556278	-80.18375
Wellington, ON	BARS	ON_010	2631-91304	43.556278	-80.18375
Wellington, ON	BARS	ON_011	2631-91027	43.556278	-80.18375
Wellington, ON	BARS	ON_012	2631-91308	43.556278	-80.18375
Wellington, ON	BARS	ON_013	2631-91313	43.556278	-80.18375
Wellington, ON	BARS	ON_014*	2631-91318	43.556278	-80.18375
Wellington, ON	BARS	ON_015*	2631-91116	43.556278	-80.18375
Wellington, ON	BARS	ON_016	2631-83900	43.556278	-80.18375
Wellington, ON	BARS	ON_017	2631-91103	43.556278	-80.18375
Wellington, ON	BARS	ON_018	2631-91119	43.556278	-80.18375
Wellington, ON	BARS	ON_019	2631-90933	43.556278	-80.18375
Wellington, ON	BARS	ON_020	2631-90904	43.556278	-80.18375
Wellington, ON	BARS	ON_021	2631-91184	43.556278	-80.18375
Wellington, ON	BARS	ON_022*	2631-91172	43.556278	-80.18375
Wellington, ON	BARS	ON_023	2631-91422	43.556278	-80.18375
Wellington, ON	BARS	ON_024*	2631-91642	43.556278	-80.18375
Wellington, ON	BARS	ON_025*	2631-91444	43.556278	-80.18375
Wellington, ON	BARS	ON_026	2631-90607	43.556278	-80.18375
Wellington, ON	BARS	ON_027	2631-90615	43.556278	-80.18375
Wellington, ON	BARS	ON_028	2631-90632	43.556278	-80.18375
Wellington, ON	BARS	ON_029	2631-90637	43.556278	-80.18375
Wellington, ON	BARS	ON_030	2631-91459	43.556278	-80.18375
Wellington, ON	BARS	ON_031	2631-90663	43.556278	-80.18375
Wellington, ON	BARS	ON_032	2631-91483	43.556278	-80.18375
Wellington, ON	BARS	ON_033	2631-91473	43.556278	-80.18375
Wellington, ON	BARS	ON_034	2631-90667	43.556278	-80.18375
Wellington, ON	BARS	ON_035	2631-90711	43.556278	-80.18375

Wellington, ON	BARS	ON_036	2631-91524	43.556278	-80.18375
Wellington, ON	BARS	ON_037	2631-91532	43.556278	-80.18375
Wellington, ON	BARS	ON_038	2631-90743	43.556278	-80.18375
Wellington, ON	BARS	ON_039	2631-90752	43.556278	-80.18375
Wellington, ON	BARS	ON_040	2631-91672	43.556278	-80.18375
Wellington, ON	BARS	ON_041	2631-91016	43.556278	-80.18375
Wellington, ON	BARS	ON_042	2631-90789	43.556278	-80.18375
Wellington, ON	BARS	ON_043	2631-91572	43.556278	-80.18375
Wellington, ON	BARS	ON_044	2631-91551	43.556278	-80.18375
Wellington, ON	BARS	ON_045	2631-91562	43.556278	-80.18375
Wellington, ON	BARS	ON_046	2631-91607	43.556278	-80.18375
Wellington, ON	BARS	ON_047	2631-91047	43.556278	-80.18375
Wellington, ON	BARS	ON_048	2631-91594	43.556278	-80.18375
Wellington, ON	BARS	ON_049	2631-91599	43.556278	-80.18375
Wellington, ON	BARS	ON_050	2631-91052	43.556278	-80.18375
Wellington, ON	BARS	ON_051	2631-91089	43.556278	-80.18375
Wellington, ON	BARS	ON_052	2631-90886	43.556278	-80.18375
Wellington, ON	BARS	ON_053	2631-90865	43.556278	-80.18375
Wellington, ON	BARS	ON_054	2631-90861	43.556278	-80.18375
Wellington, ON	BARS	ON_055	2631-90890	43.556278	-80.18375
Wellington, ON	BARS	ON_056	2631-90893	43.556278	-80.18375
Wellington, ON	BARS	ON_057	2631-91722	43.556278	-80.18375
Wellington, ON	BARS	ON_058	2631-91742	43.556278	-80.18375
Wellington, ON	BARS	ON_059	2631-91743	43.556278	-80.18375
Wellington, ON	BARS	ON_060	2631-91757	43.556278	-80.18375
Wellington, ON	BARS	ON_061	2631-91769	43.556278	-80.18375
Wellington, ON	BARS	ON_062	2631-91776	43.556278	-80.18375
Wellington, ON	BARS	ON_063	2631-91797	43.556278	-80.18375
Wellington, ON	BARS	ON_064	2631-91801	43.556278	-80.18375

Wellington, ON	BARS	ON_065	2631-91810	43.556278	-80.18375
Wellington, ON	BARS	ON_066	2631-91846	43.556278	-80.18375
Wellington, ON	BARS	ON_067	2631-91851	43.556278	-80.18375
Wellington, ON	BARS	ON_068	2631-91872	43.556278	-80.18375
Wellington, ON	BARS	ON_069	2631-91858	43.556278	-80.18375
Wellington, ON	BARS	ON_070	2631-91855	43.556278	-80.18375
Wellington, ON	BARS	ON_071	2631-91815	43.556278	-80.18375
Wellington, ON	BARS	ON_072	2631-91868	43.556278	-80.18375
Wellington, ON	BARS	ON_073	2631-91844	43.556278	-80.18375
Wellington, ON	BARS	ON_074	2631-91892	43.556278	-80.18375
Wellington, ON	BARS	ON_075	2631-91888	43.556278	-80.18375
Wellington, ON	BARS	ON_076	2631-91882	43.556278	-80.18375
Wellington, ON	BARS	ON_077	2631-91883	43.556278	-80.18375
Wellington, ON	BARS	ON_078	2631-91879	43.556278	-80.18375
Wellington, ON	BARS	ON_079	2631-91925	43.556278	-80.18375
Wellington, ON	BARS	ON_080	2631-91920	43.556278	-80.18375
Wellington, ON	BARS	ON_081	2631-91911	43.556278	-80.18375
Wellington, ON	BARS	ON_082	2631-91940	43.556278	-80.18375
Wellington, ON	BARS	ON_083	2631-91943	43.556278	-80.18375
Wellington, ON	BARS	ON_084	2631-91966	43.556278	-80.18375
Wellington, ON	BARS	ON_085	2631-91972	43.556278	-80.18375
Quebec	BARS	QC001*	13023	46.44879	-73.03525
Quebec	BARS	QC002	13024	46.44879	-73.03525
Quebec	BARS	QC003*	13025	46.44879	-73.03525
Quebec	BARS	QC004*	13026	46.44879	-73.03525
Quebec	BARS	QC005*	13027	46.44879	-73.03525
Quebec	BARS	QC006	13028	46.44879	-73.03525
Quebec	BARS	QC007*	13029	46.44879	-73.03525
Holtville, NB	BARS	NB001*	2560-01395	46.56531	-66.46191

Holtville, NB	BARS	NB002*	2560-01543	46.56531	-66.46191
Holtville, NB	BARS	NB003	2560-01580	46.56531	-66.46191
Holtville, NB	BARS	NB004	2560-01582	46.56531	-66.46191
Holtville, NB	BARS	NB005	2560-01583	46.56531	-66.46191
Holtville, NB	BARS	NB006	2560-01586	46.56531	-66.46191
Holtville, NB	BARS	NB007	2610-78496	46.56531	-66.46191

Samples used for RADseq Analyses

Auburn, AL	BARS	AL_Aub_05	2610-78419	32.609856	-85.480782
Auburn, AL	BARS	AL_Aub_06	2610-78420	32.609856	-85.480782
Auburn, AL	BARS	AL_Aub_07	2610-78379	32.609856	-85.480782
Degenhardt, NB	BARS	NB_Deg_01	2560-01427	46.56531	-66.46191
Degenhardt, NB	BARS	NB_Deg_03	2560-01436	46.56531	-66.46191
Degenhardt, NB	BARS	NB_Deg_05	2560-01438	46.56531	-66.46191
McWhirter, NB	BARS	NB_McW_01	2560-01446	46.56531	-66.46191
McWhirter, NB	BARS	NB_McW_02	2560-01447	46.56531	-66.46191
McWhirter, NB	BARS	NB_McW_04	2560-01449	46.56531	-66.46191
Wellington, ON	BARS	ON_Wel_01*	2631-91116	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_02*	2631-91172	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_03*	2631-83821	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_04*	2631-83846	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_05*	2631-83865	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_07*	2631-91642	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_08*	2631-91318	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_09*	2631-83852	43.556278	-80.18375
Wellington, ON	BARS	ON_Wel_10*	2631-91444	43.556278	-80.18375
London, ON	BARS	SWON_01	106	42.983612	-81.249725
London, ON	BARS	SWON_02	108	42.983612	-81.249725
London, ON	BARS	SWON_03	110	42.983612	-81.249725
London, ON	BARS	SWON_04	111	42.983612	-81.249725

London, ON	BARS	SWON_05	114	42.983612	-81.249725
London, ON	BARS	SWON_06	166	42.983612	-81.249725
London, ON	BARS	SWON_07	167	42.983612	-81.249725
London, ON	BARS	SWON_08	168	42.983612	-81.249725
Quill Lake, SK	BARS	SK_Qui_01	1561-25109	52.072553	-104.253704
Quill Lake, SK	BARS	SK_Qui_05	1561-25116	52.072553	-104.253704
Suderman, SK	BARS	SK_Sud_01	1561-25162	52.072553	-104.253704
Suderman, SK	BARS	SK_Sud_02	1561-25163	52.072553	-104.253704
Suderman, SK	BARS	SK_Sud_03	1561-25164	52.072553	-104.253704
Suderman, SK	BARS	SK_Sud_04	1561-25165	52.072553	-104.253704
Suderman, SK	BARS	SK_Sud_05	1561-25171	52.072553	-104.253704
Heather Husband 1124 73rd St NE Towner, ND	BARS	ND_Tow_01	ND11	48.471434	-100.33701
Heather Husband 1124 73rd St NE Towner, ND	BARS	ND_Tow_02	ND131	48.471434	-100.33701
Seattle, WA	BARS	WA_Sea_02	2610-78871	47.608013	-122.335167
Seattle, WA	BARS	WA_Sea_03	2610-78868	47.608013	-122.335167
Seattle, WA	BARS	WA_Sea_04	2610-78870	47.608013	-122.335167
Seattle, WA	BARS	WA_Sea_07	2510-41141	47.608013	-122.335167

Samples used for lcWGS Analyses

Arcata, CA	BARS	CA003	2290-64682	40.866516	-124.08284
Arcata, CA	BARS	CA007	2290-64686	40.866516	-124.08284
Denver International Airport, CO	BARS	CO020_SM	DMNS:44459	39.85833	-104.66694
Denver International Airport, CO	BARS	CO022_FM	DMNS:46202	39.85833	-104.66694
Denver International Airport, CO	BARS	CO043_FM	DMNS:53135	39.85833	-104.66694
Denver, City Park, Ferril Lake, on sidewalk, CO	BARS	CO047_SM	DMNS:55634	39.85833	-104.66694
Colombia	BARS	COL005_W*	BB09005	4.57086	-74.29733
Colombia	BARS	COL006_W*	BB09006	4.57086	-74.29733
Colombia	BARS	COL007_W*	BB09007	4.57086	-74.29733
Colombia	BARS	COL009_W*	BB09009	4.57086	-74.29733
Colombia	BARS	COL010_W*	BB09010	4.57086	-74.29733

Oak Hammock Marsh, MB	BARS	MB001*	14988	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB021_M*	2721-82291	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB022_M*	2721-82294	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB023_M*	2721-82304	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB024_M*	2721-82305	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB025_M*	2721-82306	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB026_M*	2721-82307	50.17148	-97.13222
Oak Hammock Marsh, MB	BARS	MB027_M*	2721-82308	50.17148	-97.13222
Holtville, NB	BARS	NB001*	2560-01395	46.56531	-66.46191
Holtville, NB	BARS	NB002*	2560-01543	46.56531	-66.46191
GAUDET near Memramcook, NB	BARS	NB034	2980-00701	45.983913	-64.572362
GAUDET near Memramcook, NB	BARS	NB036	2980-00703	45.983913	-64.572362
GAUDET near Memramcook, NB	BARS	NB037	2980-00704	45.983913	-64.572362
Quebec	BARS	QC001*	13023	46.44879	-73.03525
Quebec	BARS	QC003*	13025	46.44879	-73.03525
Quebec	BARS	QC004*	13026	46.44879	-73.03525
Quebec	BARS	QC005*	13027	46.44879	-73.03525
Quebec	BARS	QC007*	13029	46.44879	-73.03525
Lower Mainland, BC	BARS	SBC002	42247	49.16659	-123.133569
Port Coquitlam, Lower Mainland, BC	BARS	SBC003	41813	49.16659	-123.133569
Richmond, Lower Mainland, BC	BARS	SBC004	41819	49.16659	-123.133569
Pitt Meadows: 18598 Advert Rd. Lower Mainland, BC	BARS	SBC005	44369	49.16659	-123.133569
Oxbow, SK	BARS	SK008_M	RSKM_A-5955	49.23333	-102.1667
Estevan, SK	BARS	SK009_M	RSKM_A-9330	49.09533	-103.0324
Pangman, SK	BARS	SK010_SM	RSKM_A-9517	49.63095	-104.716
Saskatoon, SK	BARS	SK011	19BS25	52.146973	-106.64703
Saskatoon, SK	BARS	SK012	19BS26	52.146973	-106.64703
Saskatoon, SK	BARS	SK013	19BS30	52.146973	-106.64703

Saskatoon, SK	BARS	SK014	19BS28	52.146973	-106.64703
Saskatoon, SK	BARS	SK016	19BS29	52.146973	-106.64703
Washington	BARS	WA001	2140-03406	47.66661	-122.35217
Washington	BARS	WA002	2460-53100	47.66661	-122.35217
Washington	BARS	WA003	2460-98062	47.66661	-122.35217
Washington	BARS	WA004	2460-98216	47.66661	-122.35217
Washington	BARS	WA005	2510-41101	47.66661	-122.35217

Appendix 2: Supplementary Information for Chapter 3

Population Genetic Structure of North American Bank Swallow (*Riparia riparia*)

Appendix 2.1: Individual sample details for bank swallows (BANS) included in this study.

Location	Species	Sample ID	Band/Museum ID	Latitude	Longitude
Royal City, Grant, WA	BANS	WA004	UWBM:81494	46.90097	-119.6306
George, Grant, WA	BANS	WA005	UWBM:91269	47.07759	-119.8596
Royal City, Grant, WA	BANS	WA006	UWBM:81495	46.90097	-119.6306
Royal City, Grant, WA	BANS	WA008	UWBM:81497	46.90097	-119.6306
Royal City, Grant, WA	BANS	WA009	UWBM:81498	46.90097	-119.6306
Royal City, Grant, WA	BANS	WA010	UWBM:81499	46.90097	-119.6306
Okanagan, Vaseux, BC	BANS	SBC001	UBC:TB002508	49.29990	-119.5289
Cypress Hills, AB	BANS	SEAB001	RSKM:A-1118	49.66667	-109.5000
Cypress Hills, AB	BANS	SEAB002	RSKM:A-1120	49.66667	-109.5000
Cypress Hills, AB	BANS	SEAB003	RSKM:A-9273	49.83029	-109.5566
Saskatoon, SK	BANS	SK073	2840-88692	52.146973	-106.647034
Saskatoon, SK	BANS	SK074	2840-88689	52.146973	-106.647034
Brightwater Reservoir, SK	BANS	SK035	2740-99072	51.77394	-106.5443
Brightwater Reservoir, SK	BANS	SK041	2740-99087	51.77394	-106.5443
Cultus, ON	BANS	ON028	2840-95398	42.6234	-80.601
Bolin, ON	BANS	ON034	2840-95515	42.6256	-80.7306
Bossuyt, ON	BANS	ON72	2840_95587	42.6621	-80.979
Bossuyt, ON	BANS	ON070	2840-95585	42.6621	-80.979
Bossuyt, ON	BANS	ON068	2840-95583	42.6621	-80.979
PV Pit, ON	BANS	ON061	2840-95573	42.7275	-81.0641
Sackville, NB	BANS	NB003	2980-00716	45.897979	-64.34295