2023-01-26

An Assessment of Caribou (Rangifer tarandus) Genomic Diversity and Structure in Western Canada to Guide Species Conservation and Management

Michalak, Anita

UNIVERSITY OF CALGARY

An Assessment of Caribou (*Rangifer tarandus*) Genomic Diversity and Structure in Western Canada to Guide Species Conservation and Management

by

Anita Michalak

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN VETERINARY MEDICAL SCIENCES

CALGARY, ALBERTA

JANUARY, 2023

ABSTRACT

Human-induced environmental change is one of the biggest threats to global biological diversity, and the resulting environmental conditions have made it increasingly difficult for species to adapt and survive. The use of genomic technologies, such as the inference of genetic structure, can aid species conservation and prevent population declines. Particularly for caribou (*Rangifer tarandus*), which are considered an at-risk species across Canada, determining population genetic structure can help delineate units for conservation while detecting potentially cryptic population structure and diversity as well as undetected and/or mislabeled populations. For my M.Sc. thesis, I studied genomic diversity in caribou sampled throughout western Canada to better characterize population structure and supplement previous genetic studies conducted in this region. I accomplished this using several population structure inference methods and by combining individual-based genomic and spatial data for 658 individuals derived from 41 herds across British Columbia (BC) and Alberta (AB). Results indicate that population structure inferred from genomic data reflects neither past nor present caribou classification schemes. Instead, caribou genetic differentiation in BC and AB is best reflected at *K*=4 clusters, which primarily: (1) identifies a potential new conservation unit composed of individuals belonging to Itcha-Ilgachuz and neighboring subpopulations, and (2) redefines the boundaries of existing populations. Despite the need for multiple lines of evidence to provide complementary criteria for designating distinct units for conservation or populations, my work illustrates how genomics can help inform and improve the delineation of such conservation and management units for caribou.

Keywords: *conservation genomics, population structure, genetic differentiation, clustering, evolutionarily significant units, ecotypes, endangered wildlife*

PREFACE

This thesis is original, unpublished, independent work by the author, Anita Michalak. Data collection reported in this thesis was performed under Animal Care Protocol #AC20-0110 issued by the University of Calgary, in compliance with the Canadian Council on Animal Care guidelines.

ACKNOWLEDGEMENTS

The University of Calgary is situated on the traditional territories of the people of the Treaty 7 region in Southern Alberta, which includes the Blackfoot Confederacy (Siksika, Piikuni, Kainai First Nations), the Tsuut'ina, and the Stoney Nakoda (Chiniki, Bearspaw, and Wesley First Nations). The City of Calgary is also home to the Metis Nation of Alberta, Region III. Additionally, I acknowledge the many traditional and unceded territories present throughout British Columbia, part of the region covered by the work presented in this thesis; I am honored to be able to live and learn on this land.

The journey from start to finish of this M.Sc. program could not have been any more unpredictable, and I have many friends, family, mentors, collaborators, and four-legged friends to thank for guiding me through the process. First and foremost, my utmost gratitude to my supervisors, Jocelyn Poissant and Marco Musiani, for onboarding me onto this project and advising me through it every step of the way. Initiating a project during the peak of a pandemic was likely no easy feat on either the student or supervisor part, so thank you for being supportive and flexible in my navigation of (quite literal and international) new territory. Thank you also to BC Government mentors and collaborators—Helen Schwantje and Caeley Thacker, among others—who helped conceptualize and fund this project and introduced me to the wonderful study system of caribou. Likewise, a huge thank you to my committee members, Jean Polfus and Sean Rogers, for supporting my academic progression and growth as a graduate student.

My transition to being an international student in Canada would also not have been possible without the incredible support of my fellow lab members and friends both old and new. To (past and present) members of both the Poissant and Musiani Labs—Chanwoori (Diana) Jeong, Sangwook (Micky) Ahn, Mason Stothart, Amy Robinson, Stefan Gavriliuc, Nancy Ngo, Charlotte Bourbon, Rachel Reed, Sina Mohtasebi, Ihncheol Jung, Mirjam Barrueto, Margaret Hughes, and Maria Cavedon—thank you x1000 for your support and friendship. I have never felt so at home somewhere I hadn't yet stepped foot, and I

owe that to you guys, our zoom game nights, backyard beers and bonfires, and your ability to carry a mattress outside in the middle of Canadian winter.

Some honorable mentions: Maria—thank you for being one of my biggest mentors throughout this project. You have no idea how much I look up to you and appreciate how much you've helped guide me over the past two years. Margaret—thank you for surviving endless hours of travel with me (BC, Montreal, and beyond) and somehow not getting sick of my music. Near-avalanches and yeeted windshield wipers got nothing on us. Micky—thank you for letting a stranger be your roommate and being the biggest reason I was able to adjust to life in Calgary so quickly. I'm sorry you had to wait so long for me at the airport, but I'm grateful for it and every other one of the hundreds of times you've been a selfless, amazing friend. I will forever appreciate you convincing me to become a cat parent and I'm glad I was able to return the favor. Diana—I don't even know where to start. Thank you for surviving the highs and lows of having me as a roommate for almost two years. I probably don't make it easy on you, but I can't imagine my grad experience without our little, crazy-cat-lady-inspired apartment and all its nights filled with quiet friendship time, couch sessions, impromptu sound effects and breakouts into song (apologies to our neighbors), and tiny duck and baby figurines in unexpected places. You helped keep me sane (unless we're both insane, which is also likely).

To new friends made in Calgary (Jasmeen, James, Megan, Kal, Ben, Elspeth, Kariem) and old friends who made the trek up north to visit (Jillian, Jackie, Alexa), or tried to (Annabelle)—thank you for keeping me adventurous in a place as beautiful as this. Adventures shared with you guys are unforgettable and indescribably better. Kariem—thank you for an endless supply of support, affirmation, motivation, and so many things I can't even count. You helped me through the last leg of this journey more than anyone; I know it's not a debt I can (or need to) repay in TikToks and memes, but you best bet I'll try. Jill—thank you, thank you, thank you for being one of biggest sources of support,

love, laughter, and the epitome of friendship for 10+ years and counting. Our (somewhat) regular FaceTime sessions, story exchanges, and joint adventures are far from over.

To my family (mom, dad, and Matt)—thank you for being continuously supportive of my apparent need to explore hundreds of miles away from home. I know it's not easy on you, but I wouldn't be where I am today without your love. And of course, to my sweetest boys, Lemon and Kitty—you drove me crazy in all the best ways cats do, and I wouldn't have it (or have survived it) any other way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: GENERAL INTRODUCTION

1.1 Genetics for species conservation

Human-induced environmental change is one of the biggest threats to global biological diversity, presenting itself in the form of habitat degradation, fragmentation, and a changing climate, among other factors (Jump & Penuelas 2005). The resulting environmental conditions have made it increasingly difficult for many species to respond and survive without further human intervention (Ceballos et al. 2017). Current efforts to conserve species and prevent population declines—including the creation of protected habitat areas or other legal frameworks to protect endangered species—have been varied in both type and success (Supple & Shapiro 2018). Nonetheless, these efforts can be aided through the use of genetic technologies to identify and prioritize populations for conservation and better inform management actions that impact threatened and endangered wildlife.

The field of conservation genetics is aimed at preventing species loss and preserving biodiversity at the genetic diversity level (Frankham et al. 2009). Genetic diversity refers to the variability of genes within a species and is often the barometer for evaluating population fitness and ultimately long-term persistence (Teixeira & Huber 2021). Specific genetic techniques are implemented depending on the conservation questions being asked and can be focused on either the individual, population, or species level. The relationship between levels of genetic variation and demographic factors, for instance, has often been tested using highly variable neutral markers such as microsatellites or amplified fragment length polymorphisms (AFLPs; Allendorf 2017). Despite the major insight conservation genetics has provided regarding inferences about demography, gene flow, effective population size, population structure, and other aspects of population ecology, it leaves several questions unresolved because of the limitations of using a small number of markers (Ouborg et al. 2010).

1.2 The rise of conservation genomics

Conservation genomics refers to the application of genome-wide analyses to aid the preservation of both populations and biodiversity (Allendorf et al. 2010). In comparison to genetic approaches, genomics considers complete genomes or genome-wide data (Supple & Shapiro 2018). With recent technological advances, the list of genomic techniques has grown and diversified to provide a range of options for researchers. Currently, some of the most widely used techniques include whole-genome sequencing (WGS), reduced-representation sequencing (RRS), transcriptome sequencing (TS), and microarrays (Narum et al. 2013; Davey et al. 2011; Davey et al. 2010; LaFramboise 2009).

Regardless of the technique used, most recent genomic studies rely on single nucleotide polymorphisms (SNPs)—variations in individual base pairs in either neutral or functional segments of the genome—in their analyses. SNPs are relatively easy to genotype in large numbers and cover a large part of the genome, with decidedly greater coverage than microsatellites and AFLPs, and can therefore provide a more representative view of the genetic variation within individuals and populations (Ouborg et al. 2010). As a result, inferences about demography, gene flow, population history, etc. can be deduced at higher resolutions. Array-based techniques, also known as SNP chips, are particularly efficient at genotyping markers across many individuals for a range of conservation applications and are the basis for methods used in this study (LaFramboise 2009; Carrier et al. 2022).

1.2.1 Population structure

Population genomic studies are increasingly able to address multiple research questions with information from just a single genomic data set, or without as frequent of a need for supplementary data (Hohenlohe et al. 2020). For instance, genomic data can be used to assess population structure (i.e., the presence of a difference in allele frequencies between populations) from the simultaneous perspective of both neutral and adaptive markers. Genetic differentiation at neutral loci results from

genetic drift, gene flow, and mutation, while variation at adaptive loci is also shaped by selection (Wright 1931). As a result, populations may be structured differently depending on spatial patterns of drift, gene flow, mutation, and selection and thus the type of marker (neutral vs adaptive) being used (Funk et al. 2012).

From a conservation standpoint, determining the genetic structure of populations is crucial to inferring the relative importance of various evolutionary processes across them. In particular, because gene flow infuses new genetic variation into populations, understanding the levels of gene flow certain populations, especially ones impacted by fragmented landscapes, exhibit is often of strong interest (Crooks & Sanjayan 2006; Walters & Schwartz 2020). Moreover, populations may not always be continuously distributed, as is the case for species found across discrete habitat patches, so it is important to distinguish between demographic and genetic connectivity (the former depends on the relative contributions of dispersal and survival/reproduction to population growth rates, while the latter depends primarily on the absolute number of dispersers among populations; Lowe & Allendorf 2010; Waples & Gaggioti 2006). Fortunately, genomics provides the power to delineate such populations into so-called "units" for conservation efforts while detecting potentially cryptic population structure and quantifying how genetically distinct different populations are (Hohenlohe et al. 2020).

1.2.2 Overview of population structure inference methods

Many programs implementing various models or algorithmic approaches can be applied in the context of genetic structure analyses. Programs such as *GenePop* and *Structure* were among the first software packages to gain recognition throughout the population genetics community and are still widely used (Excoffier & Heckel 2006). However, new software is being continuously developed and released, particularly to keep up with advances in genomics and the increasingly large and complex datasets being generated.

Current population genetics software can be grouped into several categories, including, but not limited to, multi-purpose packages, individual-centered programs, and specialized programs. A review by Excoffier & Heckel (2006) describes these as follows: multi-purpose packages (e.g., *Arlequin*, *GenePop*, *Genetix*, etc.) are intended to compute a variety of basic statistics that describe genetic diversity within and among populations, in addition to more elaborate or specialized analyses (Excoffier et al. 2005; Raymond & Rousset 1995). These analyses typically assume putative subpopulations and examine how they relate genetically through the implementation of statistical analyses such as *F*-statistics or analyses of molecular variance (AMOVA; Excoffier et al. 1992). The goal of more specialized programs such as *Batwing*, *Colonise*, or *Migrate*, on the other hand, is to infer some population parameters under specific evolutionary scenarios (Excoffier & Heckel 2006). Finally, individual-centered programs—several of which were used in this study—are those that target the analysis of individuals and the more recent history of a given population.

Individual-centered software packages aim to detect recent immigrants from all other samples, since these immigrants are expected to present different multilocus genotypes than expected for native individuals in a population (Excoffier & Heckel 2006; Beaumont & Rannala 2004). Some of these programs sort individuals into *a priori* (predefined) populations, while others deduce theoretical groupings for which allelic frequencies are iteratively estimated (Excoffier & Heckel 2006). Until the development of more novel programs, most individual-centered packages assumed that loci were unlinked and that populations were in Hardy-Weinberg Equilibrium (HWE), which is not always the case for real systems.

Methods for allocating individuals to groups or clusters based on genetic data can be further divided into categories: model-based and distance-based (Pritchard et al. 2000; Alexander et al. 2009; Wollstein & Lao 2015). Model-based approaches are the more traditional of the two, and are typically reliant on a set of assumptions such as migration-drift or HWE (Pritchard et al. 2000). If these

assumptions are not met, resulting assessments of population structure may not be valid or accurately represent the genetic structure of a population or group of populations. Model-based approaches have been used since the early 2000s and implement Bayesian clustering and maximum likelihood techniques to evaluate the likelihood of observed data, assuming that data is randomly drawn from a predefined model of the population. Conversely, distance-based methods are typically model-free and do not require prior assumptions; instead, they aim to identify clusters through the analysis of matrices describing genetic distances or genetic similarities between individuals (Corander et al. 2003). This is often done with visualization using multidimensional scaling (MDS) methods such as principal component analyses (PCAs).

The relative advantages and disadvantages of model-based versus distance-based approaches have been previously debated (Pritchard et al. 2000; Alexander et al. 2009; Corander et al. 2003; Elhaik 2022). While distance-based methods are less constrained when it comes to assumptions, they can be limited by their dependence on the distance measure used, and it can sometimes be difficult to assess the significance of clustering results as well as incorporate additional information such as geographic sample location. Various program developments have sought to overcome these constraints (Jombart 2008; Yang et al. 2012); thus, the assortment of software currently available for population structure analyses is reasonably robust.

1.2.2.1 Model-based methods

The program *Structure* developed by Pritchard et al. (2000) is one of the most widely used frameworks for inferring population structure. The program looks at the differences in distribution of genetic variants among populations by detecting allele frequency differences and assigning individuals to clusters, typically referred to by the parameter *K*, based on an analysis of likelihoods (Porras-Hurtado et al. 2013). It has several features, including inferring the presence of distinct populations, assigning

individuals to populations, studying hybrid zones, identifying migrants and admixed individuals, and estimating population allele frequencies in cases of migrating or admixed individuals (Pritchard et al. 2000; Porras-Hurtado et al. 2013).

Structure employs four primary types of models: (1) the *no admixture* model, which assumes that individuals come from distinct populations, (2) the *admixture* model, which assumes that individuals may have mixed ancestry, (3) the *linkage* model, which accounts for admixture linkage disequilibrium (LD), or in other words, the phenomenon by which recently admixed populations have larger regions of LD among loci, and (4) the prior population information models (*LOCPRIOR*), which can use location or phenotype information to enhance the detection of population structure (Pritchard et al. 2000; Porras-Hurtado et al. 2013). The program implements a Bayesian clustering approach involving Markov chain Monte Carlo (MCMC) estimation, which is a process that includes: (1) randomly assigning individuals to a pre-determined number of clusters, (2) estimating genetic variant frequencies in each cluster, and (3) re-assigning individuals based on those frequency estimates (Pritchard et al. 2000; van Ravenzwaaij et al. 2018). This sequence is then iterated thousands of times to better determine reliable allele frequency estimates in each population, as well as the probabilities that individuals belong to their assigned cluster.

Structure can be applied to a variety of commonly used genetic markers, including SNPs, microsatellites, restriction fragment length polymorphisms (RFLPs), and AFLPs (Porras-Hurtado et al. 2013). For large genotype datasets, however, the need for a large number of posterior samples to combat convergence and mixing issues—a factor of the program's Bayesian model—can be a limiter (Raj et al. 2014). To address the computational challenges that arise when running very large SNP datasets, Raj et al. (2014) created the program *fastStructure*, which is modelled after *Structure*. *fastStructure* is unique in that it takes inspiration from both *Structure* and *Admixture*; it estimates ancestry proportions that are comparable to those estimated by *Admixture*, all while using the same variational Bayesian

framework for posterior inference as seen in *Structure* (Porras-Hurtado et al. 2013; Raj et al. 2014). Its runtimes are about two orders of magnitude faster than *Structure*'s, which makes *fastStructure* considerably more efficient when analyzing large SNP datasets (Porras-Hurtado et al. 2013). While one caveat of the program is that it does not factor in LD among genetic markers, developers claim that the bias due to unmodelled LD does not significantly impact the accuracy of cluster and ancestry estimates (Raj et al. 2014).

Some programs, such as *Admixture*, rely on a maximum likelihood (ML) approach as opposed to sampling posterior distributions using MCMC, as in *Structure* (Porras-Hurtado et al. 2013; Alexander & Lange 2011). Developed by Alexander & Lange (2011), *Admixture* uses ML estimations for individual ancestries from multilocus SNP genotypes; its statistical model is the same as that found in *Structure*, but it calculates population parameters much more rapidly using a fast numerical optimization algorithm. It utilizes a cross-validation approach to help estimate *K*, which is in contrast to *Structure*'s use of model evidence computations for each value of *K*. In essence, this allows *Admixture* to incorporate much larger marker sets in its analyses (Alexander & Lange 2011). Nonetheless, just as in *fastStructure*, models implemented in *Admixture* do not explicitly account for LD among markers; later versions of *Structure*, meanwhile, have been enhanced to consider the occurrence of LD caused by admixture among populations, though LD caused by genetic drift has not been as well supported (Kaeuffer et al. 2007). Because model-based approaches operate under the assumptions of HWE as well as the absence of LD, both departures from HWE and the presence of strong LD could lead to overestimates of the number of clusters detected (Raj et al. 2014; Alexander & Lange 2011).

Lastly, and unlike most of the programs discussed above, *Tess* is a program that investigates population structure while simultaneously incorporating genetic and geographic data (Caye et al. 2016). Much like *Structure* and its cousin *fastStructure*, *Tess* addresses spatial population genetics with a Bayesian clustering algorithm that is based on a hierarchical mixture model, where the prior distribution

on cluster labels is defined as a Hidden Markov Random Field (Francois et al. 2006). The user supplies individual geographic locations and genotypes, and *Tess* deduces population structure without assuming *a priori* populations. It is most commonly used to detect genetic barriers or discontinuities in otherwise apparently continuous populations.

1.2.2.2 Distance-based methods

In contrast to the model-based methods described above, multivariate methods for analyzing genetic markers, including both discriminant analyses of principal components (DAPCs) and standard PCAs, are useful for summarizing genetic variability without making strong assumptions about an evolutionary model; in other words, they do not rely on HWE or assume the absence of LD (Jombart 2008). This is particularly valuable when information known about a study system is limited, as is often the case in landscape genetics studies (Manel et al. 2003). DAPCs seek synthetic variables, known as discriminant functions, to show differences among groups as best as possible while minimizing variation within clusters (Jombart & Collins 2015). The discriminant functions are constructed as linear combinations of the original variables (alleles) that have the largest between-group variance and the smallest within-group variance. The DAPC then provides membership probabilities of each individual for different clusters based on the retained discriminant functions. This process is different from programs that utilize admixture coefficients, such as *Structure* and *Admixture*, but membership probabilities can still be interpreted as proximities of individuals to various clusters (Jombart & Ahmed 2011; Jombart & Collins 2015). Distance-based methods can be implemented using various software, such as *adegenet*, an R package specifically designed to handle and analyze large genome-wide SNP data (Jombart 2008; Jombart et al. 2010; Jombart & Ahmed 2011).

1.3 Biology and population genetics of caribou (*Rangifer tarandus*)

1.3.1 Global caribou distribution and ecology

The caribou (*Rangifer tarandus*) is one of North America's most iconic large mammals, and is of great ecological, economic, and cultural significance (Hebblewhite et al. 2010). The species currently has a circumpolar distribution across boreal, montane, and arctic environments. In Europe and Siberia, the species is more commonly known as reindeer. Within North America, caribou are presently found in Alaska and in all Canadian Territories and Provinces except New Brunswick, Nova Scotia, and Prince Edward Island (Bergerud 1996).

Caribou are medium-sized members of the deer family, with males typically weighing 120-200 kg and females 80-140 kg. The average maximum lifespan of males and females is 10 and 15 years, respectively. Caribou generally have a lower reproductive rate than other North American cervid species due to the fact that females give birth to only a single calf annually. Primiparity typically occurs at 3 years of age, though with good habitat conditions, females can also calf at the age of 2. Caribou generation length is estimated to range between 6-7 years (Bergerud 2000).

With their dense pelage, large fat stores, regulated metabolism, and a counter-current heat exchange system, caribou are well-adapted to cold environments (Bergerud 1996; Bergerud et al. 2008). During the winter, their diet consists primarily of arboreal and ground lichens, while in the spring, summer, and autumn, they forage mostly on vascular plants, though lichen remains an important dietary component. In the Arctic, caribou play a large ecological role as the primary food source of both arctic predators and human populations at some points of the year, contributing to the nutrient transfer that is crucial to ecosystem function and human subsistence (Brathen et al. 2007).

1.3.2 North American caribou classification

In 1961, caribou in North America were classified into four subspecies: (1) barren-ground caribou (including *R. t. groenlandicus* and *R. t. granti*), (2) woodland caribou (*R. t. caribou*), (3) Peary caribou (*R. t. pearyi*), and (4) Dawson's caribou (*R. t. dawsoni*; now extinct, but once occurring on Haida Gwaii in British Columbia; Banfield 1961). Caribou have since been further categorized by ecotype—or distinct geographic varieties that are adapted to specific environmental conditions—depending on caribou behavioral factors including habitat use, diet, and migratory pattern (Festa-Bianchet et al. 2011). The four ecotypes generally recognized across Canada are: (1) barren-ground, (2) boreal, (3) mountain, and (4) Peary. In Alaska, a migratory tundra ecotype is also acknowledged, but used interchangeably with the Canadian barren-ground ecotype. However, caribou ecotype designations are not always consistent, as there is some disagreement regarding how many ecotypes there are and where their boundaries lie.

From a phylogeographic perspective, caribou lineages found throughout North America have been shown to fall into one of two major haplogroups, or clades, that evolved around 120 kya: those originating from the Beringian-Eurasian lineage (BEL; which includes much of the barren-ground subspecies and ecotype), and those of the North American lineage (NAL; the woodland caribou subspecies that encompasses both boreal and mountain ecotypes—though see caveats below; Cronin et al. 2005; McDevitt et al. 2009; Yannic et al. 2013; Polfus et al. 2017). Caribou phylogenetic structure appears to be a direct result of postglacial expansions after the last glacial maximum (LGM), a fact that is supported by both the fossil record and ecological considerations (Klütsch et al. 2012; Yannic et al. 2013). After the LGM, these two lineages recolonized the ice-free landscape and underwent admixture upon secondary contact, a phenomenon that is referred to as a "hybrid swarm" in some instances (Klütsch et al. 2016; McDevitt et al. 2009).

1.3.3 Caribou conservation status in Canada

Endangered and threatened species monitoring in Canada occurs at federal, territorial, and provincial levels under the Canadian Species at Risk Act (SARA). Conservation and management responsibilities of caribou in Canada are thus shared among government bodies and Indigenous partners. SARA also recognizes that units below the species level may require conservation and tasked the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) with their definition and assessment (COSEWIC 2010). In 2009, COSEWIC established and formalized the concept of Designatable Units, which recognizes that there are spatially, ecologically, or genetically discrete and evolutionarily significant units that are irreplaceable components of intraspecific biodiversity (COSEWIC 2011). DUs are broadly the equivalent of evolutionarily significant units (ESUs; Ryder 1986). In the case of caribou in Canada, DU designation combines available data on species taxonomy, phylogenetics, morphology, life history, behavior, genetics, and biogeography (COSEWIC 2014a; COSEWIC 2014b). As a result, in 2011, eight "Nationally Significant Populations"—the predecessors to DUs that emerged from COSEWIC caribou assessments in the early 2000s—were reassessed and reassigned to twelve DUs across Canada (**Figure 1.1**; COSEWIC 2011).

1.3.4 Caribou in western Canada

Over recent decades, woodland caribou herds across Canada have experienced significant declines, owing primarily to the loss of critical habitat, habitat fragmentation, and increased predation by wolves. This has resulted in many populations being listed as Endangered, Threatened, or Special Concern under Schedule 1 of SARA, including numerous populations in British Columbia (BC) and Alberta (AB; Festa-Bianchet et al. 2011). Until recently, caribou were primarily listed under Schedule 1 of SARA, which initially grouped the woodland caribou subspecies into three populations: Woodland caribou (Boreal population; listed as Threatened in 2003), Woodland caribou (Northern Mountain population;

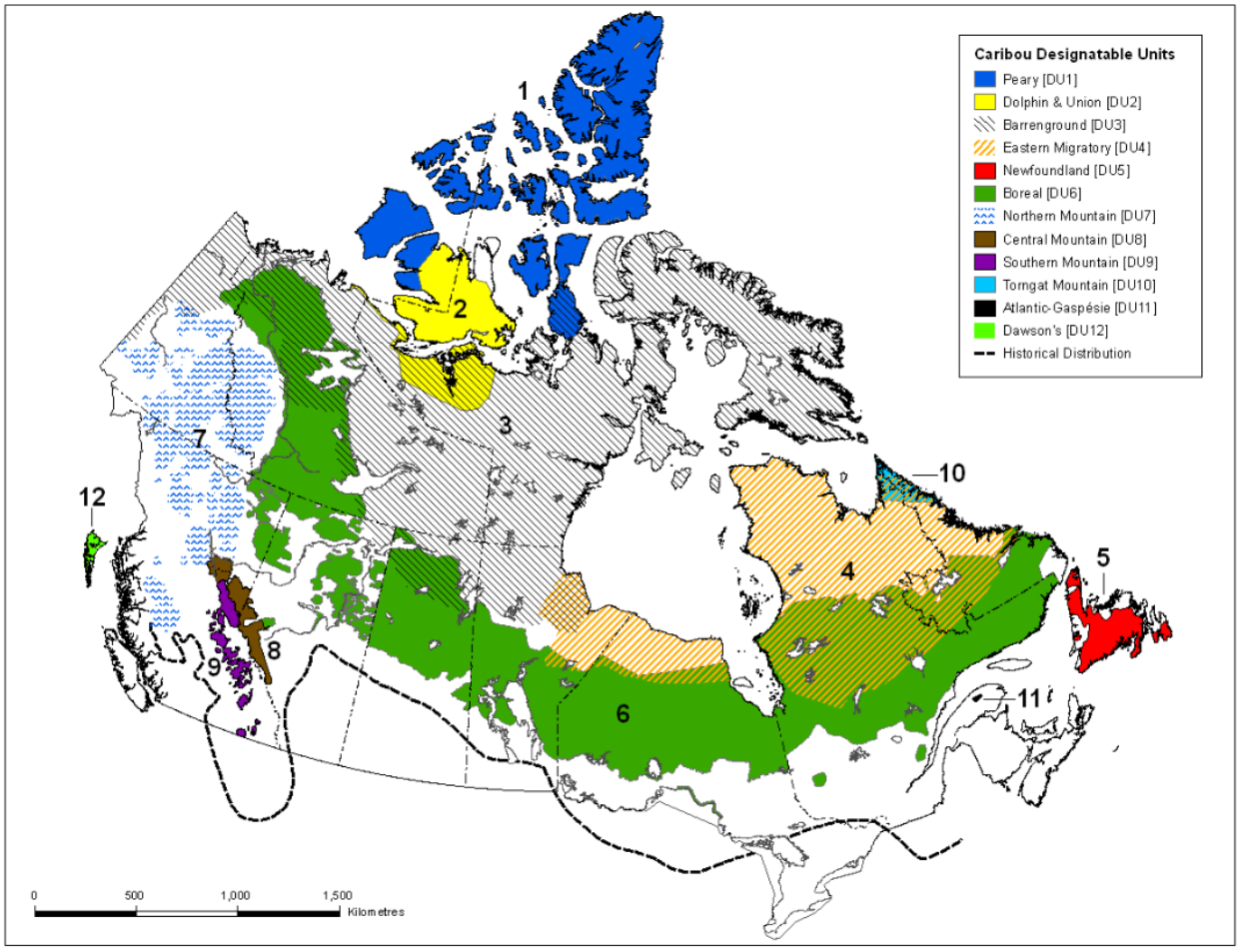

Figure 1.1. Distribution of Designatable Units (DUs) for caribou across Canada. Reprinted from "Designatable Units for Caribou (*Rangifer tarandus*) in Canada" by COSEWIC, 2011, Ottawa, 88 pp.

listed as Special Concern in 2005), and Woodland caribou (Southern Mountain population; listed as Threatened in 2003; SARA 2003a; SARA 2003b; SARA 2005). The Southern Mountain population has been further divided into three "groups" (northern group, central group, and southern group) that occur along the Rocky Mountains in BC and AB and are also managed by the Parks Canada Agency within National Park boundaries.

Since the creation of COSEWIC DUs in 2011, caribou "populations" have experienced shifts in both definition and boundaries. Currently, four of the twelve recognized caribou DUs in Canada are

found in BC and AB: Boreal (DU6), Northern Mountain (DU7), Central Mountain (DU8), and Southern Mountain (DU9; COSEWIC 2011). This reassignment and shift away from original SARA Schedule 1 populations had several implications, including: (1) separating the processes of DU/population assignment and status assessment, (2) reassigning nine caribou herds from the Schedule 1 Southern Mountain population (listed as Threatened) to Northern Mountain DU7, and (3) splitting the remainder of the herds in the Schedule 1 Southern Mountain population into the Central Mountain DU8 and Southern Mountain DU9. The individual statuses of these DUs were then assessed by COSEWIC in 2014, listing DU6 (Boreal) as Threatened, DU7 (Northern Mountain) as Special Concern, and both DU8 (Central Mountain) and DU9 (Southern Mountain) as Endangered (COSEWIC 2014a; COSEWIC 2014b).

COSEWIC's most recent assessment of DU7 as Special Concern, in particular, has caused debate, since the nine new herds belonging to the population—that in previous assessments belonged to the Schedule 1 Southern Mountain population—were effectively downgraded from Threatened to Special Concern status under COSEWIC (COSEWIC 2014b; Weckworth et al. 2018). While this assessment was referred back to COSEWIC for revision and not formally accepted under Schedule 1 of SARA, it raised initial concern for populations that have not yet been evaluated through a genetic lens in tandem with ecological and behavioral considerations. According to the most recent population estimates provided by the province of BC, seven of these nine herds are in decline (Charlotte Alplands, Graham, Itcha-Ilgachuz, Rainbows, Takla, Tweedsmuir, and Wolverine), one is stable (Chase), and only one is increasing (Telkwa; Government of British Columbia 2021).

1.3.5 Current ecological and genetic knowledge

Unlike the breakdown of caribou ecotypes across Canada that encompasses four groups (barren-ground, boreal, mountain, and Peary caribou), woodland caribou ecotypes specific to BC and AB exhibit slightly greater division, though this classification system is debatably outdated (Heard & Vagt

1998). Still, past studies have acknowledged that the Canada-wide "mountain" ecotype can be split into "northern" and "mountain" ecotypes in BC, so some authors have used this categorization to compare ecological and genetic knowledge among these caribou groups (Serrouya et al. 2012; Heard & Vagt 1998; COSEWIC 2011). The northern ecotype, which is sometimes referred to as the shallow-snow ecotype, corresponds to caribou that spend most of the year either on mountain ridges or in adjacent dry forests and dig through shallow snow for terrestrial lichens. Alternatively, the mountain ecotype caribou, also known as the deep-snow ecotype, remain high on mountains during the winter, are able to walk atop deep snowpacks (ranging 2 to 5 meters), and forage on arboreal lichens found in the canopy of conifer trees (Seip & McLellan 2008). This ecotype's exclusive reliance on arboreal lichen has justified its designation as a distinct ecotype, but the genetic basis for this classification is limited. It is also worth noting that COSEWIC has no requirements in terms of genetic differentiation when it comes to defining ecotypes, and that some biologists have also used the term ecotype to further divide caribou in BC into behavioral groups focused exclusively on behavioral differences (Edwards et al. 1960; Serrouya et al. 2012; Yannic et al. 2016). The third and final ecotype said to be present in BC is the boreal ecotype, which can be found in boreal forests in the northeastern corner of the province and digs through shallow snow to access terrestrial lichens during the winter, similarly to the northern ecotype (Seip & McLellan 2008).

The overlap of ecotype, DU, and SARA-listed population designations can be found in **Table 1.1**, but can be summarized as follows: (1) the boreal caribou ecotype, DU6, and SARA-listed population are synonymous and cover the same range, (2) the northern caribou ecotype encompasses DUs 7 (Northern Mountain) and 8 (Central Mountain) as well as the SARA-listed Northern Mountain, Southern Mountain – Northern Group, and Southern Mountain – Central Group populations, and (3) the mountain caribou ecotype covers DU9 (Southern Mountain) and the synonymous SARA-listed Southern Mountain – Southern Group population and is also the most endangered.

Table 1.1. Caribou classifications in BC and AB according to ecotypes, SARA Schedule 1 populations, and COSEWIC's Designatable Units (DUs). COSEWIC status assessment designations (where applicable) and their year of assessment are indicated in parentheses and italics. Adapted from "Caribou classifications," Government of British Columbia 2022.

<table>
<tr><th>Ecotype</th><th colspan="2">SARA Schedule 1 Population
(Status, year listed)</th><th>Designatable Unit (DU)
(Status, year assessed)</th></tr>
<tr><td>Boreal</td><td colspan="2">Woodland Caribou, Boreal population
(Threatened, 2003)</td><td>Caribou (Boreal population) DU6
(Threatened, 2002 and 2014)</td></tr>
<tr><td rowspan="3">Northern</td><td colspan="2">Woodland Caribou, Northern Mountain population
(Special Concern, 2005)</td><td rowspan="2">Caribou (Northern Mountain population) DU7
(Special Concern, 2014)</td></tr>
<tr><td rowspan="3">Woodland Caribou, Southern Mountain population
(Threatened, 2003)</td><td>Northern Group</td></tr>
<tr><td>Central Group</td><td>Caribou (Central Mountain population) DU8
(Endangered, 2014)</td></tr>
<tr><td>Mountain</td><td>Southern Group</td><td>Caribou (Southern Mountain population) DU9
(Endangered, 2014)</td></tr>
</table>

From a genetic standpoint, the justification of currently recognized population subdivisions can be attributed to both phylogeographic history and caribou population genetics studies conducted to date. The Central Mountain DU8, for instance, has been said to be dictated chiefly by a unique gene pool that was the product of the admixture of BEL and NAL caribou reported by McDevitt et al. (2009) in their analysis of mitochondrial DNA (mtDNA) in caribou in western Canada. Based on further mitochondrial haplotype analyses, however, Taylor et al. (2021) supposed that Central Mountain caribou are more likely of purely NAL origin. The same study also suggested that Southern Mountain DU9 caribou and some Northern Mountain DU7 subpopulations are of BEL origin, stating that caribou in the region have likely been impacted by differential colonization along with associated introgression events and do not represent one homogenous group with identical ancestry (Taylor et al. 2021).

Serrouya et al. (2012) carried out an extensive assessment of microsatellite genotypes of several caribou subpopulations found throughout BC and AB, with the majority of them concentrated in the eastern part of BC and representing the mountain ecotype. Their population differentiation results were consistent with the current breakdown of DUs, indicating four distinct groupings for the Boreal, Northern Mountain, Central Mountain, and Southern Mountain populations (Serrouya et al. 2012). Zittlau (2004) and Weckworth et al. (2012) also conducted studies with a focus on subpopulations belonging to the Southern Mountain DU9, but these studies used different sets of genetic markers, were inconsistent in their assignment of subpopulations to populations, and each assessed unique but incomplete combinations of subpopulations, making it difficult to compare results.

Some of the most recent work on western Canada caribou population structure was conducted by Cavedon et al. (2021, 2022b) in studies looking at the genetic basis of caribou migratory behavior and forces contributing to genetic differentiation, such as sex-biased dispersal. The earlier study found a two-cluster separation comprised of a "Northern" cluster, which included the barren-ground caribou subspecies and most of the woodland caribou subspecies found in Yukon, and a "Southern" cluster including woodland caribou belonging to the Boreal DU6 and Central Mountain DU8 (Cavedon et al. 2021). Cavedon et al. (2022b) confirmed this two-cluster genetic breakdown, with support for a transitional zone located in northern BC and vaguely corresponding to the Northern Mountain DU7. Nonetheless, this transitional zone was notable in that it did not overlap with the "hybrid swarm" previously found between the BEL and NAL lineages/clusters, indicating a potentially ambiguous and/or unresolved delineation of boundaries (McDevitt et al. 2009; Weckworth et al. 2012; Yannic et al. 2013). Additionally, very few to no samples from the Northern or Southern Mountain DUs were included in these studies, which prevents locating the exact position of population boundaries or transition zones.

1.3.6 Knowledge gaps and management implications

COSEWIC caribou reports from 2011 and 2014 stressed the uncertainties associated with the boundaries of certain DUs, specifying that the assignment of some caribou herds was uncertain due to the lack of comparative analyses and overall poor understanding of the ecology and evolutionary origin of mountain populations (COSEWIC 2011; COSEWIC 2014a; COSEWIC 2014b). In 2022, the 2014 COSEWIC assessment for the mountain caribou populations was referred back to COSEWIC for further information and consideration on the basis of these delineation uncertainties, which were also associated with insufficient and/or inadequate genetic information, recently documented changes in census size of some affected herds, as well as a lack of inclusion of Indigenous knowledge (SARA 2022).

COSEWIC assessments state that a more comprehensive analysis across all populations in BC is required, specifically for those that may have missing or incomplete data (e.g., the Northern Mountain DU7; COSEWIC 2014a; COSEWIC 2014b). Additionally, the use of a larger number of genetic markers and more rigorous structure analyses could uncover the possible presence of cryptic population structure, as well as undetected and/or mislabeled populations (Weckworth et al. 2018).

1.4 Research objectives

For my M.Sc. thesis, I aimed to address the knowledge gaps noted by COSEWIC by conducting a comprehensive study of caribou genomic diversity in western Canada. This project addressed the primary objective of combining individual-based genomic and spatial data to characterize caribou population structure in BC and AB, with the aim of guiding future conservation and management efforts. Specifically, I tested whether the population structure of caribou in western Canada generally reflects the delineation of COSEWIC Designatable Units, SARA Schedule 1 populations, and caribou ecotypes. I incorporated both distance-based (PCAs and DAPCs) and model-based (*Structure*, *Admixture*, *fastStructure*, and *Tess*) population structure inference methods and compared their respective results

to obtain a more comprehensive and robust picture of genetic structure in BC and AB caribou. Finally, I determined whether or not these new results based on genomic data, in combination with extensive sampling, warrant changes to currently recognized units or their boundaries.

CHAPTER 2: RECONSIDERING CARIBOU CONSERVATION UNITS IN WESTERN CANADA USING GENOMIC DATA

2.1. Introduction

Human-induced environmental change is one of the biggest threats to global biological diversity, presenting itself in the form of habitat degradation, fragmentation, and a changing climate, among other factors (Jump & Penuelas 2005). The resulting environmental conditions have made it increasingly difficult for many species to respond and survive without further human intervention (Ceballos et al. 2017). Current efforts to conserve species and prevent population declines—such as the translocation of individuals or the creation of protected habitat areas—have been varied in success, but can be aided through the use of genomic technologies to identify and prioritize areas for conservation and better inform management actions that impact threatened and endangered wildlife (Supple & Shapiro 2018). From a conservation standpoint, determining the genetic structure of a population or group of populations is crucial to inferring the relative importance of evolutionary processes (i.e., gene flow, selection, and drift) across them; fortunately, genomics provides the power to delineate so-called "units" for conservation efforts while detecting both potentially cryptic population structure and diversity as well as undetected and/or mislabeled populations (Weckworth et al. 2018).

Conservation efforts can focus on a variety of biological and spatial scales, such as whole ecosystems, specific geographic areas, or individual species. In recent years, the importance of maintaining adaptive variation within species has also become increasingly recognized. In particular, in both Canada and the US, groups below the species level are now often managed in terms of evolutionarily significant units (ESUs)—populations that have substantial reproductive isolation and adaptive differences, such that they represent a significant evolutionary component of the species (Ryder 1986). However, identifying such units can be a challenge (Crandall et al. 2000; Fraser & Bernatchez 2001).

Caribou (*Rangifer tarandus*) have sustained human populations in North America for thousands of years and, since colonization, have been considered one of Canada's most iconic large mammals as well as a species that has held great ecological, economic, and cultural significance, much due to its importance to Indigenous subsistence and identity (Hebblewhite et al. 2010). Despite this, the species has also been adversely affected by anthropogenic changes both globally and across North America. In Canada, caribou populations have experienced considerable shifts in population sizes, home ranges, and connectivity, primarily owing to the loss of critical habitat, habitat fragmentation, and increased predation by wolves (Festa-Bianchet et al. 2011). As a result, caribou are listed and monitored under the Canadian Species at Risk Act (SARA), which operates at federal, territorial, and provincial levels and with both Indigenous partners and independent advisory panels to best classify, assess, and manage caribou throughout Canada.

SARA recognizes that units below the species level (i.e., subspecies, populations, subpopulations, etc.) require conservation and consequently tasked the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) with their assessment (COSEWIC 2010). In 2009, COSEWIC established and formalized the concept of Designatable Units (DUs, which are considered to be the equivalent of ESUs), recognizing spatially, ecologically, or genetically discrete and evolutionarily significant units as irreplaceable components of intraspecific biodiversity (COSEWIC 2011). In 2011, caribou across Canada were assigned to twelve DUs. However, at a local scale, caribou are also referred to in the context of subpopulations (often called "herds") and ecotypes—i.e., distinct geographic varieties that are adapted to specific environmental conditions (Festa-Bianchet et al. 2011). For example, caribou in British Columbia (BC; a part of the focal region of this study) are currently discerned according to 54 subpopulations/herds and three distinct ecotypes (boreal, northern, and mountain; **Figure 2.1**; Heard & Vagt 1998).

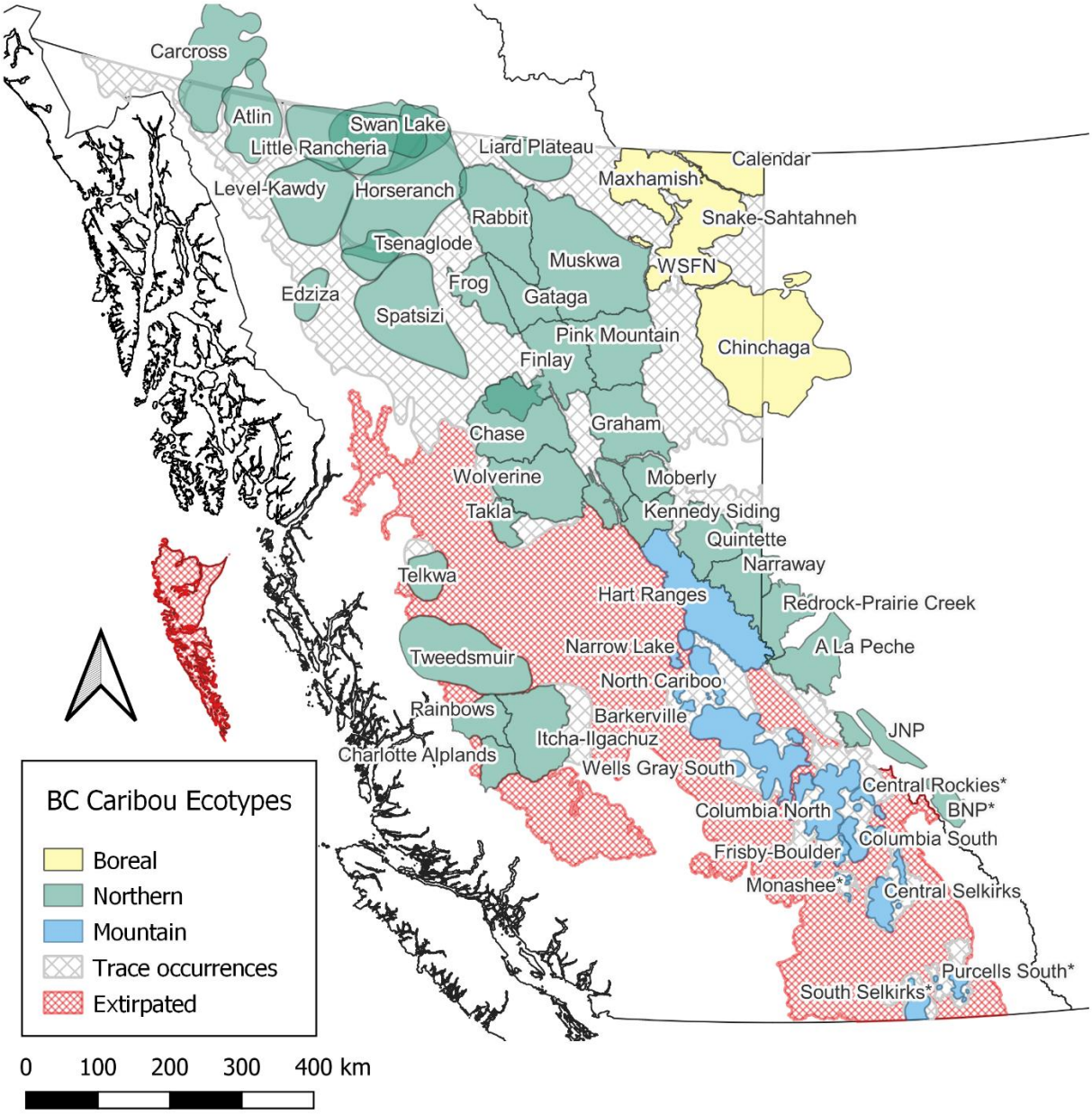

Figure 2.1. Caribou subpopulations (*N*=54) labeled and depicted by polygons within the three caribou ecotypes found in British Columbia (BC): boreal, northern, and mountain. Trace caribou occurrences are marked in grey cross-hatching, regions now extirpated but once inhabited by caribou are marked in red cross-hatching, and recently extirpated herds are indicated with an asterisk (*). Abbreviations: WSFN = Westside Fort Nelson, JNP = Jasper National Park, BNP = Banff National Park. Map created in QGIS.

Caribou in western Canada (BC and Alberta [AB]) all belong to the woodland caribou subspecies (*R. t. caribou*; one of three extant caribou subspecies currently recognized in Canada), which is listed under Schedule 1 of SARA. Until COSEWIC's assessment of DUs in 2011, woodland caribou in BC were initially grouped into three populations, which are still recognized today: Woodland caribou (Boreal population; listed as Threatened in 2003), Woodland caribou (Northern Mountain population; listed as Special Concern in 2005), and Woodland caribou (Southern Mountain population; listed as Threatened in 2003; SARA 2003a; SARA 2003b; SARA 2005). The Southern Mountain population is further divided into three "groups" (northern group, central group, and southern group) within the Recovery Strategy to facilitate communication (**Figure 2.2a**). Currently, however, most management of caribou in BC and AB happens at the above-mentioned DU-level, though some jurisdictions still rely primarily on either herd or SARA Schedule 1 population management. Of the twelve currently recognized DUs in Canada, four are found in BC: Boreal (DU6; Threatened), Northern Mountain (DU7; Special Concern), Central Mountain (DU8; Endangered), and Southern Mountain (DU9; Endangered; **Figure 2.2b**; COSEWIC 2011).

The COSEWIC DU report from 2011 and the 2014 assessment stressed the uncertainties associated with the defined boundaries of certain DUs, specifying that the assignment of some herds was "uncertain due to the lack of comparative analyses and overall poor understanding of the ecology and evolutionary origin of mountain populations" (COSEWIC 2011; COSEWIC 2014a; COSEWIC 2014b). In 2022, the 2014 COSEWIC assessment for the mountain caribou DUs was referred back to COSEWIC for further consideration as a result of these delineation uncertainties, which were also associated with insufficient and/or inadequate genetic information, recently documented changes in census size of some affected herds, as well as a lack of inclusion of Indigenous knowledge (SARA 2022). A more comprehensive analysis across all populations, especially for ones that may have missing or incomplete data, is thus required to assist with future COSEWIC assessments (COSEWIC 2014a; COSEWIC 2014b).

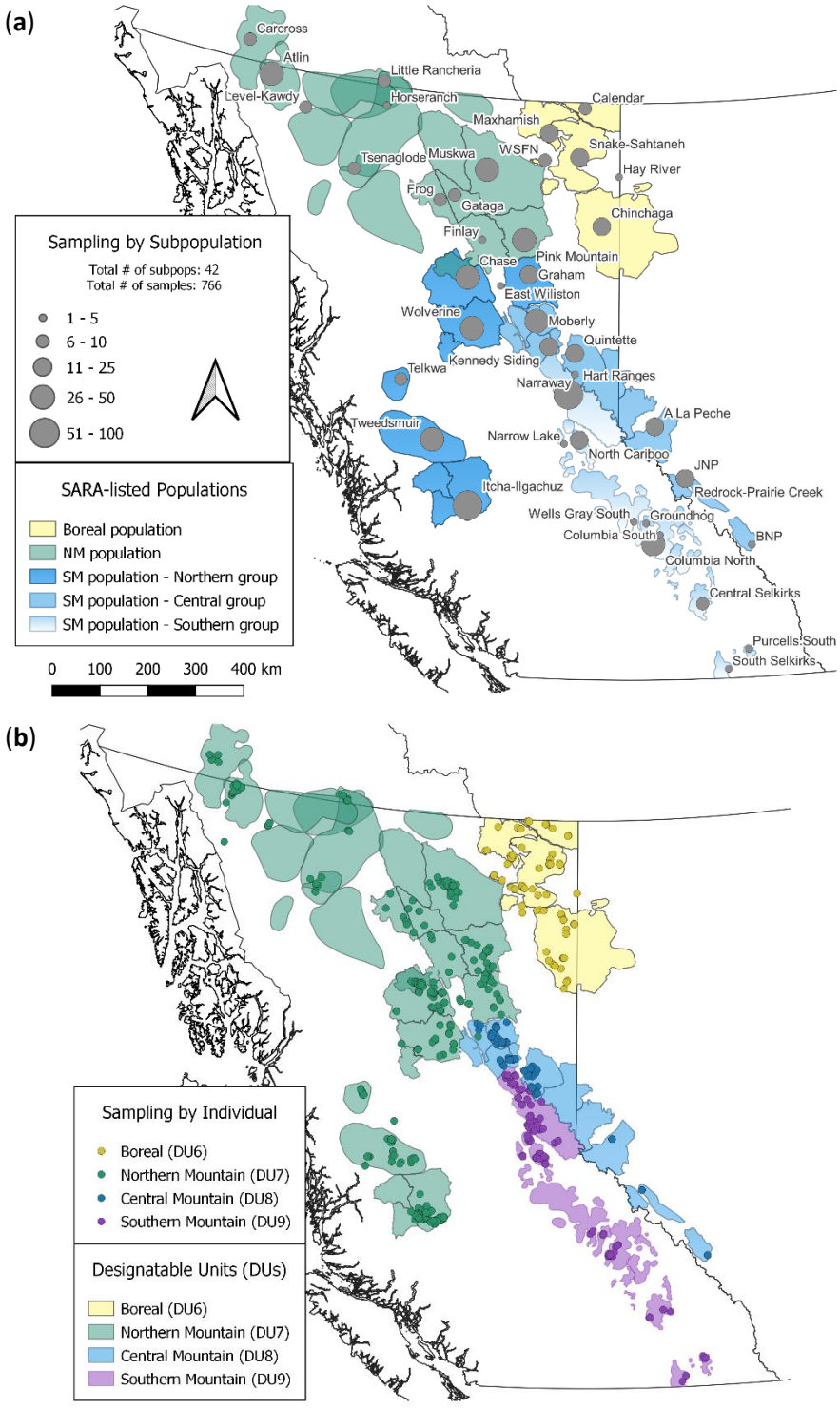

Figure 2.2. (**a**) Distribution of caribou samples by subpopulation (*N*=42) sampled throughout BC and AB. Greater sampling intensity in a given region is represented by larger circles. Color categories indicate herds' *a priori* assignment to one of the three SARA-listed populations. Abbreviations: WSFN = Westside Fort Nelson, JNP = Jasper National Park, BNP = Banff National Park. (**b**) Distribution of sampled caribou individuals (*N*=766). Color categories indicate individuals' *a priori* assignment to one of the Designatable Units (DUs). Maps created in QGIS.

From a genetic standpoint, the justification of current caribou population subdivision can be attributed to both phylogeographic history and population genetics studies conducted to date. The Central Mountain DU8, for instance, has been shown to be chiefly dictated by a unique gene pool that was the product of a "hybrid swarm" reported by McDevitt et al. (2009) in their analysis of mitochondrial DNA (mtDNA) in western Canadian caribou (Cronin et al. 2005; Klütsch et al. 2016; Yannic et al. 2013; Polfus et al. 2017), though this has also been contested with recent genomic mtDNA data (Cavedon et al. 2022b; Taylor et al. 2021). Serrouya et al. (2012) carried out one of the most extensive assessments of microsatellite genotypes in the region, considering several subpopulations found throughout BC (with the majority of them concentrated in the Southeast of the province and representing the mountain ecotype). Their population differentiation results mirrored the current breakdown of DUs throughout BC, though the authors did note potentially significant sampling gaps for Northern Mountain DU7 caribou (Serrouya et al. 2012). Zittlau (2004) and Weckworth et al. (2012) also conducted studies with a focus on subpopulations belonging to the Southern Mountain DU9, but these studies used different sets of genetic markers, were inconsistent in their assignment of subpopulations to populations, and each assessed unique but incomplete combinations of subpopulations, making it difficult to compare results.

In this study, I conducted a comprehensive evaluation of genomic diversity in western Canadian caribou by combining individual-based genomic and spatial data to characterize caribou population structure in the region. Specifically, I sought to answer the question: does population structure inferred from genomic data reflect any past or present caribou classification schemes, including but not limited to COSEWIC DUs, SARA Schedule 1 populations, and caribou ecotypes? I incorporated and compared several population structure inference methods to obtain a more complete picture of patterns in caribou genetic structure.

2.2. Materials and methods

2.2.1. *Sample collection & DNA extraction*

Caribou genomic data was obtained from samples collected by BC and AB wildlife officials and other government and parks agency partners between 2012-2021. DNA was extracted from blood (buffy coat or serum) and tissue (ear clip biopsy or muscle) following both QIAGEN DNeasy Blood & Tissue Kit and QIAamp 96 DNA QIAcube KT Kit protocols for single spin columns and 96-well plates, respectively. Extracted DNA was subsequently eluted in 400µL of molecular grade water and quantified using either the BioTek Synergy LX Multimode Reader or the Thermo Fisher Qubit 4 Fluoremeter following Thermo Fisher Quant-iT and Qubit dsDNA Assay Kits, both high sensitivity (HS) and broad range (BR). Samples containing ≥400ng of DNA were chosen for analysis, resulting in 766 individual samples representing 42 subpopulations (herds) and all four DUs present in BC (**Figure 2.2**; **Table 2.1**). The integrity of DNA for these samples was not assessed, as samples with fragmented DNA can sometimes still yield good quality genotypes using Illumina SNP arrays. However, poorly performing samples can be identified and discarded following genotyping based on abnormally low call rates (Infinium Genotyping Data Analysis 2014).

The resulting 766 samples were then normalized to a quantity of 400ng per sample, dried on the Thermo Scientific Savant SpeedVac DNA 130 Integrated Vacuum Concentrator System, and sent to Genome Quebec for genotyping, where they were processed on a newly developed genomic array providing information on a total of 76,050 SNPs across the caribou genome (Carrier et al. 2022).

2.2.2. SNP genotyping & filtering

Average genotyping success across samples for the 76,050 SNPs was 92.9%. Raw data was filtered and formatted for further analysis using PLINK v1.9 (Purcell et al. 2007). First, two samples that were run in duplicate were discarded. Another 13 samples were removed as duplicates (i.e., individuals

Table 2.1. Project samples (*N*=766) displayed by Designatable Unit (DU) and putative subpopulation (herd).

DU	Subpopulation (Herd)	# of Samples	# of Samples per DU
Boreal (DU6)	Calendar	9	74
	Chinchaga	23	
	Hay River	1	
	Maxhamish	11	
	Snake-Sahtaneh	22	
	Westside Fort Nelson	8	
Northern Mountain (DU7)	Atlin	27	398
	Carcross	6	
	Chase	47	
	East Wiliston	3	
	Finlay	5	
	Frog	7	
	Gataga	6	
	Graham	22	
	Horseranch	4	
	Itcha-Ilgachuz	78	
	Level-Kawdy	6	
	Little Rancheria	10	
	Muskwa	33	
	Pink Mountain	41	
	Telkwa	8	
	Tsenaglode	10	
	Tweedsmuir	36	
	Wolverine	49	
Central Mountain (DU8)	A La Peche	20	133
	Banff National Park	3	
	Jasper National Park	15	
	Kennedy Siding	16	
	Klinse-Za / Moberly	40	
	Narraway	4	
	Quintette	21	
	Redrock-Prairie Creek	14	
Southern Mountain (DU9)	Central Selkirks	6	161
	Columbia North	49	
	Columbia South	1	
	Groundhog	3	
	Hart Ranges	69	
	Narrow Lake	2	
	North Cariboo	21	
	Purcells South	5	
	South Selkirks	3	
	Wells Gray South	2	
TOTAL:	**42**	**766**	

with different sample IDs but the same genotypes) assuming a 0.02 mismatch rate threshold. These duplicates could be due to field sampling or lab errors, or individual caribou being unknowingly sampled twice. The PLINK flags --list-duplicate-vars and --exclude plink.dupvar were also enacted to identify and remove 1,323 duplicate SNPs.

Further filtering consisted of: (1) excluding individuals with <90% genotyping rate (--mind 0.1), which removed 14 individuals, (2) excluding 5,748 SNPs with <90% genotyping rate (--geno 0.1), (3) placing a Hardy-Weinberg equilibrium threshold (--hwe) of 1e-6, which removed 2,320 SNPs, (4) excluding 18,821 SNPs with a minor allele frequency <5% (--maf 0.05), (5) generating and removing a subset of 4,108 SNPs that are in approximate linkage equilibrium with each other (--indep-pairwise 50 5 0.5), and (6) removing 81 close relatives (parent-child or full sibling relationships, where one individual from each pair was retained) based on a --king-cutoff value of 0.177, or the equivalent of a PI_HAT value of 0.25. All filtering criteria were chosen based on relevant literature and recommendations for similar datasets (Cavedon et al. 2022a; Schweizer et al. 2016) and ultimately resulted in the retention of 658 individuals (of 766, or ~86%) and 43,730 SNPs (of 76,050, or ~58%; **Table 2.6**). Overall, most of the excluded samples and SNPs were removed because of biological reasons (e.g., close relatives, low MAF, high LD) rather than low genotyping rates.

2.2.3. Population structure analyses

Population structure analyses consisted of a variety of distance-and model-based methods. Both types of approaches are commonly used to infer population genetic structure and vary according to their respective assumptions of the dataset and difficulty of assessing the significance of clustering results.

2.2.3.1. Distance-based methods

As an initial assessment of population structure, I conducted a Principal Component Analysis (PCA) and a subsequent Discriminant Analysis of Principal Components (DAPC) in the *adegenet* v2.1.8 package in R (Jombart et al. 2010). I first used the dudi.pca function incorporating between-individual genetic distances obtained with PLINK (--distance-matrix) to generate a PCA and visualize overall patterns of variation among samples. In contrast to model-based programs such as *Structure*, PCAs and DAPCs do not rely on Hardy-Weinberg and linkage equilibrium assumptions within groups when performing *K*-means clustering. I implemented the find.clusters function in *adegenet* to identify the best-fitting number of clusters given Bayesian Information Criterion (BIC) values for *K* ranging from 1 to 20. This optimal clustering solution was then used to perform a DAPC, a multivariate analysis that minimizes within-group variance while maximizing among-group variance using variables known as discriminant functions (i.e., linear combinations of the original variables—in this case, SNPs; Jombart & Collins 2015). I then incorporated the first two discriminant functions into a scatterplot to visualize how variation is partitioned among identified groups.

2.2.3.2. Model-based methods

Population structure was further evaluated using the Bayesian clustering approach implemented in *Structure* v2.3.4, which groups individual genotypes into *K* clusters that maximize within-cluster Hardy-Weinberg and linkage equilibrium (Pritchard et al. 2000). I completed ten runs for each value of *K* ranging from 1 to 20 using the *admixture* model, both correlated and uncorrelated allele frequencies, and no *a priori* grouping of individuals by location. Each run consisted of a burn-in of 20,000 iterations followed by 50,000 Markov chain Monte Carlo (MCMC) repetitions, which was assessed as adequate based on convergence (Wang 2022). Using *Structure Harvester* v0.6.94 (Earl & vonHoldt 2012), I then obtained both a plot of log likelihood values, L(*K*), for each value of *K* and the *ΔK* statistic of Evanno et al.

(2005) to examine which value of *K* is best supported by the data. *Clumpp* v1.1.2 (Jackobsson & Rosenberg 2007) was used to consolidate and account for variation in clusters among the 10 runs/random iterations of *Structure*, while *Distruct* v1.1 (Rosenberg 2007) was used for visualization of genetic clustering results. To compare model-based results, I also ran *Admixture* v1.3, which uses a maximum likelihood approach and incorporates a cross-validation procedure that allows identification of the value of *K* for which the model has best predicative accuracy (Alexander et al. 2009). For this study, I used a 5-fold cross-validation flag for *K* clusters varying from 1 to 20.

Although population structure analysis programs are being increasingly modified to handle large SNP datasets, the accuracy of assigning individuals to clusters and/or their regions of origin is variable among approaches and models (Alvarez et al. 2021). To account for both this variability and the potential limitations associated with different programs, I also analyzed and attempted to corroborate population structure using *fastStructure*, a software based on the same variational Bayesian framework for posterior inference seen in *Structure* (Raj et al. 2014). I examined *K* ranging from 1 to 20 for 10 runs of *fastStructure* incorporating a convergence criterion of 1×10^{-6}, the simple prior, and a 5-fold cross-validation flag. Expected admixture proportions inferred by *fastStructure* were compared for different values of *K* using the chooseK.py script incorporated in the program. The script contains an algorithm for choosing the most appropriate number of model components that explain structure in the dataset without compromising prediction accuracy (Raj et al. 2014).

Lastly, I also examined population structure using the spatially-explicit program *Tess3* v1.0 implemented in R (Caye et al. 2016). In contrast to *Admixture* and *Structure*, *Tess* assigns individuals to the most likely number of clusters while incorporating information on each sample's geographic location. I conducted 10 runs of *Tess* for values of *K* ranging from 1 through 20 (tolerance = 1×10^{-7}, max. iterations = 1,000) and used the cross-entropy criterion to select the most optimal value of *K*, where the best *K* corresponds to the one with the lowest cross-validation score. The *Tess3r* package in R was also

used to create maps of the geographic distribution of genetic clusters and their corresponding geographic and/or topographic boundaries. Geographic information was obtained from BC government partners in the form of individual movement data for radio-collared caribou, in addition to location information associated with animal capture and sample collection. Individual home ranges were calculated using the *adehabitatHR* R package, which uses a 95% kernel density estimation (KDE) to create polygons associated with an individual's home range (Calenge 2006). The centroids of these polygons were then used as geographical coordinates for each sampled individual, where available. For individuals with limited or unknown movement data, capture coordinates were used instead.

2.2.4. Population differentiation

I utilized the *strataG* v2.4.905 package in R to quantify the degree of differentiation between clusters inferred using both distance- and model-based methods as well as compare the levels of diversity within them. Specifically, I calculated pairwise fixation index (F_{ST}) values, observed and expected heterozygosities (H_o and H_e), and departures from Hardy-Weinberg Equilibrium (HWE; Archer et al. 2016). To better compare clustering results obtained from various population structure programs to previously existing caribou classifications in BC, I also characterized diversity for pre-existing subpopulations (herds), populations (DUs, ecotypes, SARA-listed populations), as well as some of the optimal clusters inferred through DAPC, *Structure*, *Admixture*, *fastStructure*, and *Tess* analyses.

To visualize observed patterns in population structure and their associated phylogeny, I also calculated pairwise genetic distances between all individuals using the *ape* 5.2 R package (dist.gene function; Paradis et al. 2004). These genetic distances were then used to construct neighbor-joining trees incorporating 1000 bootstrap values, and subsequently visualized using the *ggtree* R package (Yu et al. 2017).

2.3. Results

2.3.1. Population structure analyses

2.3.1.1. Distance-based methods

A PCA incorporating the first and second principal components found preliminary clustering of individuals belonging to the Itcha-Ilgachuz and Tweedsmuir herds, which were most distinguishable along the first axis (**Figure 2.3**). There was also a gradual but distinct north-to-south geographic separation of individuals along the second axis, though neither DU structure, ecotype classification, nor SARA listing were clearly reflected.

The DAPC indicated six as the most optimal number of clusters found within the data (**Figure 2.4a**). Based on standard interpretation, the "best" number of clusters corresponds to the lowest BIC and is often indicated by an elbow in the curve of BIC values as a function of *K* (Jombart & Collins 2015). **Figure 2.4b** displays the resulting DAPC scatterplot, which reflects a transformation of the data using a PCA, followed by a Discriminant Analysis on the retained principal components. I retained a total of 100 PCs, since this was the number associated with the highest mean success as well as the lowest root mean squared error (RMSE), as recommended by Jombart & Collins (2015). Because the number of "optimal" clusters was observed to be less than 10 by the find.clusters function, I also chose to retain all eigenvalues for all four discriminant functions in the analysis (Jombart & Collins 2015). The resulting DAPC scatterplot exhibited similarities in shape to the PCA, where individuals belonging to the Itcha-Ilgachuz and Tweedsmuir herds formed their own respective clusters (**Figure 2.3**). Likewise, individuals classified as belonging to the Central and Southern Mountain DUs (DU8 and DU9, respectively) clustered together, while Northern Mountain DU7 and Boreal DU6 caribou were split into three clusters within close proximity to one another. The separation of groups along the y-axis once again reflected a north-to-south geographic gradient.

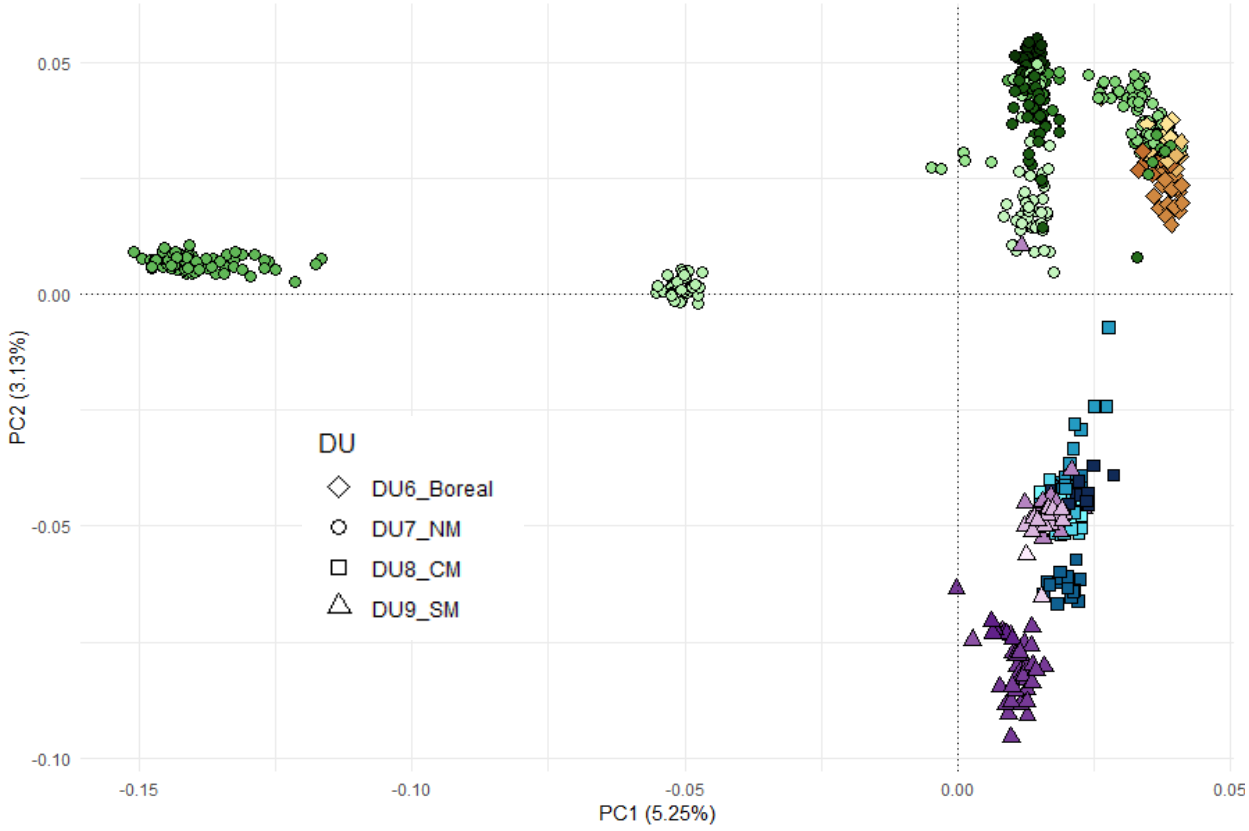

Figure 2.3. Principal Component Analysis (PCA) of caribou in British Columbia and Alberta. Points in the PCA represent sampled caribou individuals, categorized according to their respective Designatable Unit (DU; orange diamonds = Boreal DU6; green circles = Northern Mountain DU7; blue squares = Central Mountain DU8; purple triangles = Southern Mountain DU9). Variations in color shade represent different subpopulations (herds) in a given DU. The two leftmost clusters represent Itcha-Ilgachuz (far-left, medium-green circles) and Tweedsmuir (middle-left, light-green circles).

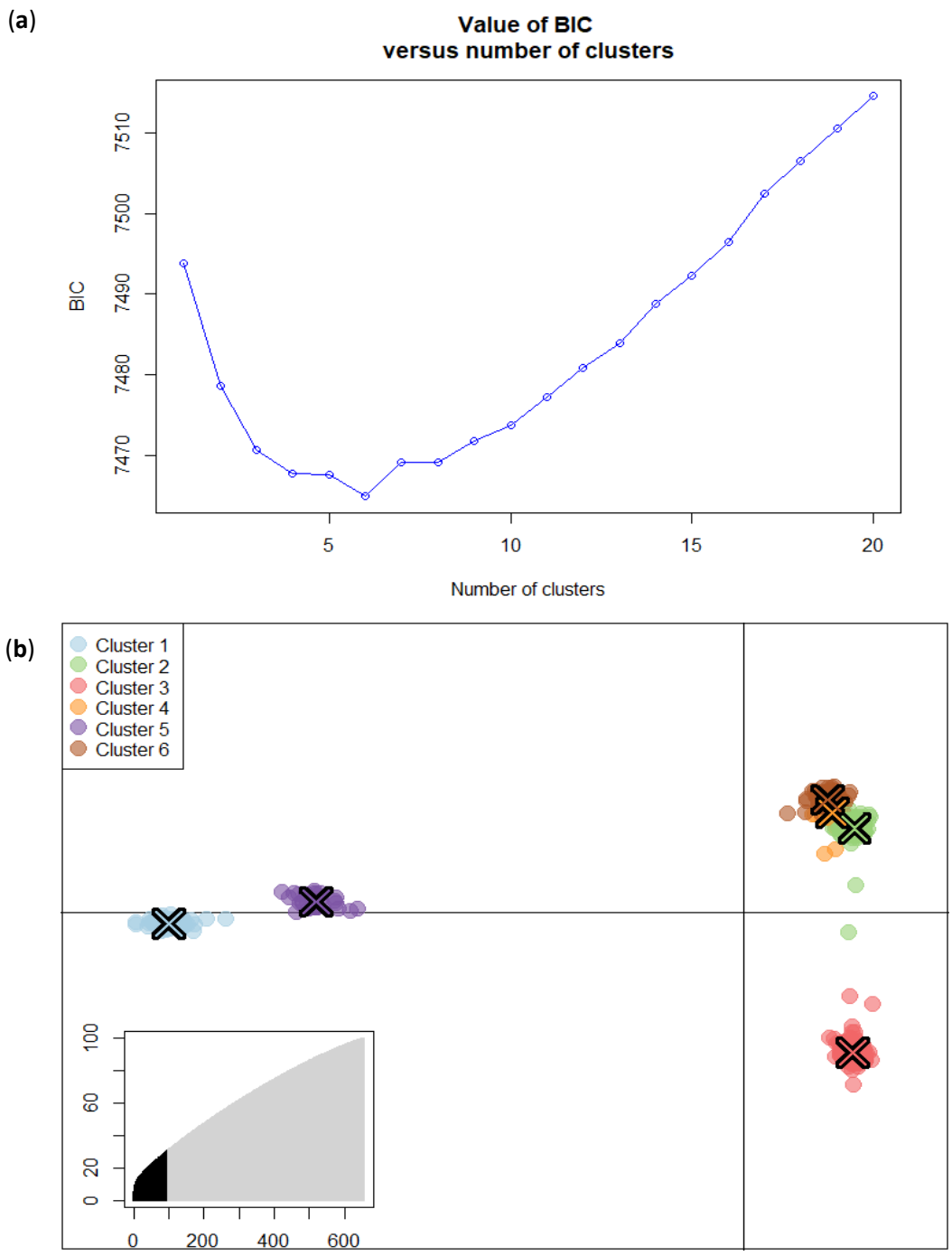

Figure 2.4. Discriminant Analysis of Principal Components (DAPC). (**a**) Value of the Bayesian Information Criterion (BIC) value as a function of *K*. The lowest value occurs at *K*=6, indicating the most likely optimal number of clusters. (**b**) Scatterplot of DAPC results, indicating the separation of individuals into six clusters. Clusters 1 (blue) and 5 (purple) represent the Itcha-Ilgachuz and Tweedsmuir herds, respectively; Cluster 3 (red) represents a merged group of Central Mountain DU8 and Southern Mountain DU9 individuals; Clusters 2 (green) and 6 (brown) are a split of Northern Mountain DU7 individuals; Cluster 4 (orange) is exclusively individuals belonging to Boreal DU6.

2.3.1.2. Structure

The most optimal *K* for *Structure* results was interpreted using a combination of the Pritchard et al. (2000) original method for inferring the most likely number of populations (i.e., Ln Pr(*X*|*K*), also known as the likelihood of *K* for each value of *K*), as well as the *ΔK* method proposed by Evanno et al. (2005), as suggested by Wang et al. (2017). These two methods provided slightly different solutions compared to both each other and to the DAPC results inferred above (**Figure 2.5**). For the Pritchard et al. (2000) original method, the most ideal clustering solution occurs when the Ln Pr(*X*|*K*) curve begins to level out and/or exhibit an asymptote; here, this value is ambiguous, with potential arguments for *K*=3 and *K*=4. The Evanno method *ΔK* plot, on the other hand, indicates an optimal *K* where *ΔK* is largest; in this dataset, these values occur at *K*=2 and *K*=9.

Bar plots for some of the most supported values of *K* inferred from both the DAPC and *Structure*, including *K*=2, *K*=3, *K*=4, and *K*=6, are presented in **Figure 2.6** and **Figure 2.7**. Notably, each of the

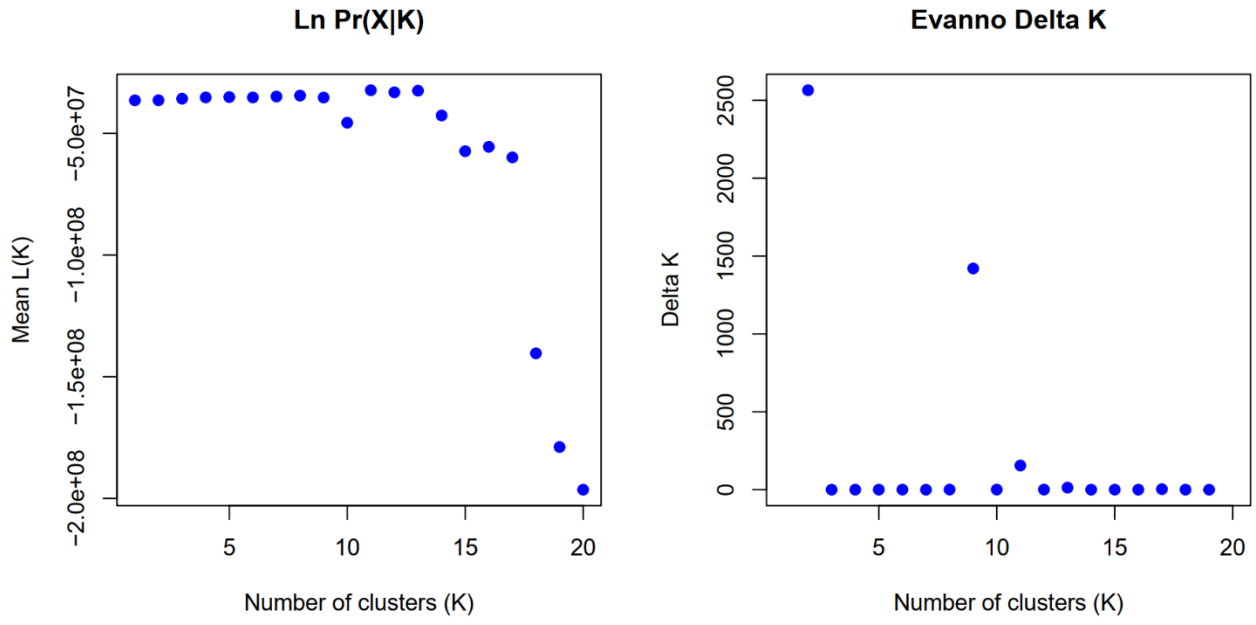

Figure 2.5. Deducing optimal *K* clusters from *Structure* results. (left) Plot incorporating the original Pritchard et al. (2000) method estimating Ln Pr(*X*|*K*). (right) Evanno *ΔK* plot proposed by Evanno et al. (2005).

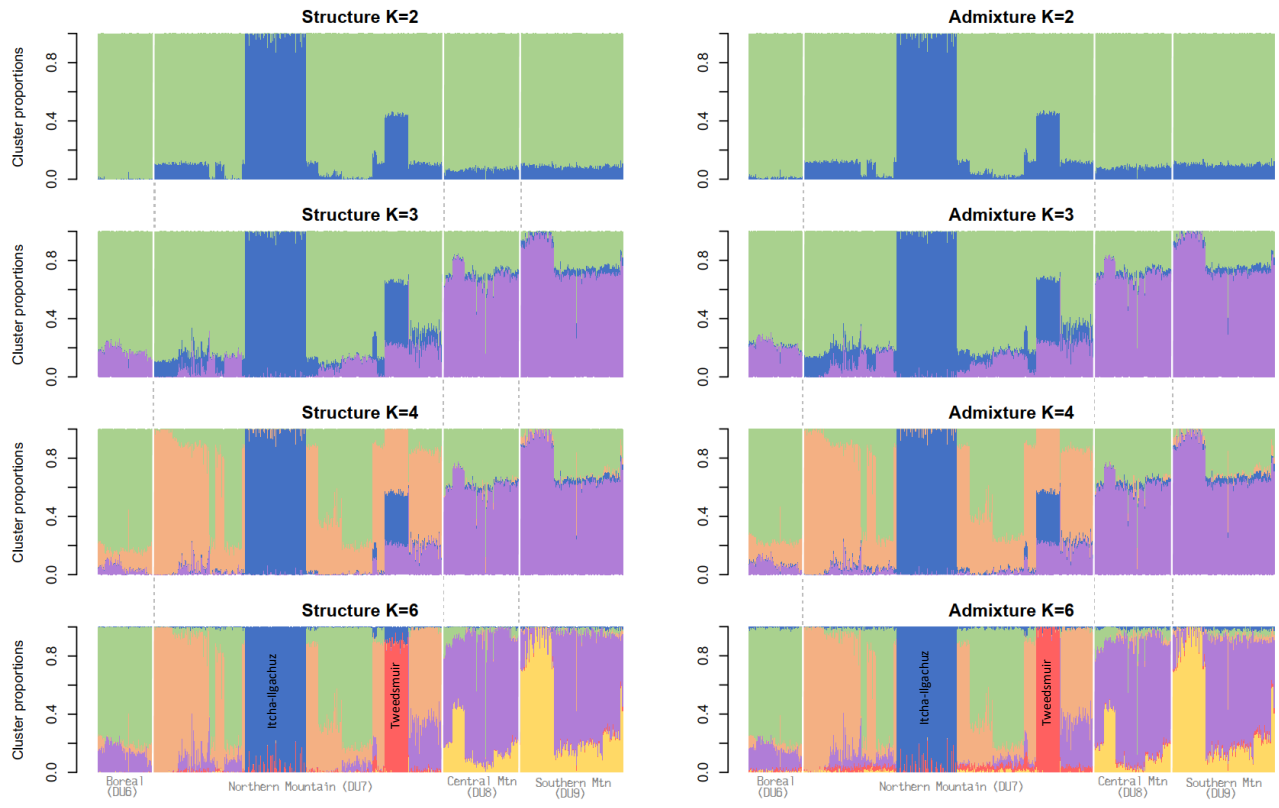

Figure 2.6. Bar plots of cluster proportions/assignment probabilities for *Structure* (left) and *Admixture* (right) for each of four values of *K*. Individuals are placed along the x-axis and grouped according to their assigned Designatable Unit (DU). The Itcha-Ilgachuz and Tweedsmuir subpopulations are labelled and grouped as the primarily blue and red clusters, respectively. As shown, differences in assignment probabilities between *Structure* and *Admixture* results were negligible.

clustering solutions created a separate group unique to individuals from the Itcha-Ilgachuz herd. Increasingly larger values of *K* first largely separated north and south individuals (*K*=3), then caribou belonging to the Boreal DU6/SARA population/ecotype (*K*=4), and finally created clusters reminiscent of—but not identical to—current DU structure, with two extra, genetically distinct groups representing the Itcha-Ilgachuz and Tweedsmuir subpopulations (*K*=6). Tweedsmuir was not recognized as its own cluster until *K*=6.

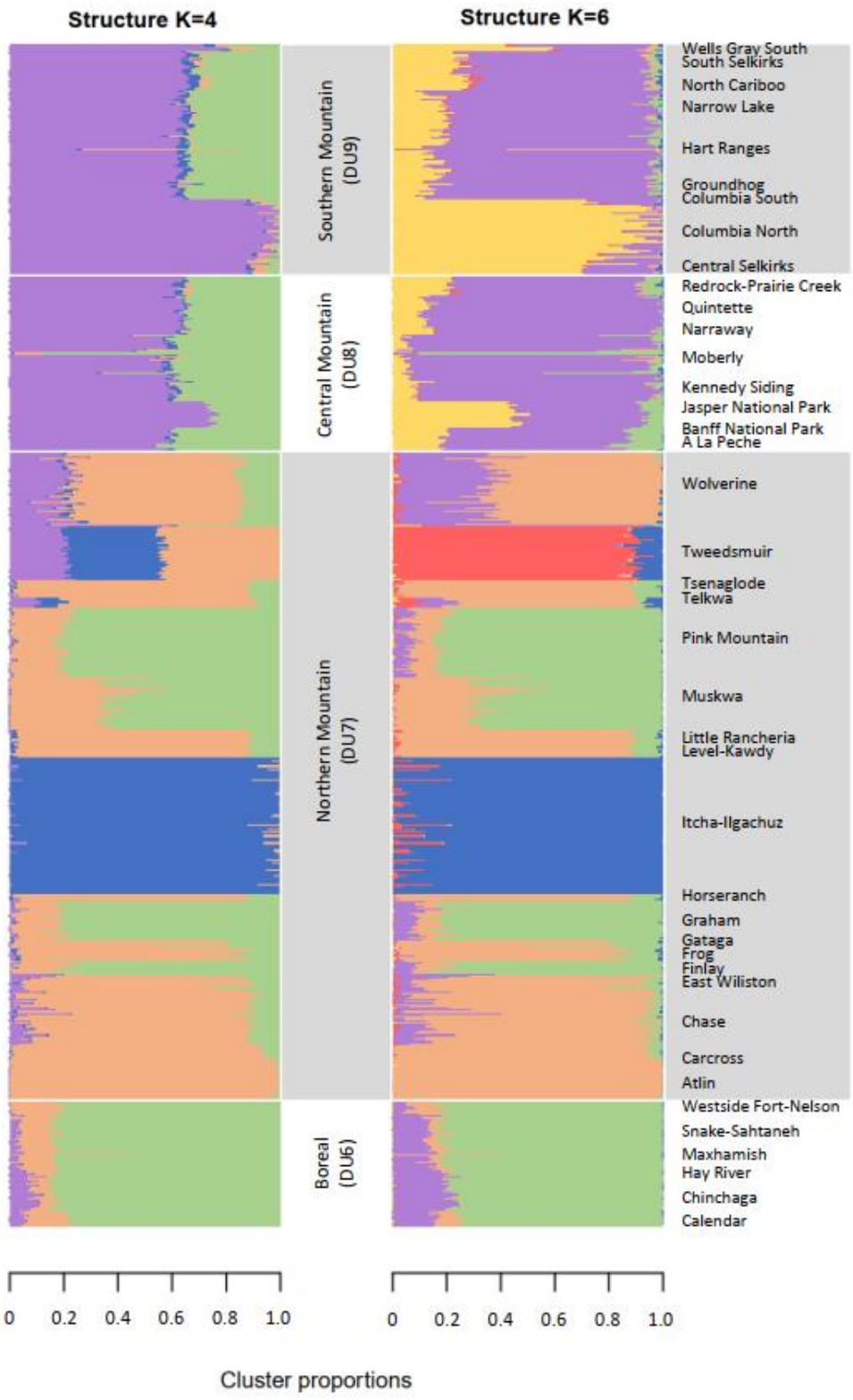

Figure 2.7. Bar plots of cluster proportions/assignment probabilities for *Structure* for values of *K*=4 and *K*=6. Individuals are placed along the y-axis and grouped according to their assigned Designatable Unit (DU) and herd.

2.3.1.3. Admixture

Cross-validation (CV) scores for the *Admixture* analysis were lowest for *K*=13 (**Figure 2.8**). Nonetheless, the curve began to level out around the *K*=3 to *K*=4 mark, yielding further support for these values as possible population solutions, similar to what was presented in *Structure* results. Differences in the probability of individual assignment to clusters between *Admixture* and *Structure* were negligible (**Figure 2.6**). The assignment of individuals to clusters for the aforementioned values of *K* further emphasized the breakdown of groups described in the section above, where Itcha-Ilgachuz individuals always clustered together first, but resulting maps also highlighted transitional zones of admixed individuals among clusters (**Figure 2.9**). Notably, the size of these transitional zones was significant between the currently recognized Boreal DU6 and Northern Mountain DU7 as well as the Northern Mountain DU7 and Central Mountain DU8, in particular.

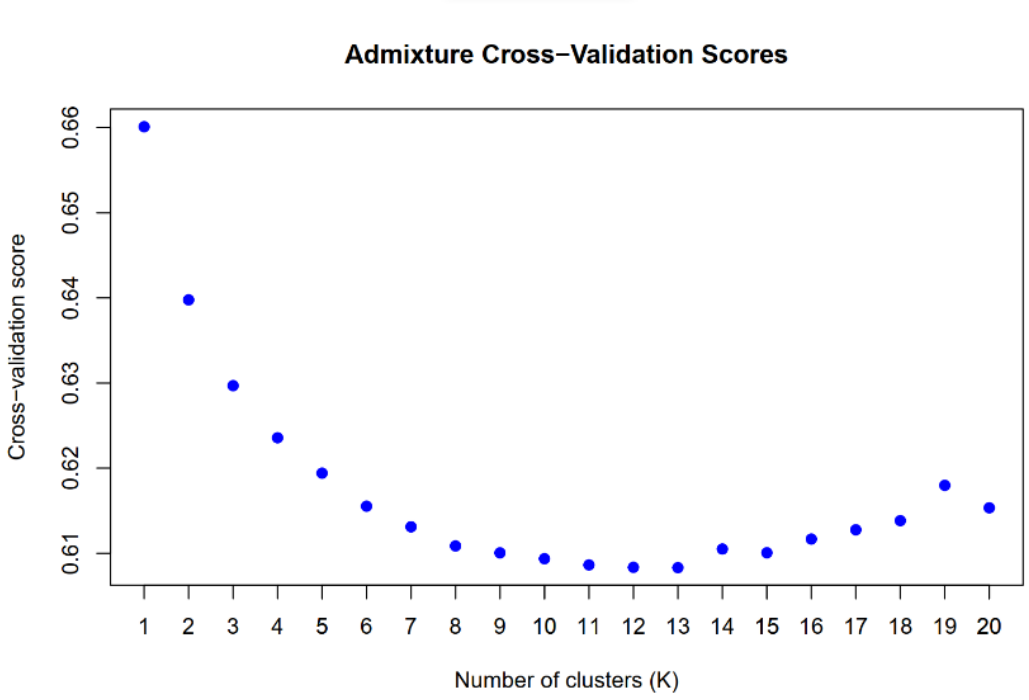

Figure 2.8. *Admixture* cross-validation (CV) scores as a function of number of clusters (*K*) for each value of *K* ranging from 1 to 20.

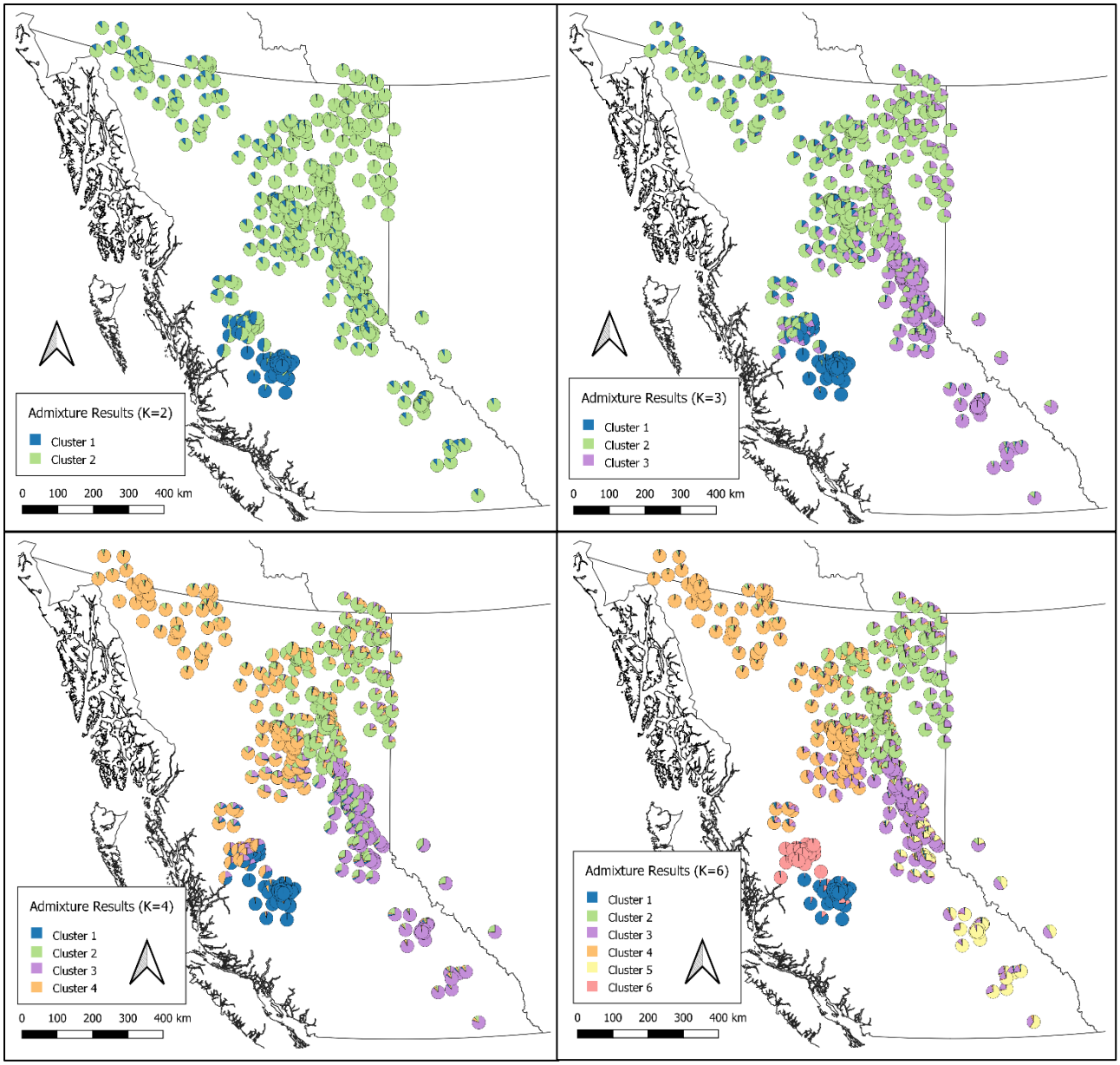

Figure 2.9. Caribou sampled in British Columbia and Alberta represented by pie chart assignment probabilities to a given number of clusters (*K*; one individual = one pie chart). Maps are displayed for *K*=2, *K*=3, *K*=4, and *K*=6 clusters derived from *Admixture* results. Maps created in QGIS.

2.3.1.4. fastStructure

fastStructure results indicated an average model complexity that maximizes marginal likelihood of 9 and an average number of model components used to explain data structure as 15. Based on simulations conducted by Raj et al. (2014) and their associated interpretations, this means that the most optimal number of clusters likely falls between *K*=9 and *K*=15 and is thus indicative of finer population structure than what was inferred by some other software. It is also worth noting that *fastStructure* did not pick up significant clustering signals for smaller values of *K*, as has been the case with most other structuring programs.

2.3.1.5. Tess

Tess results were similar to many of the patterns generated by other population structure programs. In the *Tess* cross-validation criterion plot (**Figure 2.10**), which is interpreted much the same as *Admixture*'s CV plot, values continuously declined up to *K*=20 (the largest *K* tested), which could be indicative of the presence of finer levels of population structure (i.e., *K*>6; Caye et al. 2016). However, as for *Admixture*, the curve started to level-out the most between *K*=2 and *K*=4. *Tess*'s geographic predictions for four values of *K* (*K*=2, *K*=3, *K*=4, and *K*=6) are presented in **Figure 2.11**. Geographic predictions for these values of *K* overlapped almost identically with cluster assignments derived from *Admixture* results and further highlighted the transitional zones mentioned above (**Figure 2.12**).

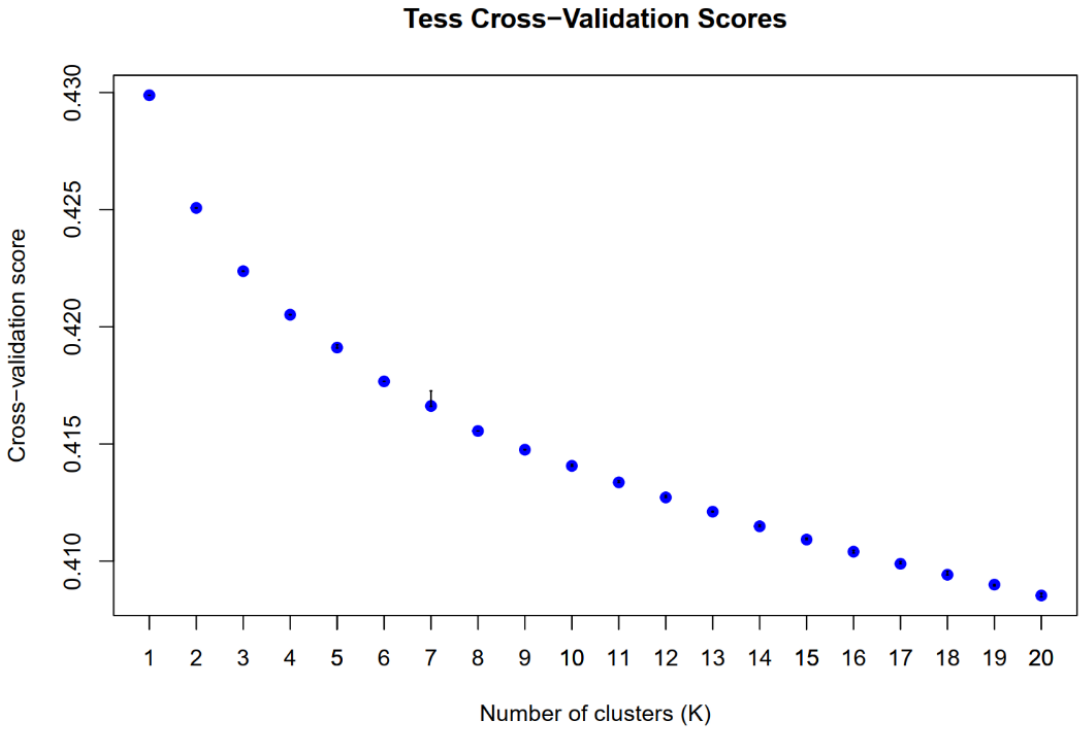

Figure 2.10. *Tess* cross-validation (CV) scores as a function of number of clusters (*K*) for each value of *K* ranging from 1 to 20.

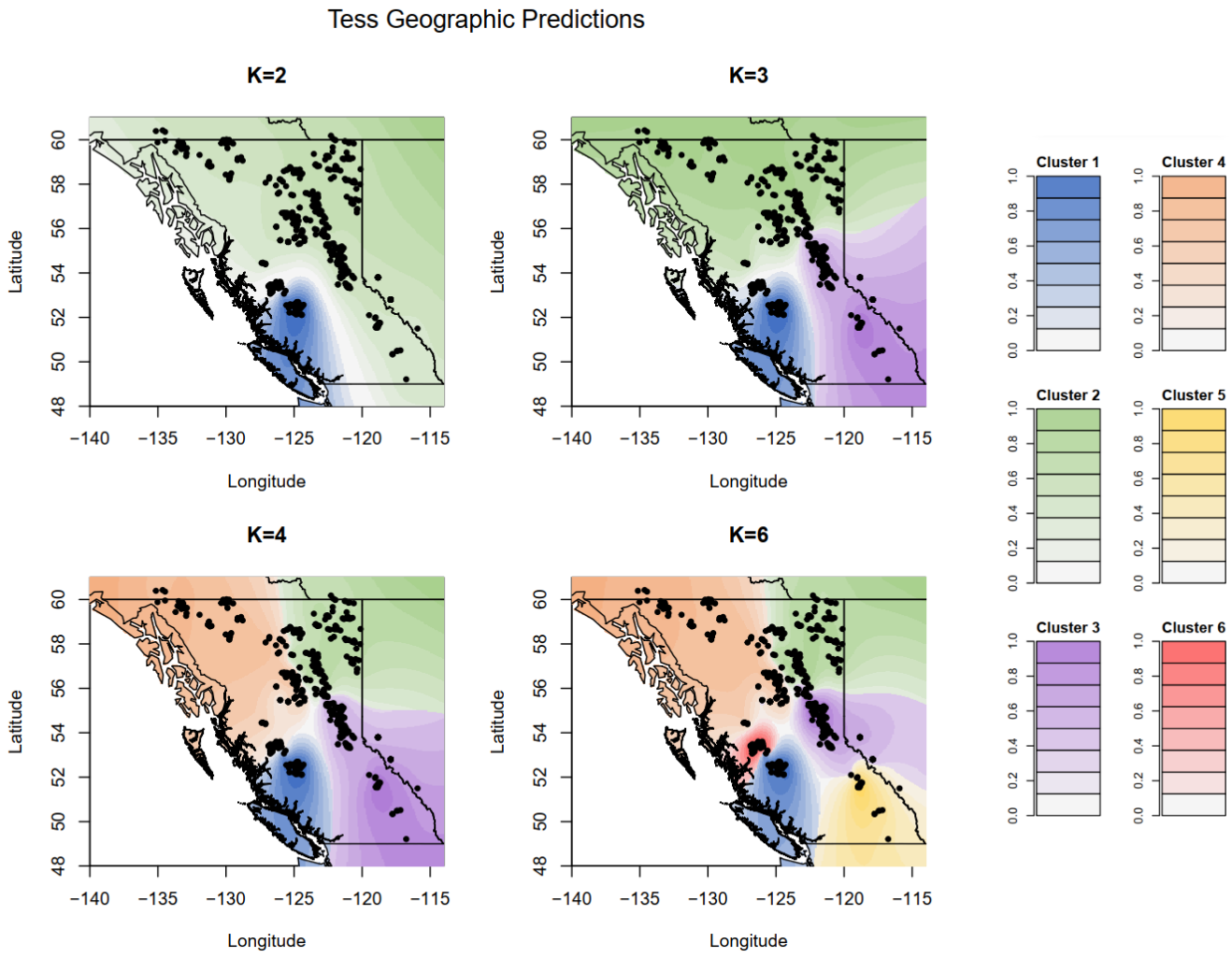

Figure 2.11. Geographic predictions inferred from *Tess* for *K*=2, *K*=3, *K*=4, and *K*=6. Sampled individuals are represented as black dots across the map. Maps created using the *tess3r* package in R.

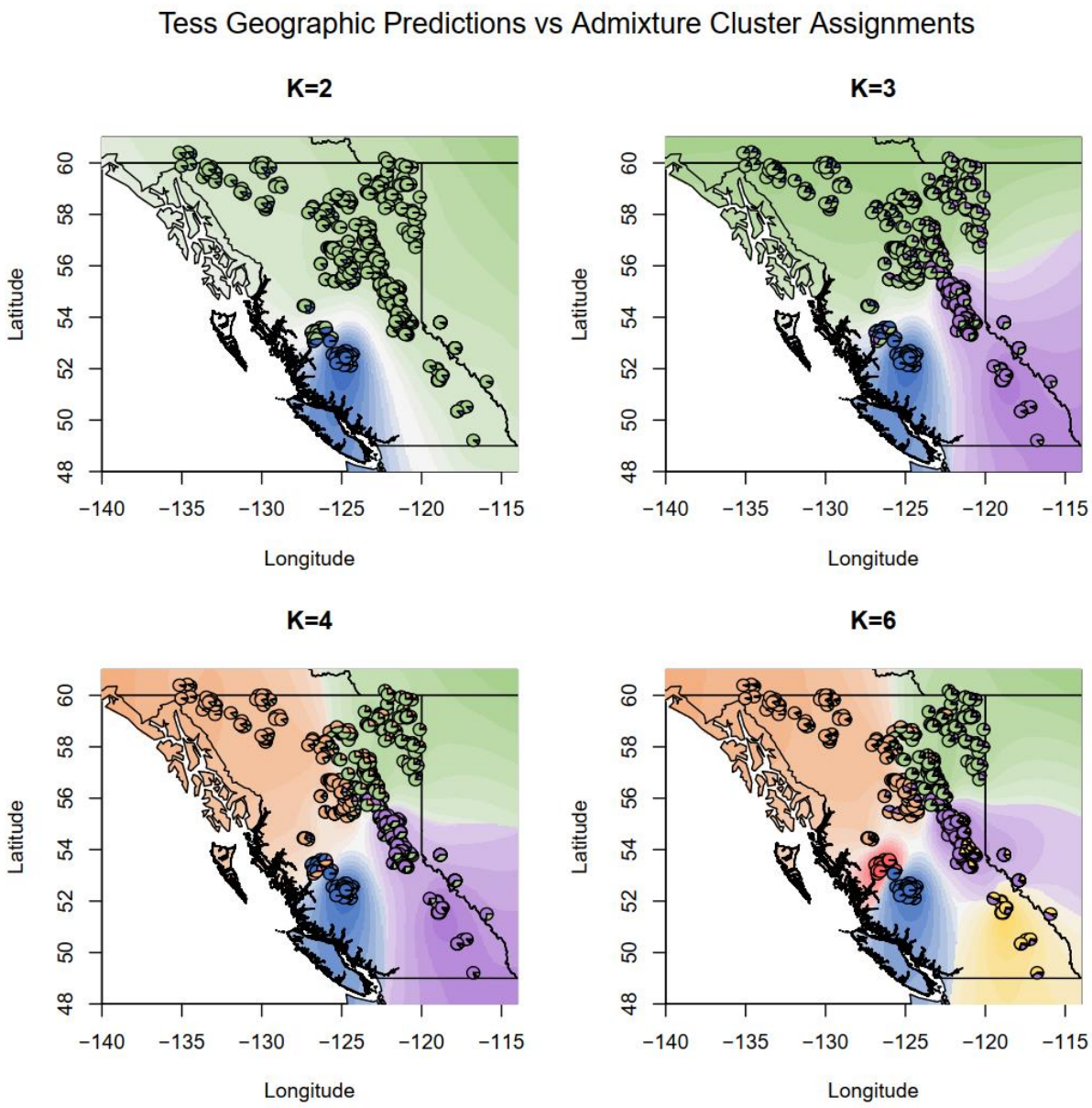

Figure 2.12. Geographic predictions inferred from *Tess* for *K*=2, *K*=3, *K*=4, and *K*=6. Sampled individuals are depicted as pie charts, which represent the probabilities of individual assignment to cluster for each value of *K*. Maps created using the *tess3r* package in R.

2.3.2. Population differentiation

Pairwise F_{ST} values between both inferred populations (clusters) and previously defined "populations" (i.e., DUs, ecotypes, and SARA-listed populations) are presented in **Table 2.2** and **Table 2.3**. All groups showed some level of genetic distinctiveness from one another. While F_{ST} values were relatively low (0.0128–0.0700 for inferred clusters), they were significantly different from zero for all pairwise comparisons, even after Bonferroni correction for multiple testing. The highest level of differentiation was observed between Clusters 1 and 5 for K=6, which broadly corresponded to Itcha-Ilgachuz and Southern Mountain individuals, respectively (F_{ST}=0.0700). Pairwise F_{ST} values calculated for DUs, ecotypes, and SARA-listed populations were all considerably lower (<0.0223). Observed and expected heterozygosity values ranged from 0.3231–0.4075 for inferred clusters and were in a similar range for existing caribou classification schemes (**Table 2.4; Table 2.5; Table 2.6**). The lowest heterozygosity values were found for inferred cluster 1—the group composed of exclusively Itcha-Ilgachuz individuals (H_e=0.3264; H_o=0.3231). I did not find evidence for systematic deviations from HWE or linkage disequilibrium in specific populations.

The observed neighbor-joining tree closely resembled population structure results obtained with other analyses, with individuals belonging to the Itcha-Ilgachuz subpopulation forming their own cluster separate from all other DUs and populations (**Figure 2.13; Figure 2.14**). The cluster deemed genetically closest to Itcha-Ilgachuz was the Tweedsmuir cluster, which also formed its own group. Notably, both Itcha-Ilgachuz and Tweedsmuir were more closely related to individuals derived from Central Mountain DU8 and Southern Mountain DU9 than to Northern Mountain DU7 and Boreal DU6, though it is also worth acknowledging that the Central and Southern Mountain DUs did not form their own distinct clusters.

Table 2.2. Pairwise F_{ST} values for inferred populations (K=2, K=3, K=4, and K=6) given below the diagonal, with respective p-values above the diagonal. Clusters are colored according to distributions in **Figure 2.9**—i.e., for K=6, Cluster 1 (blue) is Itcha-Ilgachuz, Cluster 2 (green) is Boreal, Cluster 3 (purple) is Central Mountain, Cluster 4 (orange) is Northern Mountain, Cluster 5 (yellow) is Southern Mountain, and Cluster 6 (red) is Tweedsmuir. Minimum and maximum F_{ST} values are highlighted in bold (range=0.0128–0.700). All pairwise estimates were significant (p≤0.001).

***K* = 2**

	Cluster 1	Cluster 2
Cluster 1	–	<0.001
Cluster 2	0.0467	–

***K* = 3**

	Cluster 1	Cluster 2	Cluster 3
Cluster 1	–	<0.001	<0.001
Cluster 2	0.0362	–	<0.001
Cluster 3	0.0378	**0.0128**	–

***K* = 4**

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	–	<0.001	<0.001	<0.001
Cluster 2	0.0575	–	<0.001	<0.001
Cluster 3	0.0525	0.0157	–	<0.001
Cluster 4	0.0468	0.0129	0.0151	–

***K* = 6**

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Cluster 1	–	<0.001	<0.001	<0.001	<0.001	<0.001
Cluster 2	0.0576	–	<0.001	<0.001	<0.001	<0.001
Cluster 3	0.0528	0.0151	–	<0.001	<0.001	<0.001
Cluster 4	0.0516	0.0131	0.0164	–	<0.001	<0.001
Cluster 5	**0.0700**	0.0341	0.0200	0.0327	–	<0.001
Cluster 6	0.0570	0.0448	0.0401	0.0381	0.0565	–

Table 2.3. Pairwise F_{ST} values for current caribou classification schemes (ecotypes, COSEWIC Designatable Units, and SARA-listed populations) given below the diagonal, with respective *p*-values above the diagonal. For Designatable Units (DUs): DU6 = Boreal; DU7 = Northern Mountain; DU8 = Central Mountain; DU9 = Southern Mountain. For SARA-listed populations: NM = Northern Mountain; SM – NG = Southern Mountain – Northern Group; SM – CG = Southern Mountain – Central Group; SM – SG = Southern Mountain – Southern Group. Categories are colored according to distributions in **Figures 2.1** and **2.2.** Minimum and maximum F_{ST} values are highlighted in bold (range=0.0049–0.0223). All pairwise estimates were significant ($p \leq 0.001$).

Ecotype

	Boreal	Northern	Mountain
Boreal	–	<0.001	<0.001
Northern	0.0138	–	<0.001
Mountain	0.0220	0.0102	–

Designatable Unit (DU)

	DU6	DU7	DU8	DU9
DU6	–	<0.001	<0.001	<0.001
DU7	0.0154	–	<0.001	<0.001
DU8	0.0188	0.0135	–	<0.001
DU9	0.0220	0.0146	**0.0049**	–

SARA-listed Population

	Boreal	NM	SM – NG	SM – CG	SM – SG
Boreal	–	<0.001	<0.001	<0.001	<0.001
NM	0.0129	–	<0.001	<0.001	<0.001
SM – NG	**0.0223**	0.0125	–	<0.001	<0.001
SM – CG	0.0192	0.0162	0.0175	–	<0.001
SM – SG	0.0216	0.0173	0.0174	0.0052	–

Table 2.4. Observed and expected heterozygosity values (H_e and H_o) for inferred populations (K=2, K=3, K=4, and K=6). Clusters are colored according to distributions in **Figure 2.9**—i.e., for K=6, Cluster 1 (blue) is Itcha-Ilgachuz, Cluster 2 (green) is Boreal, Cluster 3 (purple) is Central Mountain, Cluster 4 (orange) is Northern Mountain, Cluster 5 (yellow) is Southern Mountain, and Cluster 6 (red) is Tweedsmuir.

***K* = 2**

	H_e	H_o
Cluster 1	0.3264	0.3231
Cluster 2	0.4075	0.3816

***K* = 3**

	H_e	H_o
Cluster 1	0.3497	0.3310
Cluster 2	0.4073	0.3878
Cluster 3	0.3988	0.3793

***K* = 4**

	H_e	H_o
Cluster 1	0.3273	0.3241
Cluster 2	0.4047	0.3908
Cluster 3	0.3987	0.3792
Cluster 4	0.3990	0.3791

***K* = 6**

	H_e	H_o
Cluster 1	0.3264	0.3231
Cluster 2	0.4039	0.3900
Cluster 3	0.3994	0.3868
Cluster 4	0.3980	0.3836
Cluster 5	0.3677	0.3480
Cluster 6	0.3464	0.3475

Table 2.5. Observed and expected heterozygosity values (H_e and H_o) for current caribou classification schemes (ecotypes, COSEWIC Designatable Units, and SARA-listed populations). For Designatable Units (DUs): DU6 = Boreal; DU7 = Northern Mountain; DU8 = Central Mountain; DU9 = Southern Mountain. For SARA-listed populations: NM = Northern Mountain; SM – NG = Southern Mountain – Northern Group; SM – CG = Southern Mountain – Central Group; SM – SG = Southern Mountain – Southern Group. Categories are colored according to distributions in **Figures 2.1** and **2.2.**

Ecotype

	H_e	H_o
Boreal	0.4004	0.3883
Northern	0.4040	0.3737
Mountain	0.3947	0.3762

Designatable Unit (DU)

	H_e	H_o
DU6	0.4004	0.3882
DU7	0.4003	0.3712
DU8	0.4001	0.3833
DU9	0.3947	0.3762

SARA-listed Population

	H_e	H_o
Boreal	0.4004	0.3883
NM	0.4004	0.3877
SM – NG	0.4033	0.3597
SM – CG	0.3994	0.3818
SM – SM	0.3955	0.3777

Table 2.6. Filtered project samples (*N*=658) displayed with mean expected (H_e) and observed (H_o) heterozygosity values. Herds with the lowest and highest mean values are highlighted in dark and light grey, respectively.

DU	Subpopulation (Herd)	# of Samples	Mean H_e *(range)*	Mean H_o *(range)*
Boreal (DU6)	Calendar	9	0.3746 *(0.303—0.4027)*	0.3757 *(0.3052—0.4034)*
	Chinchaga	21	0.3409 *(0.3111—0.392)*	0.3421 *(0.3122—0.3935)*
	Hay River	1	0.3634 *(0.3634—0.3634)*	0.3657 *(0.3657—0.3657)*
	Maxhamish	9	0.3547 *(0.3326—0.3934)*	0.356 *(0.3339—0.3946)*
	Snake-Sahtaneh	21	0.3841 *(0.311—0.4044)*	0.385 *(0.311—0.405)*
	Westside Fort Nelson	8	0.3799 *(0.3607—0.395)*	0.3807 *(0.3614—0.3957)*
Northern Mountain (DU7)	Atlin	26	0.3474 *(0.2841—0.3987)*	0.3482 *(0.2846—0.3994)*
	Carcross	5	0.3661 *(0.3184—0.3911)*	0.3668 *(0.3185—0.3914)*
	Chase	40	0.3674 *(0.3158—0.3981)*	0.3687 *(0.3178—0.3993)*
	East Wiliston	3	0.3349 *(0.3182—0.3564)*	0.3357 *(0.3194—0.3569)*
	Finlay	5	0.3918 *(0.3899—0.3956)*	0.3922 *(0.39—0.3952)*
	Frog	6	0.3781 *(0.3665—0.389)*	0.379 *(0.3672—0.3908)*
	Gataga	5	0.345 *(0.3217—0.3663)*	0.3457 *(0.321—0.367)*
	Graham	22	0.3774 *(0.3194—0.4017)*	0.3786 *(0.3211—0.4021)*
	Horseranch	4	**0.396** *(0.3822—0.4042)*	**0.3964** *(0.3832—0.4039)*
	Itcha-Ilgachuz	76	0.3794 *(0.3185—0.4025)*	0.3801 *(0.3191—0.4028)*
	Level-Kawdy	6	0.3937 *(0.3812—0.4052)*	0.3949 *(0.3826—0.4056)*
	Little Rancheria	10	0.3951 *(0.3878—0.4049)*	0.3958 *(0.3885—0.4063)*
	Muskwa	29	0.3886 *(0.3231—0.4066)*	0.3894 *(0.3237—0.4073)*
	Pink Mountain	39	**0.3245** *(0.3012—0.3853)*	**0.3254** *(0.3032—0.3863)*
	Telkwa	5	0.3894 *(0.381—0.3945)*	0.3901 *(0.3824—0.3964)*
	Tsenaglode	10	0.3845 *(0.339—0.3996)*	0.3854 *(0.3412—0.4005)*
	Tweedsmuir	30	0.3837 *(0.2957—0.4016)*	0.3848 *(0.2966—0.4029)*
	Wolverine	42	0.387 *(0.3383—0.401)*	0.3879 *(0.3409—0.4013)*
Central Mountain (DU8)	A La Peche	13	0.3873 *(0.3685—0.3989)*	0.388 *(0.3691—0.4009)*
	Banff National Park	1	0.3905 *(0.3905—0.3905)*	0.3916 *(0.3916—0.3916)*
	Jasper National Park	14	0.3801 *(0.3405—0.3992)*	0.381 *(0.3408—0.3992)*
	Kennedy Siding	12	0.3746 *(0.2642—0.4014)*	0.3752 *(0.2654—0.4027)*
	Klinse-Za / Moberly	25	0.3825 *(0.3387—0.405)*	0.3832 *(0.3389—0.406)*
	Narraway	4	0.3846 *(0.3702—0.3971)*	0.3856 *(0.3706—0.3985)*
	Quintette	18	0.3752 *(0.2861—0.3978)*	0.3759 *(0.2863—0.3987)*
	Redrock-Prairie Creek	11	0.3818 *(0.3336—0.3961)*	0.3828 *(0.3341—0.3972)*
Southern Mountain (DU9)	Central Selkirks	6	0.3901 *(0.3733—0.3991)*	0.3913 *(0.3748—0.4002)*
	Columbia North	34	0.3809 *(0.2955—0.4079)*	0.3818 *(0.2963—0.4078)*
	Columbia South	1	0.3583 *(0.3583—0.3583)*	0.3594 *(0.3594—0.3594)*
	Groundhog	2	0.3877 *(0.384—0.3914)*	0.3893 *(0.3862—0.3924)*
	Hart Ranges	61	0.3848 *(0.319—0.4034)*	0.3856 *(0.32—0.4051)*
	Narrow Lake	2	0.3931 *(0.3907—0.3955)*	0.3937 *(0.3916—0.3957)*
	North Cariboo	20	0.3847 *(0.3504—0.4069)*	0.3852 *(0.3497—0.4071)*
	South Selkirks	1	0.3749 *(0.3749—0.3749)*	0.3757 *(0.3757—0.3757)*
	Wells Gray South	1	0.3825 *(0.3825—0.3825)*	0.3826 *(0.3826—0.3826)*
TOTAL:	**41**	**658**		

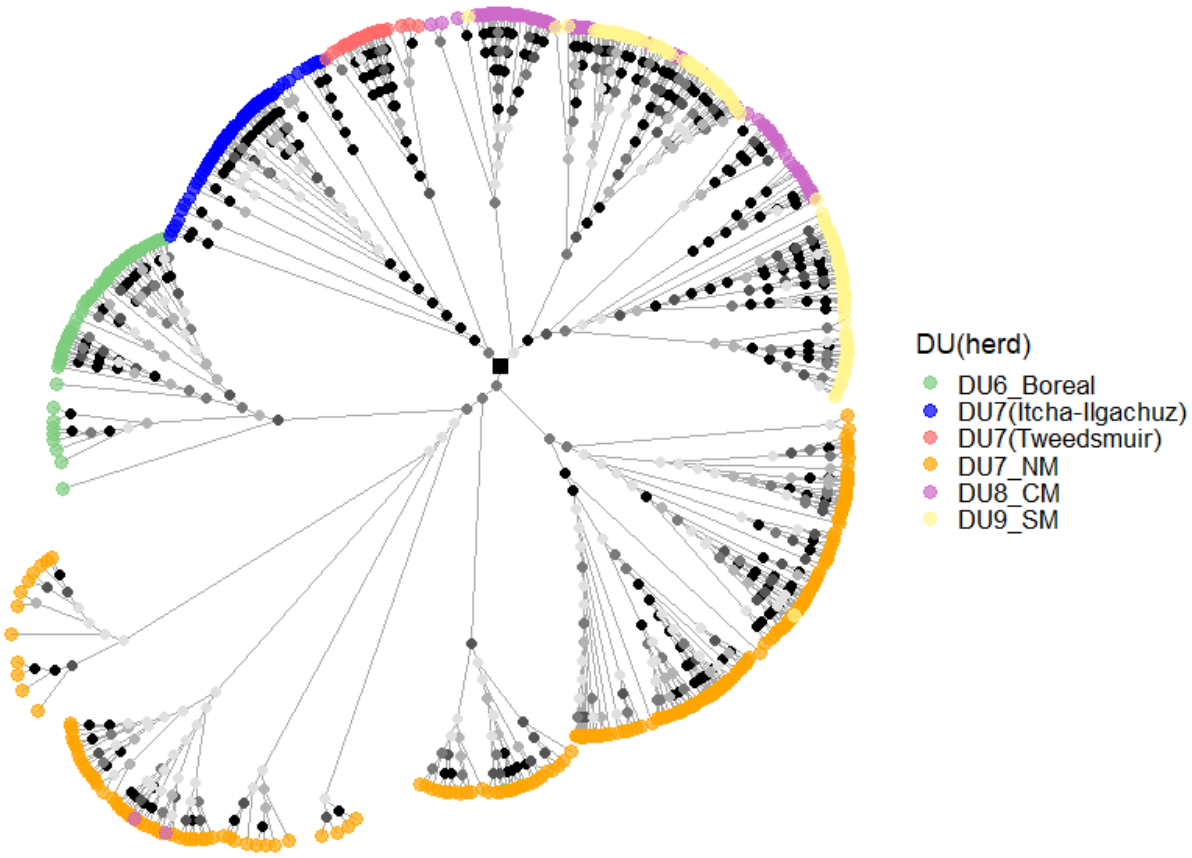

Figure 2.13. Neighbor-joining tree of caribou sampled throughout British Columbia and Alberta. Branches were calculated examining SNP data and represent caribou individuals, with branch tips colored according to each individual's corresponding Designatable Unit (DU). Bootstrap values were estimated based on 1000 replicates and are represented on internal nodes as circles in five classes/shades of grey (each class is formed by 20 bootstrap values; darker circle values range from 81–100%). The dark square represents the Itcha-Ilgachuz node and has a bootstrap value of 97%. Abbreviations: NM = Northern Mountain, CM = Central Mountain, SM = Southern Mountain.

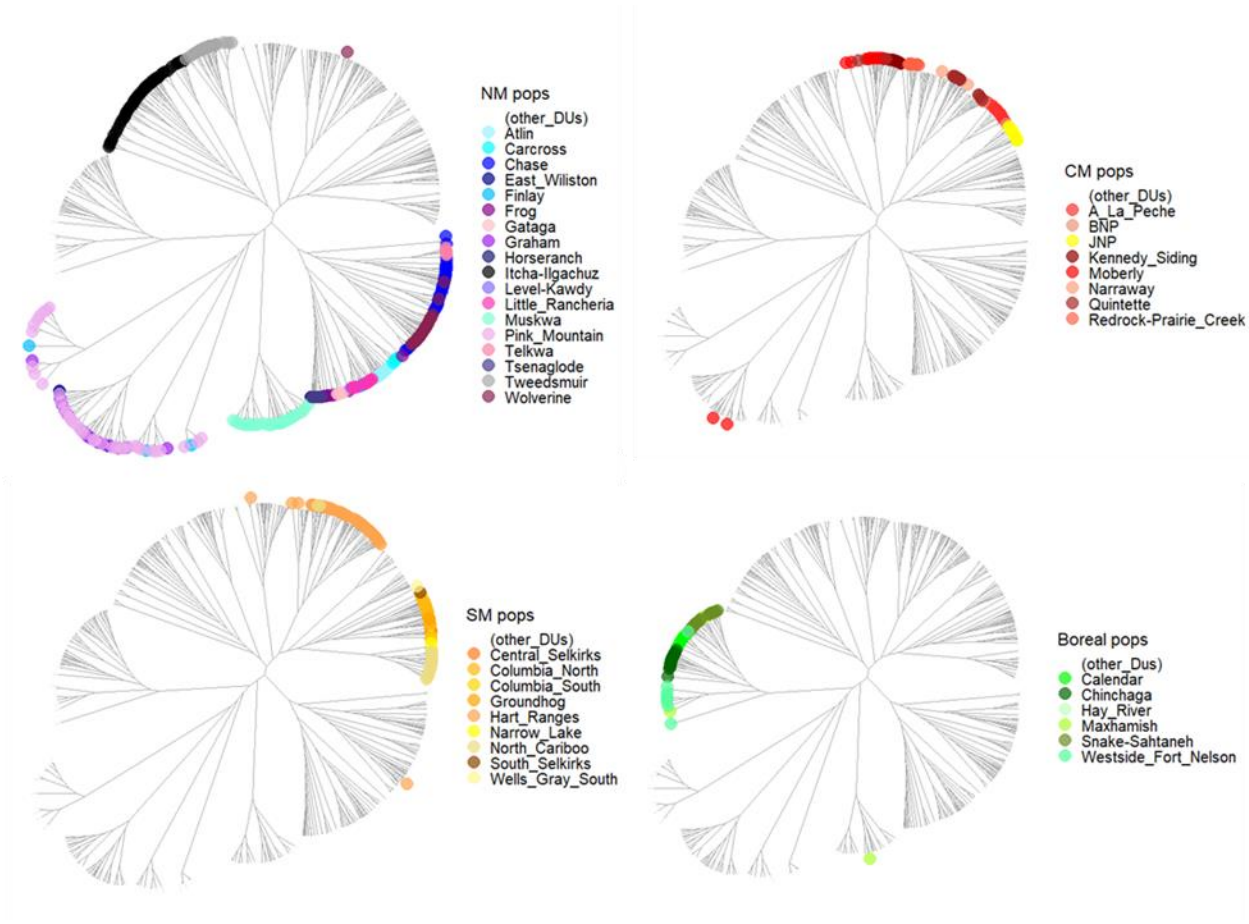

Figure 2.14. Neighbor-joining trees for each of the four Designatable Units (DUs) present in British Columbia and Alberta. Branches were calculated examining SNP data and represent caribou individuals, with branch tips colored according to each individual's corresponding subpopulation within each DU. Abbreviations: NM = Northern Mountain, CM = Central Mountain, SM = Southern Mountain, BNP = Banff National Park, JNP = Jasper National Park.

2.4. Discussion

This study builds on previous caribou genetic and genomic surveys to characterize caribou population structure in western Canada. I found support for several optimal clustering solutions, including *K*=2, *K*=3, *K*=4, and *K*=6. Importantly, my phylogenetic analysis grouped the Itcha-Ilgachuz and Tweedsmuir subpopulations of caribou closer to Central and Southern Mountain individuals, indicating that these two herds may be more closely related to caribou from the southern part of BC than to Northern Mountain or Boreal individuals. This result is supported by previous work conducted by Taylor et al. (2020) and (2021), who also found a unique signal in Itcha-Ilgachuz individuals (*n*=2) using whole-genome sequences in tandem with mtDNA. The authors hypothesized that the subpopulation likely experienced introgression preceded by an ancient colonization event, which now reflects its divergence from other northern ecotype/Northern Mountain DU7 caribou.

F_{ST} values and levels of observed and expected heterozygosities were consistent with what is expected for caribou populations from the literature and generally considered "moderate" in the context of mammals, but were not reflective of current DU structure (Cavedon et al. 2019; McDevitt et al. 2009; Goossens et al. 2016; Schweizer et al. 2016). The lowest F_{ST} value (0.0049) was estimated between existing DUs 8 (Central Mountain) and 9 (Southern Mountain), indicating little to no differentiation between the two populations at their current boundary north of the Hart Ranges subpopulation. Although this value increased slightly (F_{ST}=0.0200) between the two inferred clusters vaguely representing the Central and Southern Mountain DUs at *K*=6, phylogeny results suggested that these two populations are closely related and may not reflect two distinct units. This mirrors distance-based structuring results, where for *K*=6, the observed DAPC grouped Central and Southern Mountain individuals into a single cluster.

2.4.1. Management implications

It is clear from the results presented here that neither DUs, SARA Schedule 1 populations, nor ecotypes adequately capture the genetic diversity caribou exhibit throughout their range in western Canada. This suggests that the specific boundaries of caribou DUs ought to be reconsidered, since caribou units that do not accurately represent regional population structure and possible underlying ecological relationships may lead to recovery plans and conservation actions that are ill-suited to protect the species at risk (Crandall 2009; Polfus et al. 2016). Based on the genetic results of this study and their integration with existing caribou knowledge, a four-unit re-classification scheme for western Canadian caribou—detailed below—may be appropriate.

First, the Itcha-Ilgachuz subpopulation as well as surrounding subpopulations, including Tweedsmuir, Rainbows, and Charlotte Alplands, appear to be significantly genetically distinct. At K=2, which was most strongly supported by the program *Structure*, caribou throughout the province were divided into two clusters representing: (1) exclusively individuals belonging to the Itcha-Ilgachuz subpopulation and (2) all remaining individuals, with significant admixture apparent in Tweedsmuir individuals (F_{ST}=0.0467). It is worth noting that the greatest number of individuals was sampled from the Itcha-Ilgachuz subpopulation (n=78), while the average across all other subpopulations was approximately 17 individuals per herd, though still quite variable (range: 1–69). To account for this, I also conducted multiple runs of *Structure* incorporating different models and parameters, namely a combination of the alternative ancestry prior, an initial α value much smaller than the default, and the uncorrelated allele frequency model, as recommended by Wang (2017) for cases of unbalanced sampling. Neither one of these parameter or model changes significantly affected Ln Pr($X|K$) or ΔK results and associated optimal K solutions. Similar patterns in the context of the Tweedsmuir and Itcha-Ilgachuz subpopulations have been observed by Serrouya et al. (2012), Taylor et al. (2020), and Taylor et al. (2021). Based on microsatellites, the former study illustrated a division in clustering assignment of

the two subpopulations to both the Northern Mountain DU7 and Central Mountain DU8, though the authors acknowledged that sample sizes were small (*n*=2 and *n*=7 for Tweedsmuir and Itcha-Ilgachuz, respectively; Serrouya et al. 2012). In addition to a similar unique signal for Itcha-Ilgachuz individuals, Taylor et al. (2020) also discovered elevated inbreeding coefficient values for caribou from Itcha-Ilgachuz, which they suggest may be a result of the subpopulation's geographic isolation. In my study, a similar pattern was reflected in the lower value of observed heterozygosity for the inferred Itcha-Ilgachuz cluster (H_o=0.3231) relative to the second, larger cluster (H_o=0.3816).

Second, a genetic boundary exists between northern and southern caribou along the boundary of the Graham and Klinse-Za (Moberly) subpopulations. *Admixture*, *Tess*, and *Structure* all showed some level of support for both *K*=3 and *K*=4; for both clustering solutions, Itcha-Ilgachuz was retained as a distinct group, with the addition of a "Northern Mountain" vs "Southern Mountain" divide around the boundary between the Graham and Klinse-Za subpopulations. Serrouya et al. (2012) reported that genetic discontinuities in their analysis of western Canadian caribou were associated with two major river valleys: the North Thompson Valley and the Peace River, with the latter separating two distinct clusters along the Graham/Klinse-Za boundary (see also McLoughlin et al. 2004). The authors concluded that their inferred clusters were therefore inconsistent with ecotype designations and that observed patterns were likely a result of caribou generally preferring higher elevations for foraging as well as avoiding predation in lower elevation areas more commonly inhabited by deer, moose, and associated predators, irrespective of designated ecotype (Apps et al. 2001; Stotyn 2007; Serrouya et al. 2012). These patterns are replicated in this study specifically for Northern vs Southern Mountain caribou, where the Peace River may act as an isolation-by-resistance factor (IBR; or the resistance to gene flow caused by landscape heterogeneity and/or environmental variables) for caribou on either side of the Graham/Klinse-Za boundary, or may be indicative of a longer ancestral barrier as a result of the Peace River Valley's creation post glacial retreat.

Lastly, the genetic boundary between northern and boreal caribou appears to be along the Muskwa/Pink Mountain and Gataga/Finlay boundaries rather than the currently recognized Muskwa/Westside Fort Nelson boundary. Although Serrouya et al. (2012)'s above-mentioned study on the potential association of genetic discontinuities with major river valleys suggests that the Peace River may play a role in differentiating northern and boreal caribou ecotypes, the river's influence on the differentiation between Northern Mountain and Boreal caribou was not supported here. Instead, in this situation, high rates of genetic drift among populations/subpopulations may be masking any preconceived contributions of geographic features, such as major roads or river valleys, or even historic taxonomic boundaries to population structure (Serrouya et al. 2012; Polfus et al. 2017; Mager et al. 2014). Groups of individuals that display a continuum of genotypes do not necessarily fit the discontinuous species-based conservation model that is assumed in the process of creating management units (Polfus et al. 2016; Fitzpatrick et al. 2015); such hybrids are instead often recognized and classified according to their local ecological function (Stronen et al. 2022). Therefore, a clustering solution that recognizes a genetic division between northern and boreal caribou also respects both existing and unexplored ecological differences between these two groups, including specializations to different habitat (e.g., boreal forests vs mountain ranges) and potential variations in behavior (e.g., migratory vs sedentary).

Out of the four clustering solutions presented in this study, this four-unit subdivision that recognizes boreal, northern, southern, and Itcha-Ilgachuz/Tweedsmuir clusters remained highly supported across several population inference programs. Unlike the *K*=6 solution, it did not overemphasize the differentiation between Central and Southern Mountain caribou, which was not well supported in this study. At *K*=6, the clustering solution best supported by DAPC results, a separate Tweedsmuir cluster became apparent, as well as a separation of "Central Mountain" and "Southern Mountain" individuals in clusters deduced by *Structure*, *Admixture*, and *Tess*. In the study conducted by

Serrouya et al. (2012) discussed above, the authors also discovered the North Thompson Valley to be a likely IBR factor in the differentiation between caribou clusters north and south of the Wells Gray subpopulation. The same study's post hoc migration analyses revealed that out of 48 sampled individuals belonging to herds south of the North Thompson Valley, including Columbia North, Groundhog, and Frisby-Boulder (previously Frisby-Queest; not sampled in this study), only one potential first-generation migrant originated from Wells Gray during an almost 20-year timespan. This reiterates the idea that the spatial organization of genetic diversity can be influenced by one of two phenomena: geographic isolation, or its opposite, dispersal and gene flow (Cavedon et al. 2022a; Clobert et al. 2001).

2.4.2. The case of ambiguous units

In the case of caribou in western Canada, I uncovered several "optimal" clustering solutions (*K*=2, *K*=3, *K*=4, *K*=6, and the possibility of *K*>10 with *Admixture* and *Tess*)—depending on the population inference software and approach to selecting optimal *K* used—that each warrant scrutiny and discussion despite sharing overlap in the differentiation patterns they assume. Caribou as a study system can make the delineation of population structure particularly challenging because it is not uncommon for caribou herds to act as either closed populations (Valkenburg et al. 2002), components of metapopulations (Skoog 1968; Hinkes et al. 2005), or otherwise join together or separate based on variations in behavior (e.g., migration, periodic shifts in calving grounds, cyclic population expansions and declines, habitat availability, etc.), depending on the region (Hinkes et al. 2005; Mager et al. 2014). Nonetheless, past studies and species assessments have confirmed that caribou genetic structure does not fully overlap with existing taxonomic designations (Serrouya et al. 2012; COSEWIC 2011), which has likewise been confirmed in this study, regardless of the value of *K* determined to be most optimal.

It is understandable and not unexpected for *K* to be arbitrary depending on the study system in question and its associated sampling scheme. Samples may be taken from one large continuous

population and therefore not exhibit apparent population genetic structure (Wright 1946), or otherwise exhibit significant transitional zones. Despite forming groups vaguely reminiscent of DU structure, the boundaries of the inferred clusters presented here were not strict; instead, significant transitional zones between most clusters were evident (**Figure 2.15**). There was a particularly obvious transitional zone in central BC—where the Northern, Central, and Southern Mountain DUs generally converge—for values of $K \geq 3$. Individuals belonging to the Wolverine and Chase subpopulations, in particular, seemed to be significantly admixed between the Northern and Central Mountain inferred clusters, despite both officially belonging to the Northern Mountain DU7. The Wells Gray subpopulation, on the other hand, seemed to be the transitional zone between the Central and Southern Mountain inferred clusters in the southeastern part of BC, which is in contrast to the currently acknowledged division between Central and Southern Mountain caribou around the Hart Ranges subpopulation/region. However, it is of note that I only sampled one Wells Gray individual and do not have a full representation of this transitional zone. Previous studies have acknowledged that the Wells Gray subpopulation is slightly distinct from herds both to the north and south of it, which may support its signal as an admixed—though not entirely unique—zone in this study (Serrouya et al. 2012).

In the northeastern part of the province, the boundary of the Boreal DU/ecotype/SARA Schedule 1 population (which has been the only uncontested and consistent population in caribou classification schemes in BC thus far) also warrants scrutiny. For values of $K \geq 4$, the Boreal cluster became apparent relative to other clusters, but with modifications from what is currently known about the population; namely, the Muskwa, Pink Mountain, and Graham herds (currently classified as Northern Mountain DU) grouped together with the Boreal herds with confident assignment probabilities ($q > 0.5$ for both *Structure* and *Admixture*). This brings into question the genetic distinctiveness between the Boreal and Northern caribou ecotypes, among others, though it is also crucial to be aware of the history of specific caribou herds within this region. In the late 1990s, for instance, a total of 32

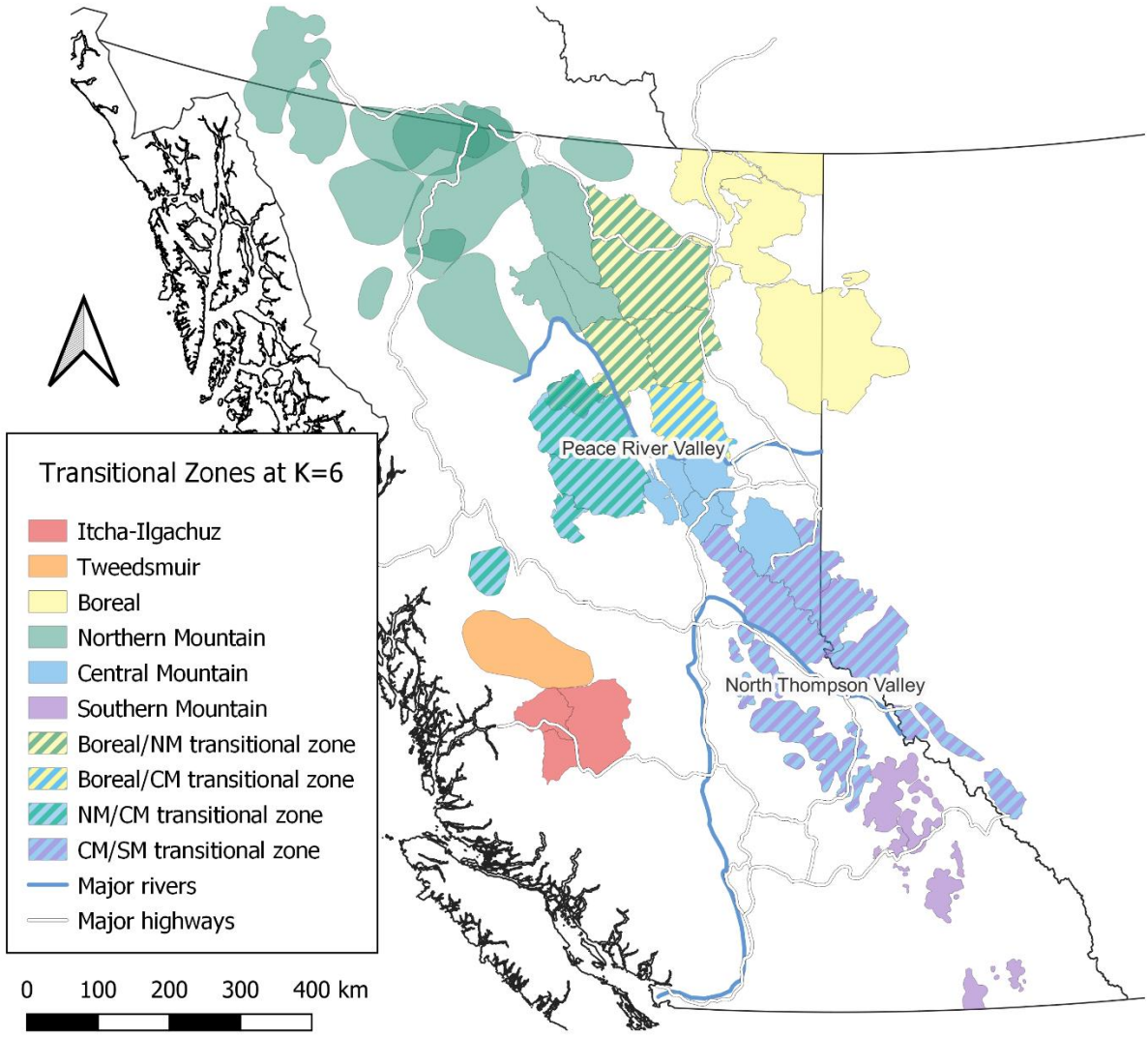

Figure 2.15. Transitional zones for caribou genetic subdivisions at *K*=6 using *Admixture* results. Zone names are based off original Designatable Unit (DU) classifications—Boreal, Northern Mountain (NM), Central Mountain (CM), and Southern Mountain (SM). Transitional zones were identified as any subpopulation having an average individual assignment probability (*q*) between 0.4 and 0.6. Map created in QGIS.

individuals were transplanted from the Chase subpopulation to the Telkwa subpopulation, which are both currently grouped within the Northern Mountain DU7 (Cichowski 2014; Houwers 2006; Stronen et al. 2007). In light of the results presented in this study, this would explain why Telkwa individuals are genetically more similar to other subpopulations belonging to the Northern/Central Mountain

transitional zone (e.g., Chase, Wolverine, Takla) than to the Tweedsmuir and Itcha-Ilgachuz subpopulations, despite closer geographic proximity.

Study systems such as the caribou studied here could also arguably have two or more optimal values of *K* that explain the presented population structure, especially when sampled individuals are taken from hierarchically structured populations where various levels of *K* correspond to the number of populations defined at different hierarchical levels (i.e., provinces, regions, subregions, etc.; Evanno et al. 2005; Wang 2017). Hierarchical clustering—creating subsets of individuals and reassessing population structure at smaller, finer scales—has been recommended for mobile species such as caribou to help reveal and better explain substructure at progressively finer scales; previous studies have successfully been able to detect populations within larger clusters in several species, including caribou belonging to the Peary and barren-ground subspecies (Jenkins et al. 2018) as well as Atlantic salmon (Vaha et al. 2007). A brief look into iterative clustering in this dataset revealed patterns consistent with the progressive values of *K* inferred from the presented mix of population structure programs, confirming the existence of fine-scale population structure that was already hinted at in nearly all the software used in this study.

Nonetheless, it can be arduous to disentangle hierarchical population structure into evolutionarily significant units, especially because managers and decision-makers often use species classification schemes that can be influenced by jurisdictional boundaries to manage species instead of biologically relevant units (O'Donnell et al. 2022). Population structure defined exclusively by genetic information may miss other important information regarding factors influencing processes such as habitat selection and gene flow, which can vary considerably at different temporal and spatial scales (Braunisch & Suchant 2010; Spear et al. 2010). As a result, detecting population structure within fine-scaled landscape patterns and continuously distributed species can be difficult because populations may

be organized into both clinal patterns and discrete clusters (Priadka et al. 2018; Stronen et al. 2022, Ryberg et al. 2013).

2.4.3. Conclusion

The results presented in this study suggest the existence of multiple genetically distinct populations and subpopulations in western Canadian caribou along with introgression-caused transitional zones between them. While genetic structure analyses provided evidence for both fine-scale and hierarchical population structure, I was able to deduce four primary new units that capture the evolutionary significance of BC and AB caribou more accurately than current caribou classification schemes. As human habitat modification continues to fragment habitats as well as propagate range expansions, contractions, or shifts, it is likely that differentiation but also possible hybridization and introgression between populations will become more apparent, which can in turn make it increasingly difficult to delineate conservation units. Further research will be needed to determine how to best conserve and manage caribou in this context, including at multiple hierarchical scales (e.g., DUs, groups of populations, or local population units), an approach that can be applied to any species that exhibits hierarchical population structure. Overall, this study is testament to the power of genomic data in helping better delineate conservation units and can be of great use to caribou conservation and management in western Canada.

CHAPTER 3: GENERAL CONCLUSION

In this thesis, I studied the population structure of caribou throughout western Canada. Results among population inference programs were varied but ultimately culminated to four most likely clustering solutions occurring at *K*=2, *K*=3, *K*=4, and *K*=6, in addition to showing signals of more fine-scale structure. The following sections contain supplemental considerations not discussed in the previous chapter and encompass some of the limitations and potential biases encountered in the midst of analyses. Recommendations for future work and general reflections of the project are also discussed.

3.1. Discussion of population structure inference approaches

In cases where population structure estimation is ambiguous, such as that exemplified here—but certainly not limited to caribou—it is important to understand and integrate results among the multiple performed analyses and then assess them in the context of additional non-genetic (ecological, behavioral, etc.) information regarding the species. Studies oftentimes apply *Structure* by default, owing to the software's pervasiveness and popularity as one of the original population structure programs. Nonetheless, despite its versatility in being able to handle different genetic markers and their respective characteristics and assumptions, one of *Structure*'s most persistent challenges has been inferring the "correct" number of source populations. *Structure*'s creators have themselves acknowledged the problems associated with cluster inference and interpretation, stating that the software should be used as an "exploratory tool" and that a given range of *K* values should be thoroughly inspected (Novembre 2016). Several studies have commented on the difficulty in determining, for instance, the point of plateau for *Structure*'s Ln Pr($X|K$) plot (a question that is transferable to i.e., *Admixture* and *Tess*'s cross-validation plots as well; Evanno et al. 2005; Falush et al. 2003; Francois & Durand 2009; Latch et al. 2006). High rates of gene flow among populations, the presence of close relatives, small numbers of

individuals or loci, and study design have been cited as potential factors for this, in addition to unaddressed, uneven sample sizes (Waples & Gagiotti 2006; vonHoldt et al. 2010; Puechmaille 2016). In the context of this study, small numbers of individuals or loci and the presence of close relatives can most likely be ruled out because they are either inapplicable (this study deals with a large number of both SNPs and sampled individuals) or accounted for (close relatives were filtered out prior to population structure analyses).

Similar questions of ambiguous clustering also exist in the context of spatial population genetic structure programs, such as *Tess*. Several studies have reported *Tess* having a tendency to infer a greater number of populations than estimates inferred from other spatial genetics clustering software (Gauffre et al. 2008; Guillot 2009). There are also warnings, however, of *Tess* inferring spatial patterns arising from a spatial model interacting with a poor statistical inference method, especially when relying on default program parameters. Critics have claimed that *Tess* is highly sensitive to the provided value of ψ, the interaction parameter, and that its default value of 0.6 may lead to overestimation of *K*, high error rate in assignments, and inference of spurious populations (Guillot 2009). These factors are worth considering in the context of both population structure results presented here and future work involving genetic structuring supplemented by geographic information for any species.

3.2. Integrating ecological and genetic information

Though there is a lot of discourse surrounding the issue of obtaining an optimal *K*, some authors stress that the seemingly unbreakable focus and attention being placed on *K* may not be entirely justified (e.g., Novembre 2016). Janes et al. (2017), for instance, recommends an approach where rather than focusing on the interpretation of clusters as isolated, panmictic "populations," more emphasis should be placed on the concept of genetic subdivision. Applying and integrating knowledge related to population genetic structure obtained from *Structure* and other programs is impossible without equal

consideration of the species' biology, as well as an integration of complementary genetic analyses, natural history, and Indigenous Knowledge (Jombart & Collins 2015; Raj et al. 2014; Verity & Nichols 2016).

As discussed in previous sections, studies that have looked at caribou population structure in other parts of the country have concluded that caribou ecotypes do not ideally match apparent genetic differences (Serrouya et al. 2012; Yannic et al. 2016). Across the species' North American range, caribou ecotypes have often been represented by multiple genetic entities rather than correspondence to one specific genetic group (Boulet et al. 2007; Cronin et al. 2005; Serrouya et al. 2012; Weckworth et al. 2012). This contradiction in caribou group delineation highlights the importance of differentiating groups using both genetic and ecological criteria (Cronin et al. 2005). Yannic et al. (2016) claim that for caribou, this translates to genetic clusters being defined using genetic criteria, as expected, but herds and ecotypes found within these clusters being defined using complementary ecological criteria to aid species management. The justification and assessment of evolutionarily significant units or DUs should therefore consider both lines of evidence while still acknowledging the need to manage species and populations at different levels (i.e., province, territories, regions, etc.) representative of a species' hierarchical structure. In the context of results generated by my study, there is pressing need to characterize (any) ecological differences between the Itcha-Ilgachuz and surrounding subpopulations compared to other caribou in western Canada if these are to be designated as a new DU, as suggested herein (Crandall et al. 2000; Fraser & Bernatchez 2001).

3.2.1. The balance of gene flow, dispersal, and geographic isolation

The question of whether to either limit or preserve gene flow between clusters or DUs is one that requires careful consideration, since unique evolutionary lineages may be characterized by both known and unknown rare genetic variants and require special protection as part of an ESU (Stronen et

al. 2022; Crandall et al. 2000). This has been the case for the Spirit bear, a white-coated color morph of the black bear (*Ursus americanus*) that inhabits the Canadian north Pacific Coast and is associated with a dietary niche that relies on marine resources more greatly than its black color morph counterpart (Service et al. 2020; Reimchen & Klinka 2017). It is therefore suspected that ecological segregation plays a large role in preserving such polymorphisms, implying that a broad temporal perspective and adaptive management are required to conserve species with ecologically relevant traits affected by, as of yet, unknown genes or genomic regions (Stronen et al. 2022).

Other studies have discussed the difficulty associated with accurately characterizing genetic structure when such ecological segregation, also known as isolation by distance (IBD), and genetic clines occur (Rowe & Beebee 2007; Guillot et al. 2008; Frantz et al. 2009). The assumed risk here is that underlying IBD patterns may lead to overestimations of genetic structure (Ball et al. 2010). Disentangling the effects of both IBD and IBR is crucial to assessing population-level delineation and identifying the potential factors contributing to population structure (Guillot et al. 2009; Priadka et al. 2018). Signals of genetic differentiation can also be the result of both historic (dating back to the LGM) and contemporary events (Thompson et al. 2019). In other species found in BC and known for their large dispersing movements, such as wolves, strong differentiation has been attributed to evolutionary adaptations to different ecological conditions (Stronen et al. 2014). Such genotype-environment associations have thus far not been greatly explored in the context of BC caribou but can be one solution to integrating genetic and ecological data; only a handful of studies have looked at whether patterns in observed genetic structure parallel other ecological, behavioral, or life history traits, which is a severe detriment to the continued management of the species (Weckworth et al. 2012, Cavedon et al. 2019; Cavedon et al. 2022a).

While high rates of genetic drift in small populations may override the influence taxonomic boundaries or landscape features have on population structure, population size may also influence

population dynamics through its effects on behavior (Mager et al. 2014). Caribou herds have been reported to modify their ranges according to fluctuations in census size, which could subsequently alter their geographic proximity to other herds and the probability of genetic exchange (Hinkes et al. 2005). This is no surprise given the theory of genetic drift (Wright 1931), so the nonlinear relationship between population size and population structure is worth better exploring in future studies, especially with the support of genomic data. Caribou as a species have an especially interesting threshold in this context: below a census population size of approximately 150 individuals, the magnitude and variation of differentiation significantly increases between adjacent herds (Serrouya et al. 2012).

For caribou found in western Canada and Alaska, both geographic barriers and population size have been the factors most correlated with genetic differentiation and variation (Mager et al. 2014; Serrouya et al. 2012). Moreover, studies have suggested that genetic drift in herds with small effective population sizes and reduced gene flow may explain the existence of small and genetically discrete herds within otherwise contiguous and/or interconnected populations, which may be the case for the unique genetic signals of the Itcha-Ilgachuz and Tweedsmuir herds presented here. Similar relationships have been reported in grizzly bears (*Ursus arctos*) in the northwestern part of the United States and western Canada, where population differentiation was largely driven by both population size and degree of isolation (Proctor et al. 2012). An increased focus on effective population size as an explanatory factor in future analyses and discussions of caribou population structure in western Canada would therefore also be valuable.

3.2.2. The effect of translocations on present and future population structure

One of the most widely discussed recovery tactics for at-risk populations, particularly those with low or declining effective population sizes, has been translocation—in other words, the artificial movement or relocation of individuals from one population to another to bolster struggling but

otherwise extant populations (Armstrong & Seddon 2008). Despite the hypothesis that anthropogenic land use is the biggest driver of caribou population declines throughout the region, protected areas alone may not guarantee species conservation and survival, as has been evidenced by the already high rates of incidence of extirpated populations (i.e., Banff National Park; Hebblewhite et al. 2010; Decesare et al. 2010). Some authors claim that translocation may be the best and most aggressive approach to caribou conservation, especially in the face of a lack of societal support for predator control or other conservation and protection measures (Garrot et al. 1993; Bruskotter et al. 2009; Decesare et al. 2010). While translocations have been associated with increased means and reduced variances in population growth rates in several species of ungulates (Komers & Curman 2000; Van Houtan et al. 2010), they have also been reported to have mixed success, since many translocation programs do not distinguish between initial population establishment and its long-term persistence (Armstrong & Seddon 2008; Decesare et al. 2010). Moreover, it may be difficult to successfully translocate caribou and other species that exhibit locally adapted behaviors, which is where it becomes valuable to consider both ecological and genetic information to best target populations for translocation.

To date, caribou in BC have been subject to several translocation events. In 2019, the South Selkirks and Purcells South subpopulations (both belonging to the Southern Mountain DU9) were also functionally extirpated, so the remaining five individuals were translocated to the Columbia North subpopulation using a soft-release method by means of the Revelstoke maternity pen (Mathieu et al. 2022). While my project's sampling design and the recent timeline of this translocation may not have made it possible to detect associated effects on regional population structure, it is worth nothing that both the source (South Selkirks and Purcells South) and sink (Columbia North) subpopulations clustered into one "Southern Mountain" population with limited differentiation in presented analyses. Given the recommendation that conservationists ought to capture ≥95% of the source population's genetic diversity during translocations to limit any possible bottleneck effects caused by the process (Weeks et

al. 2011), this instance of caribou translocation seems justified. Still, other translocation studies—e.g., in Seychelles warbler—have found evidence of significant, albeit low to moderate, genetic differentiation between populations involved in translocation (Wright et al. 2014). Specifically, populations established with fewer founders were observed to cluster separately, which alludes to the existence of possible transitional zones, such as those presented here, and calls for further investigation into how translocations affect neutral and functional genetic diversity as well as population structure.

3.3. Future research

As is the case with most research projects, multiple questions remain, and new ones arise. As mentioned in previous sections, future studies of caribou in western Canada could particularly explore: (1) ecological differences between the Itcha-Ilgachuz and surrounding subpopulations compared to other caribou in western Canada, (2) how Indigenous Knowledge supports, amends, or otherwise disputes observed caribou genetic structure, (3) both historic and contemporary effective population sizes and their relationship to population genetic diversity and structure, (4) the possible presence of genotype-environment associations in various caribou groups (i.e., populations or ecotypes), (5) gene flow analyses relying on genomic data (such as *TreeMix* or *SpaceMix* analyses, which could indicate the direction of admixture), (6) additional divergence analyses between inferred and existing caribou classification schemes (e.g., using *Bayescan* or *PCAdapt*), (7) the presence and importance of inbreeding depression, as well as many other directions. Ultimately, to ensure the survival and evolutionary potential of a species such as caribou, the threats posed by environmental changes must be mitigated, but the availability of genetic variation within well-defined subspecies, ecotypes, and populations ought to remain a top conservation priority; it is therefore my hope that this project, as well as those to come, will positively contribute to caribou recovery and persistence.

REFERENCES

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*(1), 246. https://doi.org/10.1186/1471-2105-12-246

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: Allozymes to genomes. *Molecular Ecology*, *26*(2), 420–430. https://doi.org/10.1111/mec.13948

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Álvarez, I., Fernández, I., Traoré, A., Menéndez-Arias, N. A., & Goyache, F. (2021). Population Structure Assessed Using Microsatellite and SNP Data: An Empirical Comparison in West African Cattle. *Animals*, *11*(1), Article 1. https://doi.org/10.3390/ani11010151

Apps, C. D., McLellan, B. N., Kinley, T. A., & Flaa, J. P. (2001). Scale-Dependent Habitat Selection by Mountain Caribou, Columbia Mountains, British Columbia. *The Journal of Wildlife Management*, *65*(1), 65. https://doi.org/10.2307/3803278

Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11. https://doi.org/10.1111/1755-0998.12559

Armstrong, D. P., & Seddon, P. J. (2008). Directions in reintroduction biology. *Trends in Ecology & Evolution*, *23*(1), 20–25. https://doi.org/10.1016/j.tree.2007.10.003

Ball, M. C., Finnegan, L., Manseau, M., & Wilson, P. (2010). Integrating multiple analytical approaches to spatially delineate and characterize genetic population structure: An application to boreal caribou (Rangifer tarandus caribou) in central Canada. *Conservation Genetics*, *11*(6), 2131–2143. https://doi.org/10.1007/s10592-010-0099-3

Banfield, A. W. F. (1961). *A revision of the reindeer and caribou, genus Rangifer*.

Beaumont, M. A. & Rannala, B. (2004). The Bayesian revolution in genetics. *Nature Reviews Genetics*, 5(4), 251-261.

Bergerud, A. T. (1996). Evolving perspectives on caribou population dynamics, have we got it right yet? *Rangifer*, 95–116. https://doi.org/10.7557/2.16.4.1225

Bergerud, A. T. (2000). Caribou. In: *Demerais S, Krausmann P (Eds) Ecology and management of large mammals in North America*. (eds S. Demarais & P.R. Krausman). Prentice Hall, New Jersey, 658–693.

Bergerud, A.T., Luttich, S.N., & Camps, L. (2008). *The Return of Caribou to Ungava*. McGill-Queen's University Press, Montreal and Kingston.

Boulet, M., Couturier, S., Côté, S. D., Otto, R. D., & Bernatchez, L. (2007). Integrative use of spatial, genetic, and demographic analyses for investigating genetic connectivity between migratory, montane, and sedentary caribou herds. *Molecular Ecology*, *16*(20), 4223–4240. https://doi.org/10.1111/j.1365-294X.2007.03476.x

Brathen, K. A., Ims, R. A., Yoccoz, N. G., Fauchald, P., Tveraa, T., & Hausner, V. H. (2007). Induced shift in ecosystem productivity? Extensive scale effects of abundant large herbivores. *Ecosystems*, 10, 773-789.

Braunisch, V., & Suchant, R. (2010). Predicting species distributions based on incomplete survey data: The trade-off between precision and scale. *Ecography*, 33, 826–840. https://doi.org/10.1111/j.1600-0587.2009.05891.x

Bruskotter, J. T., Vaske, J. J., & Schmidt, R. H. (2009). Social and Cognitive Correlates of Utah Residents' Acceptance of the Lethal Control of Wolves. *Human Dimensions of Wildlife*, *14*(2), 119–132. https://doi.org/10.1080/10871200802712571

Calenge, C. (2006). The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling*, *197*, 1035.

Carrier, A., Prunier, J., Poisson, W., Trottier-Lavoie, M., Gilbert, I., Cavedon, M., Pokharel, K., Kantanen, J., Musiani, M., Côté, S. D., Albert, V., Taillon, J., Bourret, V., Droit, A., & Robert, C. (2022). Design and validation of a 63K genome-wide SNP-genotyping platform for caribou/reindeer (Rangifer tarandus). *BMC Genomics*, *23*(1), 687. https://doi.org/10.1186/s12864-022-08899-6

Cavedon, M., Gubili, C., Heppenheimer, E., vonHoldt, B., Mariani, S., Hebblewhite, M., Hegel, T., Hervieux, D., Serrouya, R., Steenweg, R., Weckworth, B. V., & Musiani, M. (2019). Genomics, environment and balancing selection in behaviourally bimodal populations: The caribou case. *Molecular Ecology*, *28*(8), 1946–1963. https://doi.org/10.1111/mec.15039

Cavedon, M., vonHoldt, B., Hebblewhite, M., Hegel, T., Heppenheimer, E., Hervieux, D., Mariani, S., Schwantje, H., Steenweg, R., Theoret, J., Watters, M., & Musiani, M. (2022a). Genomic legacy of migration in endangered caribou. *PLOS Genetics*, *18*(2), e1009974. https://doi.org/10.1371/journal.pgen.1009974

Cavedon, M., Poissant, J., vonHoldt, B., Michalak, A., Hegel, T., Heppenheimer, E., Hervieux, D., Neufeld, L., Polfus, J. L., Schwantje, H., Steenweg, R., & Musiani, M. (2022b). Population structure of threatened caribou in western Canada inferred from genome-wide SNP data. *Conservation Genetics*. https://doi.org/10.1007/s10592-022-01475-1

Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, *16*(2), 540–548. https://doi.org/10.1111/1755-0998.12471

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*, *114*(30), E6089–E6096. https://doi.org/10.1073/pnas.1704949114

Cichowski, D. (2014). *Telkwa Caribou Population Status and Background Information Summary 2014*. 110.

Clobert, J., Danchin, E., Dhondt, A.A. & Nichols, J.D. (2001). *Dispersal*. Oxford University Press, Oxford.

Corander, J., Waldmann, P. & Sillanpaa, M.J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics*, 163(1), 367-374.

COSEWIC. (2010). *COSEWIC Annual Report*. 88.

COSEWIC. (2011). *Designatable Units for Caribou (Rangifer tarandus) in Canada* (p. 88). Committee on the Status of Endangered Wildlife in Canada.

COSEWIC. (2014a). *COSEWIC assessment and status report on the caribou, Rangifer tarandus: Newfoundland population, Atlantic-Gaspésie population, boreal population, in Canada.*

COSEWIC. (2014b). *COSEWIC assessment and status report on the Caribou Rangifer tarandus, Northern Mountain population, Central Mountain population and Southern Mountain population in Canada* (p. 113). Committee on the Status of Endangered Wildlife in Canada.

Coulon, A., Fitzpatrick, J. W., Bowman, R., Stith, B. M., Makarewich, C. A., Stenzler, L. M., & Lovette, I. J. (2008). Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma cœrulescens). *Molecular Ecology*, *17*(7), 1685–1701. https://doi.org/10.1111/j.1365-294X.2008.03705.x

Crandall, K. A., Bininda-Emonds, O. R. P., Mace, G. M., & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, *15*(7), 290–295. https://doi.org/10.1016/S0169-5347(00)01876-0

Crandall, K. A. (2009). A multifaceted approach to species conservation. *Animal Conservation*, *12*(2), 105–106. https://doi.org/10.1111/j.1469-1795.2009.00254.x

Cronin, M. A., MacNeil, M. D., & Patton, J. C. (2005). Variation in Mitochondrial DNA and Microsatellite DNA in Caribou (Rangifer tarandus) in North America. *Journal of Mammalogy*, *86*(3), 495–505. https://doi.org/10.1644/1545-1542(2005)86[495:VIMDAM]2.0.CO;2

Crooks, K. R., & Sanjayan, M. (2006). Connectivity conservation: Maintaining connections for nature. In K. R. Crooks & M. Sanjayan (Eds.), *Connectivity Conservation* (1st ed., pp. 1–20). Cambridge University Press. https://doi.org/10.1017/CBO9780511754821.001

Cullingham, C. I., Miller, J. M., Peery, R. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., & Janes, J. K. (2020). Confidently identifying the correct K value using the ΔK method: When does K = 2? *Molecular Ecology*, *29*(5), 862–869. https://doi.org/10.1111/mec.15374

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*(5–6), 416–423. https://doi.org/10.1093/bfgp/elq031

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), Article 7. https://doi.org/10.1038/nrg3012

Decesare, N. J., Whittington, J., Hebblewhite, M., Robinson, H., Bradley, M., Neufeld, L., & Musiani, M. (2010). The Role of Translocation in Recovery of Woodland Caribou Populations. *Conservation Biology*, *25*(2): 365–373. https://doi.org/10.1111/j.1523-1739.2010.01609.x

Durand, E., Chen, C., & François, O. (2009). Comment on 'On the inference of spatial structure from population genetics data.' *Bioinformatics*, *25*(14), 1802–1804. https://doi.org/10.1093/bioinformatics/btp337

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Edwards, R. Y., Soos, J., & Ritcey, R. W. (1960). Quantitative Observations on Epidendric Lichens Used as Food by Caribou. *Ecology*, *41*(3), 425–431. https://doi.org/10.2307/1933317

Elhaik, E. (2022). Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-14395-4

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992). Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics*, 131(2), 479-491.

Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin ver. 3. 0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1, 47-50.

Excoffier, L. & Heckel, G. (2006). Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics*, 7(10), 745-758.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164, 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Festa-Bianchet, M., Ray, J. C., Boutin, S., Côté, S. D., & Gunn, A. (2011). Conservation of caribou (Rangifer tarandus) in Canada: An uncertain future. *Canadian Journal of Zoology*, *89*(5), 419–434. https://doi.org/10.1139/z11-025

Fitzpatrick, B. M., Ryan, M. E, Johnson, J. R., Corush, J., & Carter, E. T. (2015). Hybridization and the species problem in conservation. *Current Zoology, 61*:206-216. http://dx.doi.org/10.1093/czoolo/61.1.206

François, O., Ancelet, S., & Guillot, G. (2006). Bayesian Clustering Using Hidden Markov Random Fields in Spatial Population Genetics. *Genetics*, *174*(2), 805–816. https://doi.org/10.1534/genetics.106.059923

Francois, O., & Durand, E. (2010). Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources*, 10, 773784. https://doi.org/10.1111/j.1755-0998.2010.02868.x

Frankham, R. (1995). Conservation Genetics. *Annual Review of Genetics*, *29*(1), 305–327. https://doi.org/10.1146/annurev.ge.29.120195.001513

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2nd ed). Cambridge University Press.

Frantz, A. C., Cellina, S., Krier, A., Schley, L., & Burke, T. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: Clusters or isolation by distance? *Journal of Applied Ecology*, *46*(2), 493–505. https://doi.org/10.1111/j.1365-2664.2008.01606.x

Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology*, *10*(12), 2741–2752. https://doi.org/10.1046/j.0962-1083.2001.01411.x

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. https://doi.org/10.1016/j.tree.2012.05.012

Garrott, R. A., White, P. J., & Vanderbilt White, C. A. (1993). Overabundance: An Issue for Conservation Biologists? *Conservation Biology*, *7*(4), 946–949. https://doi.org/10.1046/j.1523-1739.1993.740946.x

Gauffre, B., Estoup, A., Bretagnolle, V., & Cosson, J. F. (2008). Spatial genetic structure of a small rodent in a heterogeneous landscape. *Molecular Ecology*, *17*(21), 4619–4629. https://doi.org/10.1111/j.1365-294X.2008.03950.x

Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J.-S., Moyers, B. T., Renaut, S., Rennison, D. J., Veen, T., & Vines, T. H. (2012). Recommendations for utilizing and reporting population genetic analyses: The reproducibility of genetic clustering using the program structure. *Molecular Ecology*, *21*(20), 4925–4930. https://doi.org/10.1111/j.1365-294X.2012.05754.x

Goossens, B., Sharma, R., Othman, N., Kun-Rodrigues, C., Sakong, R., Ancrenaz, M., Ambu, L. N., Jue, N. K., O'Neill, R. J., Bruford, M. W., & Chikhi, L. (2016). Habitat fragmentation and genetic diversity in natural populations of the Bornean elephant: Implications for conservation. *Biological Conservation*, *196*, 80–92. https://doi.org/10.1016/j.biocon.2016.02.008

Government of British Columbia. (2021). *Population Estimates for Caribou Herds of British Columbia*. BC Caribou Recovery Program.

Guillot, G. (2009). On the inference of spatial structure from population genetics data. *Bioinformatics*, *25*(14), 1796–1801. https://doi.org/10.1093/bioinformatics/btp267

Guillot, G., Santos, F., & Estoup, A. (2008). Analysing georeferenced population genetics data with Geneland: A new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, *24*(11), 1406–1407. https://doi.org/10.1093/bioinformatics/btn136

Guillot, G., Leblois, R., Coulon, A., & Frantz, A. C. (2009). Statistical methods in spatial genetics. *Molecular Ecology*, *18*(23), 4734–4756. https://doi.org/10.1111/j.1365-294X.2009.04410.x

Harding, L. E. (2022). Available names for Rangifer (Mammalia, Artiodactyla, Cervidae) species and subspecies. *ZooKeys*, *1119*, 117–151. https://doi.org/10.3897/zookeys.1119.80233

Heard, D. C., & Vagt, K. L. (1998). Caribou in British Columbia: A 1996 status report. *Rangifer*, 117–123. https://doi.org/10.7557/2.18.5.1548

Hebblewhite, M., White, C., & Musiani, M. (2010). Revisiting Extinction in National Parks: Mountain Caribou in Banff. *Conservation Biology*, *24*(1), 341–344. JSTOR.

Hinkes, M. T., Collins, G. H., Van Daele, L. J., Kovach, S. D., Aderman, A. R., Woolington, J. D., & Seavoy, R. J. (2005). Influence of Population Growth on Caribou Herd Identity, Calving Ground Fidelity, and Behavior. *The Journal of Wildlife Management*, *69*(3), 1147–1162. https://doi.org/10.2193/0022-541X(2005)069[1147:IOPGOC]2.0.CO;2

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2020). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62–82. https://doi.org/10.1111/mec.15720

Infinium Genotyping Data Analysis. (2014).

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*(14), 3594–3602. https://doi.org/10.1111/mec.14187

Jenkins, Deborah. A., Yannic, G., Schaefer, J. A., Conolly, J., & Lecomte, N. (2018). Population structure of caribou in an ice-bound archipelago. *Diversity and Distributions*, *24*(8), 1092–1108. https://doi.org/10.1111/ddi.12748

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403-1405.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94. https://doi.org/10.1186/1471-2156-11-94

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070-3071.

Jombart, T., & Collins, C. (2015). *A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.0.0*. 43.

Jump, A. S., & Peñuelas, J. (2005). Running to stand still: Adaptation and the response of plants to rapid climate change. *Ecology Letters*, *8*(9), 1010–1020. https://doi.org/10.1111/j.1461-0248.2005.00796.x

Kaeuffer, R., Reale, D., Coltman, D.W. & Pontier, D. (2007). Detecting population structure using STRUCTURE software: effect of background linkage disequilibrium. *Heredity*, 99(4), 374-380.

Komers, P. E., & Curman, G. P. (2000). The effect of demographic characteristics on the success of ungulate re-introductions. *Biological Conservation*, *93*(2), 187–193. https://doi.org/10.1016/S0006-3207(99)00141-X

Klütsch, C. F. C., Manseau, M., & Wilson, P. J. (2012). Phylogeographical Analysis of mtDNA Data Indicates Postglacial Expansion from Multiple Glacial Refugia in Woodland Caribou (Rangifer tarandus caribou). *PLOS ONE*, *7*(12), e52661. https://doi.org/10.1371/journal.pone.0052661

Klütsch, C. F. C., Manseau, M., Trim, V., Polfus, J., & Wilson, P. J. (n.d.). The eastern migratory caribou: The role of genetic introgression in ecotype evolution. *Royal Society Open Science*, *3*(2), 150469. https://doi.org/10.1098/rsos.150469

LaFramboise, T. (2009). Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances. *Nucleic Acids Research*, *37*(13), 4181–4193. https://doi.org/10.1093/nar/gkp552

Latch, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E. (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels

of population differentiation. *Conservation Genetics*, 7, 295–302. https://doi.org/10.1007/s10592-005-9098-1

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*(15), 3038–3051. https://doi.org/10.1111/j.1365-294X.2010.04688.x

Mager, K. H., Colson, K. E., Groves, P., & Hundertmark, K. J. (2014). Population structure over a broad spatial scale driven by nonanthropogenic factors in a wide-ranging migratory mammal, Alaskan caribou. *Molecular Ecology*, *23*(24), 6045–6057. https://doi.org/10.1111/mec.12999

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.

Mathieu, A., Vennen, L. M. V., Ried, A., Legebokow, C., & Schwantje, H. M. (2022). Translocation of southern mountain caribou using a soft-release technique. *Rangifer*, *42*(1), Article 1. https://doi.org/10.7557/2.42.1.5930

McDevitt, A. D., Mariani, S., Hebblewhite, M., Decesare, N. J., Morgantini, L., Seip, D., Weckworth, B. V., & Musiani, M. (2009). Survival in the Rockies of an endangered hybrid swarm from diverged caribou (Rangifer tarandus) lineages. *Molecular Ecology*, *18*(4), 665–679. https://doi.org/10.1111/j.1365-294X.2008.04050.x

McLoughlin, P. D., Paetkau, D., Duda, M., & Boutin, S. (2004). Genetic diversity and relatedness of boreal caribou populations in western Canada. *Biological Conservation*, *118*(5), 593–598. https://doi.org/10.1016/j.biocon.2003.10.008

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*(11), 2841–2847. https://doi.org/10.1111/mec.12350

Novembre, J. (2016). Pritchard, Stephens, and Donnelly on Population Structure. *Genetics*, 204(2), 391-393.

O'Donnell, M. S., Edmunds, D. R., Aldridge, C. L., Heinrichs, J. A., Monroe, A. P., Coates, P. S., Prochazka, B. G., Hanser, S. E., & Wiechman, L. A. (2022). Defining fine-scaled population structure among continuously distributed populations. *Methods in Ecology and Evolution*, *13*(10), 2222–2235. https://doi.org/10.1111/2041-210X.13949

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. (Kuke), & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics*, *26*(4), 177–187. https://doi.org/10.1016/j.tig.2010.01.001

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, *20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Polfus, J., Manseau, M., Simmons, D., Neyelle, M., Bayha, W., Andrew, F., Andrew, L., Klütsch, C., Rice, K., & Wilson, P. (2016). Łeghágots'enetę (learning together): The importance of indigenous perspectives in the identification of biological variation. *Ecology and Society*, *21*(2). https://doi.org/10.5751/ES-08284-210218

Polfus, J. L., Manseau, M., Klütsch, C. F. C., Simmons, D., & Wilson, P. J. (2017). Ancient diversification in glacial refugia leads to intraspecific diversity in a Holarctic mammal. *Journal of Biogeography*, *44*(2), 386–396. https://doi.org/10.1111/jbi.12918

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Genetics*, *4*. https://www.frontiersin.org/articles/10.3389/fgene.2013.00098

Priadka, P., Manseau, M., Trottier, T., Hervieux, D., Galpern, P., McLoughlin, P. D., & Wilson, P. J. (2018). Partitioning drivers of spatial genetic variation for a continuously distributed population of boreal caribou: Implications for management unit delineation. *Ecology and Evolution*, *9*(1), 141–153. https://doi.org/10.1002/ece3.4682

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Proctor, M. F., Paetkau, D., Mclellan, B. N., Stenhouse, G. B., Kendall, K. C., Mace, R. D., Kasworm, W. F., Servheen, C., Lausen, C. L., Gibeau, M. L., Wakkinen, W. L., Haroldson, M. A., Mowat, G., Apps, C. D., Ciarniello, L. M., Barclay, R. M. R., Boyce, M. S., Schwartz, C. C., & Strobeck, C. (2012). Population fragmentation and inter-ecosystem movements of grizzly bears in western Canada and the northern United States. *Wildlife Monographs*, *180*(1), 1–46. https://doi.org/10.1002/wmon.6

Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, *16*(3), 608–627. https://doi.org/10.1111/1755-0998.12512

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Raymond, M. & Rousset, F. (1995). Genepop (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248-249.

Reimchen, T. E., & Klinka, D. R. (2017). Niche differentiation between coat colour morphs in the Kermode bear (Ursidae) of coastal British Columbia. *Biological Journal of the Linnean Society*, *122*(2), 274–285. https://doi.org/10.1093/biolinnean/blx079

Rosenberg, N. A. (2003). distruct: A program for the graphical display of population structure: PROGRAM NOTE. *Molecular Ecology Notes*, *4*(1), 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rowe, G., & Beebee, T. J. C. (2007). Defining population boundaries: Use of three Bayesian approaches with microsatellite data from British natterjack toads (Bufo calamita). *Molecular Ecology*, *16*(4), 785–796. https://doi.org/10.1111/j.1365-294X.2006.03188.x

Ryberg, W. A., Hill, M. T., Painter, C. W., & Fitzgerald, L. A. (2013). Landscape Pattern Determines Neighborhood Size and Structure within a Lizard Population. *PLOS ONE*, *8*(2), e56856. https://doi.org/10.1371/journal.pone.0056856

Ryder, O. A. (1986). Species conservation and systematics: The dilemma of subspecies. *Trends in Ecology & Evolution*. https://doi.org/10.1016/0169-5347(86)90059-5

SARA. (2003a). *Caribou (Rangifer tarandus), Boreal population*. Species at Risk Act.

SARA. (2003b). *Woodland Caribou (Rangifer tarandus caribou), Southern Mountain population*. Species at Risk Act.

SARA. (2005). *Woodland Caribou (Rangifer tarandus caribou), Northern Mountain population*. Species at Risk Act.

SARA. (2022). *List of Wildlife Species at Risk (referral back to COSEWIC) Order*. Species at Risk Act.

Schaefer, J. A. (2006). Towards maturation of the population concept. *Oikos*, *112*(1), 236–240. https://doi.org/10.1111/j.0030-1299.2006.13763.x

Schluter, D., Clifford, E. A., Nemethy, M., & McKinnon, J. S. (2004). Parallel Evolution and Inheritance of Quantitative Traits. *The American Naturalist*, *163*(6), 809–822. https://doi.org/10.1086/383621

Schweizer, R. M., vonHoldt, B. M., Harrigan, R., Knowles, J. C., Musiani, M., Coltman, D., Novembre, J., & Wayne, R. K. (2016). Genetic subdivision and candidate genes under selection in North American grey wolves. *Molecular Ecology*, *25*(1), 380–402. https://doi.org/10.1111/mec.13364

Seip, D., & McLellan, B. 2008. Mountain caribou. *In* Caribou and the North: a shared future. *Edited by* M. Hummel and J. Ray. Dundurn Press Ltd., Toronto, Ont. pp. 240–255

Serrouya, R., Paetkau, D., McLellan, B. N., Boutin, S., Campbell, M., & Jenkins, D. A. (2012). Population size and major valleys explain microsatellite variation better than taxonomic units for caribou in western Canada. *Molecular Ecology*, *21*(11), 2588–2601. https://doi.org/10.1111/j.1365-294X.2012.05570.x

Service, C. N., Bourbonnais, M., Adams, M. S., Henson, L., Neasloss, D., Picard, C., Paquet, P. C., & Darimont, C. T. (2020). Spatial patterns and rarity of the white-phased 'Spirit bear' allele reveal gaps in habitat protection. *Ecological Solutions and Evidence*, *1*(2), e12014. https://doi.org/10.1002/2688-8319.12014

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalén, L., De Meester, L., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Höglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., … Zieliński, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*(2), 78–87. https://doi.org/10.1016/j.tree.2014.11.009

Skoog, R.O. (1968). Ecology of the caribou (Rangifer tarandus granti) in Alaska. PhD Dissertation, University of California, Berkeley.

Spear, S. F., Balkenhol, N., Fortin, M. J., McRae, B. H., & Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology*, 19,3576–3591. https://doi.org/10.1111/j.1365-294X.2010.04657.x

Stotyn, S. (2007). Ecological Interactions of Mountain Caribou, Wolves and Moose in the North Columbia Mountains, British Columbia M.Sc., Thesis, University of Alberta, Canada.

Stronen, A. V., Paquet, P., Herrero, S., Sharpe, S., & Waters, N. (2007). Translocation and Recovery Efforts for the Telkwa Caribou, Rangifer tarandus caribou, Herd in Westcentral British

Columbia, 1997-2005. *The Canadian Field-Naturalist*, *121*(2), Article 2. https://doi.org/10.22621/cfn.v121i2.440

Stronen, A. V., Navid, E. L., Quinn, M. S., Paquet, P. C., Bryan, H. M., & Darimont, C. T. (2014). Population genetic structure of gray wolves (Canis lupus) in a marine archipelago suggests island-mainland differentiation consistent with dietary niche. *BMC Ecology*, *14*(1), 11. https://doi.org/10.1186/1472-6785-14-11

Stronen, A. V., Norman, A. J., Vander Wal, E., & Paquet, P. C. (2022). The relevance of genetic structure in ecotype designation and conservation management. *Evolutionary Applications*, *15*(2), 185–202. https://doi.org/10.1111/eva.13339

Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome Biology*, *19*(1), 131. https://doi.org/10.1186/s13059-018-1520-3

Taylor, R. S., Manseau, M., Horn, R. L., Keobouasone, S., Golding, G. B., & Wilson, P. J. (2020). The role of introgression and ecotypic parallelism in delineating intraspecific conservation units. *Molecular Ecology*, *29*(15), 2793–2809. https://doi.org/10.1111/mec.15522

Taylor, R. S., Manseau, M., Klütsch, C. F. C., Polfus, J. L., Steedman, A., Hervieux, D., Kelly, A., Larter, N. C., Gamberg, M., Schwantje, H., & Wilson, P. J. (2021). Population dynamics of caribou shaped by glacial cycles before the last glacial maximum. *Molecular Ecology*, *n/a*(n/a). https://doi.org/10.1111/mec.16166

Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, *118*(10), e2015096118. https://doi.org/10.1073/pnas.2015096118

Thompson, L. M., Klütsch, C. F. C., Manseau, M., & Wilson, P. J. (2019). Spatial differences in genetic diversity and northward migration suggest genetic erosion along the boreal caribou southern

range limit and continued range retraction. *Ecology and Evolution*, *9*(12), 7030–7046. https://doi.org/10.1002/ece3.5269

Vähä, J.-P., Erkinaro, J., Niemelä, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology*, *16*(13), 2638–2654. https://doi.org/10.1111/j.1365-294X.2007.03329.x

Valkenburg, P., Keech, M. A., Sellers, R. A., Tobey, R. W., & Dale B. W. (2002). *Investigation of Regulating and Limiting Factors in the Delta Caribou Herd*. Final Research Technical Report, July 1996-June 2002. Project 3.42. Alaska Department of Fish and Game, Juneau, Alaska.

Van Houtan, K. S., Halley, J. M., Van Aarde, R., & Pimm, S. L. (2009). Achieving success with small, translocated mammal populations. *Conservation Letters*, *2*(6), 254–262. https://doi.org/10.1111/j.1755-263X.2009.00081.x

van Ravenzwaaij, D., Cassey, P. & Brown, S.D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin & Review*, 25(1), 143-154.

Verity, R., & Nichols, R. A. (2016). Estimating the Number of Subpopulations (K) in Structured Populations. *Genetics*, *203*(4), 1827–1839. https://doi.org/10.1534/genetics.115.180992

vonHoldt, B. M., Pollinger, J. P., Lohmueller, K. E., Han, E., Parker, H. G., Quignon, P., Degenhardt, J. D., Boyko, A. R., Earl, D. A., Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J. C., Mosher, D. S., Spady, T. C., Elkahloun, A., Geffen, E., Pilot, M., … Wayne, R. K. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*, *464*(7290), Article 7290. https://doi.org/10.1038/nature08837

Walters, A. D., & Schwartz, M. K. (2020). Population genomics for the management of wild vertebrate populations. *In: Hohenlohe P. A.; Rajora O. P., Eds. Population Genomics: Wildlife. Springer, Cham. p. 419-436.*, 419–436.

Wang, J. (2017). The computer program structure for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources*, *17*(5), 981–990. https://doi.org/10.1111/1755-0998.12650

Waples, R. S., & Gaggiotti, O. (2006). INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419–1439. https://doi.org/10.1111/j.1365-294X.2006.02890.x

Weckworth, B. V., Musiani, M., McDEVITT, A. D., Hebblewhite, M., & Mariani, S. (2012). Reconstruction of caribou evolutionary history in Western North America and its implications for conservation: CARIBOU GENETICS IN WESTERN NORTH AMERICA. *Molecular Ecology*, *21*(14), 3610–3624. https://doi.org/10.1111/j.1365-294X.2012.05621.x

Weckworth, B. V., Hebblewhite, M., Mariani, S., & Musiani, M. (2018). Lines on a map: Conservation units, meta-population dynamics, and recovery of woodland caribou in Canada. *Ecosphere*, *9*(7), e02323. https://doi.org/10.1002/ecs2.2323

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., Byrne, M., Coates, D. J., Eldridge, M. D. B., Sunnucks, P., Breed, M. F., James, E. A., & Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: A genetic perspective. *Evolutionary Applications*, *4*(6), 709–725. https://doi.org/10.1111/j.1752-4571.2011.00192.x

Wollstein, A. & Lao, O. (2015). Detecting individual ancestry in the human genome. *Investigative Genetics*, 6, 1-12.

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T., & Richardson, D. S. (2014). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, *23*(9), 2165–2177. https://doi.org/10.1111/mec.12740

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97–159.

Wright, S. (1946). Isolation by Distance under Diverse Systems of Mating. *Genetics*, *31*(1), 39–59.

Yang, W., Novembre, J., Eskin, E. & Halperin, E. (2012). A model-based approach for analysis of spatial structure in genetic data. *Nature Genetics*, 44(6), 725-731.

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., Dussault, C., Hundertmark, K. J., Irvine, R. J., Jenkins, D. A., Kolpashikov, L., Mager, K., Musiani, M., Parker, K. L., Røed, K. H., Sipko, T., Þórisson, S. G., Weckworth, B. V., Guisan, A., … Côté, S. D. (2013). Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change*, *4*(2), Article 2. https://doi.org/10.1038/nclimate2074

Yannic, G., St-Laurent, M.-H., Ortego, J., Taillon, J., Beauchemin, A., Bernatchez, L., Dussault, C., & Côté, S. D. (2016). Integrating ecological and genetic structure to define management units for caribou in Eastern Canada. *Conservation Genetics*, *17*(2), 437–453. https://doi.org/10.1007/s10592-015-0795-0

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Zittlau, K. A. (2004). Population genetic analyses of North American caribou (*Rangifer tarandus*). PhD Thesis, University of Alberta, Canada.