2019

Genomic Inference of Inbreeding in Alexander Archipelago Wolves (Canis lupus ligoni) on Prince of Wales Island, Southeast Alaska

Katherine Emily Zarn
University of Montana, Missoula

GENOMIC INFERENCE OF INBREEDING IN ALEXANDER ARCHIPELAGO WOLVES (*CANIS LUPUS LIGONI*) ON PRINCE OF WALES ISLAND, SOUTHEAST ALASKA

By

KATHERINE EMILY ZARN

Bachelor of Science in Biology, Gonzaga University, Spokane, Washington, 2011

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Wildlife Biology

The University of Montana
Missoula, MT

December 2019

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Michael K. Schwartz, Ph.D., Co-Chair
Wildlife Biology Program; USDA Forest Service Rocky Mountain Research Station

Gretchen H. Roffler, Ph.D., Co-Chair
Wildlife Biology Program; Alaska Department of Fish and Game

Marty Kardos, Ph.D.
Wildlife Biology Program; Flathead Lake Biological Station

Jeffrey M. Good, Ph.D.
Division of Biological Sciences; Wildlife Biology Program

Zarn, Katherine, M.S., Fall 2019 Wildlife Biology

Genomic Inference of Inbreeding in Alexander Archipelago Wolves (*Canis lupus ligoni*) on Prince of Wales Island, Southeast Alaska

Co-Chairperson: Michael K. Schwartz, Ph.D.

Co-Chairperson: Gretchen H. Roffler, Ph.D.

Abstract

Habitat loss and climate change are increasingly resulting in reduction and fragmentation of wildlife populations. Populations that have experienced fragmentation and decreases in abundance are at heightened risk of inbreeding due to reduced opportunities to mate with unrelated conspecifics. Prolonged or extensive inbreeding can result in inbreeding depression via the exposure of deleterious alleles in long runs of homozygosity. Alexander Archipelago wolves (*Canis lupus ligoni*) on Prince of Wales Island (POW) in Southeast Alaska are a small, isolated population of conservation concern that have experienced habitat loss and high harvest rates, and present an ideal system in which to explore questions about inbreeding. We present a custom hybridization capture that generates 37,082 genome-wide single nucleotide polymorphism (SNP) genotypes and over 500,000 SNP genotype likelihoods, allowing us to characterize and understand the extent of genomic inbreeding.

We found that F_{ROH} revealed different patterns of inbreeding across three populations of Alexander Archipelago wolves when compared to F_H. Estimates of F_H revealed no differences between the three populations, while F_{ROH} revealed that wolves in the southeast portion of the study area had the greatest total proportion of the genome in runs of homozygosity, including short runs from inbreeding in distant ancestors. Wolves on POW had more long runs of homozygosity, indicating more frequent mating between individuals with recent common ancestors, likely due to smaller recent historical population size on POW. We demonstrate the advantage of using genomics to assess inbreeding via F_{ROH} (versus F_H), which allows for inference about the timing and severity of inbreeding. Wolves on POW exhibit similar inbreeding patterns as wolves in Isle Royale National Park, a population that was founded by two to three individuals and has demonstrated severe inbreeding depression. This result is important for management (e.g. informing habitat use, population targets, and harvest quotas) and demonstrates the benefit of using high-resolution genomic data to infer individual inbreeding so proactive management can be applied to ensure the long-term sustainability of the populations.

Acknowledgements

I would not be who I am and where I am today without my family, who have made my life brighter, funnier, and more exciting and enjoyable. Thank you to: **David Zarn** for always fostering my curiosity by answering any and all questions I have been able to come up with to the best of your ability, for memorable and inspiring explorations and adventures, and for always supporting me however you could, no matter the distance. **Julie Zarn** for great hiking adventures, enjoying lots of delicious food together, and for encouragement, positivity, and support when I needed it most. **Justin Zarn** for being an inspirational, level-headed, loving, creative, hilarious, and dependable brother. **Rod Hanna** for our mutual appreciation and excitement for nature and wildlife, for much appreciated vet advice, and for always being up for exciting road trips and hikes in Washington, Wyoming, Montana, and Arizona. Thanks also to **Don and Donnie Van Brackel** for updates on fireflies, support, sound advice, and for welcoming me to the lake for relaxing and enjoyable visits.

A million thanks to **Matthew R. Jones** for fun hiking and backpacking adventures, never-ending patience, encouragement, understanding, support, and unconditional love through the highs and lows of life over the last 7 years, and especially through graduate school. You have made all the difference. Thanks also to **Cody** for making me laugh at hilarious faces and antics, putting your chin on my knee when I was feeling low, holding me accountable for daily walks through the woods, and for keeping the mailman from breaking into our home on a daily basis.

Life isn't complete without great friends. Thanks to **Alex Japhet** for being a steadfast, witty, insightful, and valued friend through thick and thin over the last twenty (really??) years. **Beth Fitzpatrick** for many fun months of working and living together, lots of laughs, and for always being there to pick up right where we left off like no time has passed at all. **Brooke Alcamo** for many hours spent trading stories and laughing, picking huckleberries, looking for morels and hiding from moose with Griffin, and for fantastic baking adventures. **Emiko Sano** for dog walks, introducing me to the most perfect chai latte I have ever had, and for many wonderful and funny conversations and lots of laughs. **Bobbi Ecklund** for being a wonderful and admirable person and friend and for showing me how great neighbors make a house feel so much more like home.

I have had the privilege and honor of working with and getting to know a great number of inspirational and dedicated wildlife conservation professionals at the USFS National Genomics Center. Thank you to all of the members of the NGC, and especially to **Kellie Carim, Todd Cross**, **Taylor Wilcox**, **Kelton Verble**, and **Julie Weckworth** for kindness, encouragement, valuable feedback on presentations and manuscripts, and exciting discussions about science over the last several years. Thanks also to **Kristy Pilgrim** and **Roberta Steele** for providing valued support both in the lab and in the office. **Kris Larson** at the Alaska Department of Fish and Game for letting me tag along in fieldwork on POW and for sharing tons of fantastic photos of and anecdotes about wolves and wildlife in Southeast Alaska..

This work would not have been possible without my outstanding committee. Thanks to **Mike Schwartz** for bestowing valuable career and life lessons, and for productive and engaging discussions regarding wildlife conservation science and how to provide the highest quality data for informed management decisions. **Gretchen Roffler** for facilitating travel and participation in field work, for valuable discussions about the status and background of wolves in Southeast

Alaska, and for being an admirable and encouraging advisor. **Marty Kardos** for always being approachable, encouraging, and thoughtful, and for checking in often and proactively offering much-appreciated support and guidance. **Jeff Good** for consistently welcoming my questions with a smile, sharing a wealth of knowledge about molecular genomics techniques, and for making it possible for me to be productive and able to pay rent and buy groceries during the longest government furlough in U.S. history.

Lastly, thank you to the people who have played instrumental roles in my scientific career so far: **Joey Haydock**, **Walt Koenig**, and **Eric Walters** for providing me with my first job opportunities in the lab and in the field, which helped me discover my love for genetics, field work, Acorn Woodpeckers, and wildlife. **Lila Fishman** for being so inspirational in demonstrating curiosity and enthusiasm for science. **Colin Callahan** for sharing expertise and demonstrating lots of patience in answering so many of my questions about library preps and captures many years ago, and for teaching me that listening to heavy metal in the lab isn't nearly as terrible as I once feared. **David Xing** for being a shining example of someone who is professional, dependable, proficient, easy to work with, and productive, and for help in continuing my thesis work during the government furlough. **John McCutcheon** for teaching valuable programming skills and techniques and opening my eyes to a new and exciting language of scientific analysis. Last, but not least, thanks to **Dave Patterson** for the best statistics classes I have taken and for teaching skills that I will carry with me throughout my career.

Thesis Format

Chapter 2 is written with the intent to publish in a peer-reviewed journal, and as such I use “we” throughout this thesis to recognize the contributions by my coauthors.

TABLE OF CONTENTS

CHAPTER 1: Introduction & Overview

Wildlife Conservation Genetics & Genomics

Habitat loss and climate change are causing wildlife populations around the world to decline at record rates. To ensure long-term persistence of wildlife species, it is essential to manage populations using biologically relevant management boundaries, especially in light of the evolutionary past and potential of populations and species of conservation concern (Allendorf et al. 2010; Funk et al. 2012; Supple and Shapiro 2018). A sound understanding of population demographic history (e.g. historic population expansions and bottlenecks and inbreeding), local adaptation, and population connectivity can facilitate effective and efficient use of conservation resources to ensure long term persistence of wildlife species. Efficient use of resources is especially imperative as anthropogenic landscape alterations rapidly reduce and alter the quantity and quality of available wildlife habitat.

Wildlife conservation and management efforts have traditionally been informed by genetics approaches to address pressing questions about populations of conservation concern. These approaches are typically characterized by the use of a small number (dozens) of coarsely spaced neutral genetic markers, which are suitable for addressing questions about population size and genetic structure, inbreeding, drift, and demographic vital rates. Genomic approaches, which have become more accessible to wildlife conservation research in recent years due to methodological improvements and decreased sequencing costs, utilize genome-wide sequence data to explore many more loci (hundreds to millions). With this increased number of loci, genomic methodologies offer a much higher resolution view into the genomes of wild animals. This facilitates a more complete characterization of population traits that were previously difficult or impossible to investigate including migration rates, local adaptation, population demographic history, genomic inbreeding, and inbreeding depression. Genomic data provide a more comprehensive understanding of these population characteristics, the forces that shape these characteristics, and how these populations can be effectively and efficiently managed to ensure sustainability into the future.

Inbreeding & Genomics

Inbreeding is particularly likely to occur in small populations, leading to increased homozygosity (Allendorf et al. 2013). Sustained or high levels of inbreeding may result in reduced fitness effects known as inbreeding depression, which may not be realized for many generations (Allendorf et al. 2013). One way inbreeding depression occurs is when deleterious partially recessive alleles become exposed at homozygous loci, and may manifest in increased susceptibility to disease, physical malformations, and decreased survival and reproduction, all of which may result in decreased fitness and population growth rates (Wright 1922; Allendorf et al. 2013; Hedrick et al. 2014). Inbreeding depression can arise in an extinction vortex, which describes a positive feedback loop of population decline, where the combination of demographic and environmental stochasticity, reduced genetic variation, and inbreeding depression results in an increased risk of extinction (Allendorf et al. 2013). Assessing and monitoring inbreeding in small, isolated

populations of conservation concern can provide valuable insight into a population's potential for exhibiting inbreeding depression and entering an extinction vortex.

Estimates of the inbreeding coefficient, F, seek to quantify the proportion of the genome that is identical by descent, where both copies of an allele in diploid organisms are descended from a single allele copy in a recent common ancestor. One traditional method for estimating F is to use a pedigree to calculate the probability that a diploid individual has two identical copies of an allele that descended from a single recent common ancestor. This estimate is called F_P (Templeton and Read 1994). Calculating and interpreting F_P for wild populations is difficult because it requires a pedigree, which is not often available for wildlife populations. Furthermore, a major assumption of F_P is that founding individuals in a pedigree are unrelated, which is seldom accurate, and results in F_P underestimating F (Kardos et al. 2015). Lastly, estimates of F_P can be inaccurate because true inbreeding coefficients vary greatly for individuals with the same pedigree (Kardos et al. 2015). Another traditional way to estimate inbreeding is by assessing the deficiency of heterozygous loci within an individual given expected heterozygosity within the subpopulation: $F_H = \frac{H_{exp} - H_{obs}}{H_{exp}}$ where H_{exp} is expected heterozygosity and H_{obs} is observed heterozygosity (Keller and Waller 2002). This method is flexible in that it can utilize a small number of coarsely spaced genetic markers (e.g. microsatellites) or genomic data. Calculations of F_H can, however, be biased by sampling schemes: if samples from two or more subpopulations are used to calculate F_H, expected heterozygosity will increase and result in artificially elevated F_H estimates. This can result in an incorrect inference that an individual is inbred, when in fact the inbreeding signal may simply be the result of population substructure, a phenomenon also known as the Wahlund effect (Wahlund 1928; Waples and Allendorf 2015).

One way to address the problems inherent in measures of F_P and F_H is to instead directly measure the proportion of the genome that is identical by descent by using genomic sequence data to identify runs of homozygosity throughout the genome. This measure is called F_{ROH}, and is the most direct and reliable way to calculate the proportion of the genome that is identical by descent (Keller et al. 2011). Beyond providing the most accurate measure of individual inbreeding, characterizing runs of homozygosity also allows for inference about the number of generations back to the common ancestor(s) of the parents of inbred offspring. Recent common parental ancestors produce long ROH because few recombination events have occurred since the time to the most recent common ancestor. More distant common parental ancestors produce shorter ROH because more recombination events have occurred in the time to the most recent common ancestor. Therefore, we are able to use the length distribution of ROH to infer how recently inbreeding may have occurred in an individual's lineage and demographic history. This greatly expands our capacity to understand relationships between environmental variables and changes in population dynamics like inbreeding.

Study System & Overarching Goal

Alexander Archipelago wolves (*Canis lupus ligoni*) in Southeast Alaska (Figure 1) are a population of conservation concern that has been petitioned for listing as threatened or endangered under the Endangered Species Act three times in the last two decades (Fish and Wildlife Service 1997; Fish and Wildlife Service 2010; Fish and Wildlife Service 2014). Additionally, the Tongass

National Forest's Land and Resource Management Plan identifies the Alexander Archipelago wolf as a management indicator species for which a long-term sustainable population is a priority (USDA Forest Service 2016). The species' genetic population structure and connectivity in this region has been explored and described previously in comparison to wolves from interior Alaska and Yukon Territory (Weckworth et al. 2005; Weckworth et al. 2010; Weckworth et al. 2011; Weckworth et al. 2015), and alongside wolves in British Columbia (Breed 2007), but these analyses were limited by a small number of coarsely-spaced genetic markers and, in the work by Weckworth et al. (2010, 2011, and 2015), a relatively small amount (611 nucleotides) of mitochondrial genome sequence data. Collectively, these previous data are not well-suited for addressing questions regarding demographic rates like movement and dispersal, or for precise measures of individual inbreeding in Alexander Archipelago wolves. Cronin et al. (2014) explored genome-wide single nucleotide polymorphism (SNP) variation in wolves throughout North America, including wolves in Southeast Alaska. However, this work utilized a commercially available SNP chip designed for domestic dog breed association and genetic disease studies (e.g. as performed by Tengvall 2015). Their conclusions were disputed (Fredrickson et al. 2015; Weckworth et al. 2015, but see Cronin et al. 2015) and questions remained about fine scale population genetic structure, connectivity, and inbreeding in Alexander Archipelago wolves in Southeast Alaska, and about isolation of and inbreeding in wolves on Prince of Wales Island in particular.

For effective management that ensures long-term viability of wolves in Southeast Alaska, it is essential to understand population demographic rates like inbreeding, movement, and dispersal. The overarching goals of this thesis were to (1) develop a SNP genotyping tool suitable for characterizing individual inbreeding in Alexander Archipelago wolves and, (2) as part of a broader research effort undertaken by the Alaska Department of Fish and Game, to evaluate whether wolves on Prince of Wales Islands are isolated and more inbred relative to wolves elsewhere in the region. Ultimately, these metrics can be used to inform management decisions (e.g. annual harvest quotas and old-growth logging) to ensure the long-term persistence of Alexander Archipelago wolves on Prince of Wales Island and in Southeast Alaska.

Although it is unknown when wolves first arrived on POW, it has been hypothesized that they initially colonized Southeast Alaska from a southern Pleistocene refugium coinciding with the retreat of glacial ice approximately 12,000 years ago (Nowak and Paradiso 1983; Nowak 1995). Over the last half century, old-growth forests on POW have been heavily logged (Albert and Schoen 2013), which has led to substantial habitat fragmentation and clear-cuts negatively impacting Sitka black-tailed deer (*Odocoileus hemionus sitkensis*), the wolves' main prey (Hanley 1984, Albert and Schoen 2013). The dense road network supporting these logging efforts has also allowed increased access for legal and illegal wolf harvest (Person et al. 1996; Person 2001). The Alaska Department of Fish and Game (ADF&G) has monitored wolf harvest rates in Game Management Unit 2, which is comprised of POW and the surrounding complex of nearby islands (total land area 9,025 km^2, sometimes referred to in previously published literature as GMU2, but hereafter collectively referred to as POW) from 1980 to present. During this time the annual harvest rate has varied from 7 to 131 wolves (Figure 2). ADF&G has also monitored wolf abundance in the POW island complex since 2012. During that time, the annual wolf population estimate has ranged from 89 to 231 individuals (Figure 3). Alexander Archipelago wolves were petitioned for emergency listing as threatened or endangered under the Endangered Species Act

(ESA) in 2015 due to declining population estimates, high legal and illegal annual harvest, and habitat loss from old growth logging (Toppenberg et al. 2015). This was the third listing consideration in two decades (Fish and Wildlife Service 1997; Fish and Wildlife Service 2010; Fish and Wildlife Service 2016). In response to the 2015 petition for listing, the U.S. Fish and Wildlife Service determined that the wolves were not warranted for listing under the ESA (Fish and Wildlife Service 2016), but there remains a strong need to understand genetic connectivity and inbreeding to ensure long-term persistence of the population.

Selection of the Genotyping Method

The genotyping method for this work addressed two main goals: (1) provide data sufficient to address smaller scale questions about individual inbreeding and runs of homozygosity for this thesis, requiring hundreds of thousands of single nucleotide polymorphisms, or SNPs, and (2) provide data sufficient to address larger scale questions about wolf population genetic structure and connectivity in Southeast Alaska as part of a larger research effort at ADF&G, requiring hundreds to thousands of SNPs. In addition to being suitable for processing 60 samples for this thesis, the method also needed to be compatible with and in budget for processing an additional 540 tissue, hair, and scat samples as part of the larger ADF&G population genetic structure study.

We considered several pre-existing methods for SNP discovery and genotyping, including commercially available canine SNP chips and restriction site associated DNA (RAD) sequencing approaches. However, these approaches had a variety of incompatibilities with our research needs including that they were unsuitable for use with non-invasive samples like hairs and scats, would incur costly SNP discovery steps, and/or would be likely to induce ascertainment bias, which could impact population genetic structure results. We ultimately decided that a custom-designed capture was the most appropriate approach to meet our research goals. For additional details on our genotyping method selection, refer to Appendix 1.

Characterizing genome-wide runs of homozygosity requires high resolution, genome-wide sequence data (hundreds of thousands of genome-wide SNPs) to accurately identify the start and stop points of chromosomal segments that are identical by descent. High resolution SNP data ensure that small runs of homozygosity are not entirely missed and that smaller, distinct runs of homozygosity are not erroneously concatenated into a larger, single run. Characterizing population genetic structure requires fewer (hundreds or thousands) putatively neutral SNPs to test hypotheses about gene flow, migration, and dispersal between populations. We sought to achieve a balance between a fine-scale genotyping effort at many loci for a higher per-sample cost versus a coarse-scale genotyping effort with fewer loci at a lower per-sample cost.

CHAPTER 2: Inbreeding in Wolves on Prince of Wales Island, Southeast Alaska

Introduction

Habitat loss and climate change are increasingly causing wildlife populations to become smaller, fragmented, and isolated. Additional pressures like hunting and trapping can exacerbate these pressures by further reducing population abundance, which can result in a cascade of negative outcomes (Keller and Waller 2002). One of the consequences of population decline and isolation is inbreeding, or mating between relatives. Understanding the degree to which inbreeding occurs in small and isolated populations can be crucial in detecting the negative effects of isolation, habitat loss, and other anthropogenic changes before the outcomes of these influences begin to manifest in ways that are difficult to correct (e.g. reduced reproduction, survival, and fitness, also known as inbreeding depression). Monitoring inbreeding allows for proactive management to reduce and potentially reverse negative pressures on populations and ensure the maintenance of sustainable populations.

Alexander Archipelago wolves (*Canis lupus ligoni*) on Prince of Wales Island (POW) in Southeast Alaska are an excellent case study for detecting and characterizing inbreeding in a small, geographically isolated, wild population. Wolves are highly social carnivores living in packs typically comprised of a single breeding male and female, and may also include non-breeding individuals (Mech and Boitani 2007). Wolves on Prince of Wales Island are isolated from the mainland by one long swim (~6.2 km) or at least five shorter swims (longest straight-line swim 2.7 km) through strong ocean currents. Though wolves on POW have been closely monitored since 2012 (Roffler et al. 2016; Roffler et al. 2019) and there is evidence that wolves in this region can swim up to 13 km (U.S. Fish and Wildlife Service), they have never been observed traveling between POW and mainland Southeast Alaska. In addition to geographic isolation, old-growth forests on POW have been heavily logged since the 1950s (Durbin 1990), resulting in decreased habitat for both wolves and their main prey, Sitka black-tailed deer (Hanley 1984, Albert and Schoen 2013). The wolf population on POW has also experienced heavy harvest rates in recent years. The population has been annually estimated at 89 to 231 individuals since 2012 (Figure 3), and annual harvest rates have ranged from 7 to 131 wolves since 1980 (Figure 2). Given the geographic isolation, habitat loss, and harvest pressure, we sought to understand whether wolves on POW were more inbred than wolves elsewhere in Southeast Alaska. We also included recently published data on inbreeding in a highly inbred population of wolves in Isle Royale National Park to serve as a reference point for observations of severe inbreeding depression.

To assess inbreeding in POW wolves, we required a method that was sensitive enough to detect and differentiate inbreeding that arose due to mating between individuals with more recent or more ancient common ancestors. This would allow us to understand whether wolves on POW have experienced low levels of inbreeding over a long period of time, or if they have experienced a recent increase in inbreeding, perhaps as the result of anthropogenic pressures on the population. One statistic traditionally used to estimate inbreeding is F_H, which is the deficiency of observed heterozygosity in an individual relative to expected heterozygosity given allele frequencies for a subpopulation. The F_H coefficient can be estimated using either low-resolution genetic markers (e.g. microsatellites) or higher resolution genomic markers (e.g. single-nucleotide polymorphisms

or SNPs). However, F_H estimates can be artificially inflated by erroneously combining samples from multiple subpopulations (Wahlund effect; Wahlund 1928; Waples and Allendorf 2015) and it is not suitable for inference of the timing of inbreeding events. High throughput sequencing approaches have greatly expanded the accuracy and precision of many population demographic measures, including estimates of individual inbreeding using runs of homozygosity (ROH; Kardos et al. 2016). Runs of homozygosity are stretches of the genome that are homozygous and identical by descent and serve as the genetic basis of inbreeding depression (Szpiech et al. 2013). Long ROH occur in the offspring produced by inbreeding between parents with recent common ancestors. These long ROH are broken down into shorter ROH in subsequent generations by recombination events during meiosis. The progressive shortening of haplotypes through generations by recombination allows for inference of the timing of inbreeding events using ROH lengths, because long ROH serve as evidence of inbreeding by individuals with more recent common ancestors, and short ROH indicate inbreeding by individuals with more ancient common ancestors.

We designed a hybridization-based, targeted DNA capture to generate a high-resolution SNP genotypes to infer genome-wide ROH in Alexander Archipelago wolves. We compared estimates of F_{IS} and F_{ROH} to demonstrate the advantages of using F_{ROH} to more precisely detect inbreeding and understand the timing of inbreeding events. Because of the versatility and flexibility of capture, our capture design is compatible with populations throughout the species range and will better inform a population's inbreeding status than traditional methods permit, facilitating proactive management to ensure the long-term persistence of small and isolated populations.

Materials & Methods

Study Area

Prince of Wales Island is located in Southeast Alaska (Figure 1) and is the fourth-largest island in the United States, covering 6,670 km^2 with 1,593 km of coastline characterized by numerous bays, inlets, and fjords. The POW island complex includes an additional eleven large islands and hundreds of small islands less than 1 km from the shoreline of POW. The island complex is geographically isolated from the mainland by the Clarence Strait (Figure 1), and wolves would need to swim up to ~3 km between multiple stepping-stone islands to travel between the mainland and POW. Prince of Wales Island is composed of lakes, streams, river valleys, and mountainous terrain up to 1,160 m in elevation. In addition to Alexander Archipelago wolves, POW also supports populations of other mammals, such as Sitka black-tailed deer, North American beaver (*Castor canadensis*), North American river otter (*Lontra canadensis*), American marten (*Martes americana*), and black bear (*Ursus americanus*). The vegetative communities on POW include old-growth conifer forests, post-clear-cut forest stands of uniform ages at varying successional stages, and, in less abundance, freshwater muskeg bogs and riparian and alpine areas. The region receives 130 to 400 cm of precipitation annually, primarily via rainfall, but winter snowfall may accumulate over 50 cm (Shanley et al. 2015). Extensive old-growth logging has resulted in a 94% reduction of contiguous high-volume forests between 1954 and 2004 on northern POW (Albert and Schoen 2013), and old-growth timber sales and logging throughout POW continues into the present day, though these actions are regularly challenged by litigation from environmental

conservation groups (e.g. SEACC, et al. v. U.S. Forest Serv., et al. 2019). To support logging efforts, more than 6,800 km of roads have been constructed on POW. Regions with highest road density have also experienced the highest wolf harvest relative to other locations on POW (Person and Russell 2008; Person and Logan 2012).

Sample Collection

Samples for this study were obtained opportunistically from a larger sampling effort undertaken by ADF&G to survey population genetic structure and connectivity of wolves in Southeast Alaska. Tissue samples were collected from 2002 through 2015 (Table 1) in accordance with guidelines established by ADF&G Animal Care and Use Committee (ACUC #2012–028 and #2014–15) and the American Society of Mammologists (Sikes et al. 2011) or by voluntary tissue sample submission to ADF&G from hunter harvest.

Capture Design & Optimization

We used the domestic dog nuclear genome (canFam3.1, Lindblad-Toh et al. 2005) as a reference to design a hybridization capture probe set to target 100,144 500-bp regions spaced ≥19,700 nucleotides apart. Hybridization-based capture will enrich target DNA that is up to 40% divergent (though enrichment efficiency begins to decline at 4-10% divergence; Paijmans et al. 2016), and because genome-wide pairwise sequence divergence between dogs and wolves is ~0.11% on average (Freedman et al. 2014), this approach is well-suited for targeted enrichment and genomic sequencing in wolves. To avoid challenges with mapping reads, we excluded repetitive sequence regions like microsatellites, nested repeats, short interspersed elements, long interspersed elements, and other repeating regions that we identified using RepeatMasker (Smit et al.), Tandem Repeats Finder (Benson 1999), and WindowMasker (Morgulis et al. 2006). These excluded regions were already masked in the sequence file for the domestic dog genome or available on UCSC's TableBrowser (https://genome.ucsc.edu/cgi-bin/hgTables). Our design also excluded exons (regions of the genome that code for proteins) and 100,000 base pairs (bp) flanking exons to avoid targeting SNPs under selection. We submitted our design to Roche and after their proprietary screening and optimization to ensure efficient capture of target regions, the resulting approved probe set (SeqCap EZ Choice Probes) comprised 131,816 targets covering a total of 43,564,811 base pairs in the reference genome. The increased number of targets relative to the original capture design reflects the division of some targeted regions into one or more smaller fragments during Roche's proprietary screening and optimization process.

To further improve the capture design, we prepared ten tissue DNA samples for sequencing following the library preparation and capture protocol detailed in *Library Preparation* below, then sequenced these samples on an Illumina MiSeq (v2 Nano 150 bp paired end reads) at the University of Montana Genomics Core (UMGC, Missoula, MT). Briefly, reads were processed through the removal of PCR duplicates step as detailed below in *Read Filtering, SNP Calling, & SNP Filtering*. After removing PCR duplicates, we used picard (v2.9.2, http://broadinstitute.github.io/picard/) CollectHsMetrics to assess fold-enrichment and mean on-target coverage. We then redesigned the capture, excluding all probes for which read coverage was three times or higher than the mean on-target coverage for all targets with above 0× coverage. High coverage in these excluded regions is likely the result of reads from multiple paralogs mapping to a single region of the reference

genome. These high coverage regions have lower mapping qualities and artificially inflated heterozygosity, so their removal from the capture design facilitated improved mapping quality scores, more accurate SNP genotype calls, and more accurate identification of ROH in subsequent analyses. We submitted the modified capture design to Roche and the final probe set comprised 136,542 targets covering a total of 45,985,696 bp in the reference genome.

Library Preparation

We prepared libraries according to Meyer and Kircher (2010) with the following modifications: we used a Covaris E220 to randomly fragment DNA to a mean size of 300 bp (125 µl sample volume in a microTUBE Snap-Cap AFA Fiber, 140 W peak incident power, 10% duty factor, 200 cycles per burst, 55 second treatment time). For every post-reaction solid phase reversible immobilization clean-up step, we used Serapure beads (Faircloth and Glenn 2011). We used unique indexing oligos on both the P7 and P5 ends for each sample to avoid any cross-sample contamination from tag-switching during post-pool PCR amplification steps.

Capture and Sequencing

We pooled 10 sample libraries (125 ng each) for each capture. For hybridization preparation, hybridization, recovery, and post-capture amplification, we followed the protocol outlined in chapters five through seven in SeqCap EZ HyperCap Workflow User's Guide Version 2.1 (Roche 2017), replacing the LM-PCR oligos with HPLC-purified primers IS5 and IS6 from Meyer and Kircher (2010). Thirty captured libraries were pooled in equimolar concentrations and the pooled library was diluted to 10 nM total concentration and sent for sequencing on a single lane of a HiSeq 4000 (paired end 150 bp reads) at the University of Oregon Genomics & Cell Characterization Core Facility (GC3F; Eugene, OR).

Read Filtering, SNP Calling, & SNP Filtering

Demultiplexed reads were downloaded from the GC3F to the UMGC Carnation server. We used Trimmomatic v0.33 (Bolger et al. 2014) to remove adapters and discard reads less than 100 bp in length to improve mapping quality scores. Adapter removal and read filtering was confirmed using FastQC v0.11.5 (Andrews et al. 2012). Unpaired reads were discarded to ensure high mapping quality scores and we used PEAR v0.9.6 (Zhang et al. 2014) to assemble overlapping reads. Assembled and unassembled paired reads were aligned to the domestic dog genome (canFam3.1, Lindblad-Toh et al. 2005) using bwa mem 0.7.15-r1140 (Li 2013). We used picard (http://broadinstitute.github.io/picard/) MergeSamFiles to merge SAM files for unassembled and assembled read pairs for each sample, MarkDuplicates to identify PCR duplicates, and samtools view (Li et al. 2009) to remove PCR duplicates. We then used picard CollectHsMetrics to assess fold-enrichment and mean on-target coverage. We used GATK v3.8 (McKenna et al. 2010; DePristo et al. 2011; Van der Auwera et al. 2013; version and citations apply to all subsequent mentions of GATK) to realign reads around indels and picard (http://broadinstitute.github.io/picard/) to correct read mate information after realignment around indels. We used GATK HaplotypeCaller to call raw SNPs for each individual within targeted regions and a 250 bp flanking region around each target. Following methods by Kardos et al. (2017), we used GATK GenotypeGVCFs to call raw SNPs using all samples, SelectVariants to

filter out indels, and bcftools view to filter SNPs by allele frequency and MQRankSum ('INFO/AF>0.01 & INFO/MQRankSum>-0.2') resulting in a reference dataset of known SNP sites. We used GATK BaseRecalibrator to recalibrate base quality scores using the aforementioned reference dataset and the realigned, mate-corrected BAM files. We called SNPs a final time using the recalibrated BAM files and GATK HaplotypeCaller and GenotypeGVCFs. Finally, we used vcftools (Danecek et al. 2011) to filter SNPs by minor allele frequency (--maf 0.05), minimum genotype quality (--minGQ 20), minimum mean coverage (--min-meanDP 10), missingness (--max-missing 0.85), Hardy-Weinberg proportions (--hwe 0.001), and to only retain biallelic sites (--min-alleles 2 --max-alleles 2) on autosomal chromosomes 1 through 38. We used vcftools --relatedness2 (Danecek et al. 2011) to screen for and remove duplicated samples.

Population Genetic Structure

To test for population genetic structure and identify populations within the sampled region, we used 37,082 SNPs and Admixture v1.3.0 (Alexander et al. 2009) to explore estimated individual ancestry based one to eight potential ancestral populations. To visualize proportion ancestry for each individual, we used R v3.5.3 (R Core Team 2019) and R Studio v1.1.463 (RStudio Team 2016) with packages marmap v1.0.3 (Pante and Simon-Bouhet 2013), tidyverse v1.2.1 (Wickham and RStudio 2017), and mapdata v2.3.0 (Becker et al. 2018). To visualize genetic differentiation among individuals, we generated a principle components analysis (PCA) plot using EIGENSOFT v6.1.4 (Price et al. 2006), R v3.5.3 (R Core Team 2019), and R Studio v1.1.463 (RStudio Team 2016) with package scales (Wickham and RStudio 2018).

We grouped individuals into three populations (best supported number of ancestral populations as determined by Admixture) using highest proportion ancestry from the Admixture analysis to test the extent of genetic structure among these groups and to determine whether samples should be split into subpopulations to ensure accurate identification of runs of homozygosity. The first group consisted of 15 individuals from Prince of Wales, Dall, Long, and Suemez Islands (GMU 2; hereafter referred to as the POW group). The second group consisted of 31 individuals total from Kuiu, Kupreanof, Mitkof, Prince of Wales (n=1), and Duke Islands and the mainland east of Lynn Canal (GMUs 1A, 1B, 1C, 2, 3; hereafter referred to as the southeast group). The third group consisted of 13 individuals from Pleasant Island, Spurt Cove, and the mainland west of Lynn Canal (GMUs 1C, 1D, 4Z, 5A; hereafter referred to as the northwest group; Figure 1, Appendix 2). To evaluate genetic differentiation between the three Admixture-identified populations, we estimated pairwise Weir and Cockerham weighted F_{ST} using vcftools --weir-fst-pop.

Heterozygosity & Individual Inbreeding Coefficients

We calculated heterozygosity for each individual as the proportion of heterozygous SNPs out of all SNPs genotyped for that individual. We calculated the individual inbreeding coefficient, F_H, for each individual within each Admixture-identified population using vcftools --het. We visualized estimates using R v3.5.3 (R Core Team 2019) and R package scales v1.0.0 R package (Wickham and RStudio 2018).

Genotype Likelihoods & Identification of Runs of Homozygosity

The number of SNPs (37,082) retained after quality filtering was insufficient to accurately identify genome-wide runs of homozygosity. To overcome this insufficient genome-wide SNP density we used ANGSD (Korneliussen et al. 2014) to calculate SNP genotype likelihoods from realigned sequence data. Genotype likelihoods incorporate potential sequencing errors along with prior data on allele frequencies and linkage to generate hundreds of thousands of genome-wide SNP genotype likelihoods from sequence data that would otherwise be discarded during filtering for minimum coverage when calling SNP genotypes (Nielsen et al. 2011). We used SNP genotype likelihoods to identify runs of homozygosity in R v3.5.3 (R Core Team 2019) and R Studio v1.1.463 (RStudio Team 2016) using an R script modified from Kardos et al. (2017b). Additional details about SNP genotype likelihoods can be found in Appendix 3.

We calculated F_{ROH} as the proportion of 38 autosomes in ROH for each individual. We calculated $F_{ROH \geq 100kb}$, $F_{ROH \geq 1Mb}$, and $F_{ROH \geq 10Mb}$, the proportion of the genome in runs of homozygosity longer than 100 kb, 1 Mb, and 10 Mb, respectively, for each individual by summing runs of homozygosity greater than each indicated length and dividing the sum by the total length of autosomes. Estimates for F_{ROH} and $F_{ROH \geq 1Mb}$ were visualized using the R package scales v1.0.0 (Wickham and RStudio 2018).

Results

Identification of Duplicated Tissue Sample

When evaluating relatedness results, we found that two sequenced tissue samples with a pairwise relatedness estimate of phi = 0.49 originated from a single wolf. For a monozygotic twin or duplicated sample, expected phi = 0.5 (Manichaikul et al. 2010). These tissue samples had different field identification numbers, but the same hunter certification and seal numbers, indicating that the two samples originated from a single individual. One sample from this individual had a mean on-target coverage of 36× and the other sample had a mean on-target coverage of 22×. We removed the lower coverage sample and reperformed all SNP calling and filtering steps on the 59 remaining tissue samples to ensure that sample duplication did not result in biased SNP calling and filtering.

SNP Filtering

In Table 2 we show the number of SNPs remaining after each filter step was applied *individually* to the initial dataset of 360,354 SNPs, then the number of SNPs remaining after all filters were applied together. This is to illustrate the magnitude of SNP removal that each filter had on an individual basis, which is masked when presenting the number of SNPs removed by each filter in sequential order. Filtering sites with <10× coverage resulted in the removal of 270,367 (75%) of SNPs from the original dataset. Filtering sites with data missing in nine or more individuals (≥15%) resulted in the removal of 164,818 (46%) of sites from the original dataset. Filtering sites with a minor allele frequency of 0.05 or lower resulted in the removal of 133,625 (37%) of SNPs from the original dataset. The remaining filters individually removed 22% or fewer SNPs from the

original dataset. When all filters were combined 323,272 or 90% of SNPs were removed for a final dataset of 37,082 SNPs (Table 2).

Population Genetic Structure

In our Admixture analysis of one through eight potential ancestral populations, we found that three ancestral populations were best supported (Figure 4, Appendix 4). Of the 16 wolves that we analyzed from POW and nearby islands (Dall, Warm Chuck, Long, and Suemez Islands), 15 had majority assignment to the same ancestral population (POW group) and one wolf had majority ancestry assignment to the southeast group (n= 31; Kuiu, Kupreanof, Mitkof, Prince of Wales, and Duke Islands and the mainland east of Lynn Canal). The northwest group consisted of 13 wolves from Pleasant Island and the mainland west of Lynn Canal. No wolves outside of POW had majority ancestry assignment to the POW group (Figures 4 and 5).

We used PCA to visualize and describe the first three principal components in 37,082 SNPs (Figure 6) for all 59 individuals. Reflecting the results observed in the Admixture analysis, samples grouped into three clusters, with 15 of 16 tissue samples collected on POW comprising one cluster (red circles in Figure 6A). The other two clusters were comprised of a mixture of samples from the remaining GMUs and one sample collected in the Yukon Territory (Figure 6). POW wolves were not distinct from other wolves on the first or third principle components axes, which explained 11.2% and 5.1% of the genetic variation observed in all samples, respectively (Figure 6A and 6B). POW wolves separated from other wolves on the second principle components axis, which explained 7.1% of the genetic variation observed in all samples (Figures 6A and 6C).

When comparing the POW group (Prince of Wales Island, Dall Island, Warm Chuck Island, Long Island, and Suemez Island) to the southeast group (Kuiu, Kupreanof, Mitkof, Prince of Wales, and Duke Islands and the mainland east of Lynn Canal), F_{ST} was 0.11. Between the POW group and the northwest groups (Pleasant Island and the mainland west of Lynn Canal), F_{ST} was 0.21. Between the southeast and northwest groups, F_{ST} was 0.14.

Heterozygosity & Individual Inbreeding Coefficients

For wolves in the POW group, mean heterozygosity was 0.32 (standard deviation (SD) = 0.04), southeast group mean heterozygosity was 0.30 (SD = 0.06), and northwest group mean heterozygosity was 0.39 (SD = 0.08). There was moderate evidence for a difference in heterozygosity between the POW and northwest groups and strong evidence for a difference between the southeast and northwest groups (one-way analysis of variance (ANOVA), df = 2, F = 10.217, $P = 1.65\times10^{-4}$; Figure 7A; Tukey multiple comparisons of means *P*-values in Appendix 5). The individuals with the lowest heterozygosity estimates were from Duke Island (heterozygosity = 0.14), Pennock Island (heterozygosity = 0.15), and Mitkof Island (heterozygosity = 0.21).

For wolves in the POW group, the mean F_H was 0.05 (SD = 0.12), the southeast group mean F_H was 0.08 (SD = 0.18), and the northwest group mean F_H was -0.04 (SD = 0.21, Figure 7B). There was very little evidence for a difference in F_H between any of the populations (one-way analysis of variance (ANOVA), df = 2, F = 2.182, $P = 0.12$). The two individuals with the highest F_H

estimates were from Pennock Island ($F_H = 0.55$) and Duke Island ($F_H = 0.57$, map showing these locations in Appendix 2). The two individuals with the lowest F_H estimates were from the northwest group on the mainland northwest of Lynn Canal (3, $F_H = -0.51$) and the mainland west of Lynn Canal ($F_H = -0.28$).

Genotype Likelihoods & Identification of Runs of Homozygosity

Each individual in the POW group (n = 15, Prince of Wales Island, Dall Island, Warm Chuck Island, Long Island, and Suemez Island) had 543,964 SNP genotype likelihoods, each individual in the southeast group (n = 31, Kuiu, Kupreanof, Mitkof, Prince of Wales, and Duke Islands and the mainland east of Lynn Canal) had 1,876,582 SNP genotype likelihoods, and each individual in the northwest group (n = 13, samples from Pleasant Island, Spurt Cove, and the mainland west of Lynn Canal) had 753,259 SNP genotype likelihoods.

We compared genomic inbreeding coefficients, F_{ROH}, across the three Southeast Alaska groups and a group of 11 Isle Royale National Park (IRNP) wolves (Robinson et al. 2019). There was very strong evidence of a difference in $F_{ROH \geq 100kb}$ between all four groups. The highest mean $F_{ROH \geq 100kb}$ was in the southeast group (mean = 0.60, range = 0.45 – 0.89), followed by the POW group (mean = 0.49, range = 0.33 – 0.61), followed by IRNP (mean = 0.36, range = 0.23 – 0.48), then the northwest group (mean = 0.27, range = 0.13 – 0.46; ANOVA df = 3, F = 45.99, $p = 3.66 \times 10^{-16}$; Figure 7C; Tukey multiple comparisons of means *P*-values in Appendix 5). The three wolves with the highest $F_{ROH \geq 100kb}$ estimates were from Kupreanof Island ($F_{ROH \geq 100kb} = 0.89$), Duke Island ($F_{ROH \geq 100kb} = 0.78$), and Pennock Island ($F_{ROH \geq 100kb} = 0.77$; map in Appendix 2).

The highest mean $F_{ROH \geq 1Mb}$ was in the POW group (mean = 0.44, range = 0.27 – 0.55), followed by the southeast group (mean = 0.32, range = 0.09 – 0.66), then the northwest group (mean = 0.18, range = 0.07 – 0.41; ANOVA df = 3, F = 11.03, $p = 5.81 \times 10^{-6}$; Figure 7D; Tukey multiple comparisons of means *p*-values in Appendix 5). When we evaluated $F_{ROH \geq 10Mb}$, indicative of the most recent inbreeding events, we found very strong evidence for the POW group and IRNP wolves having higher $F_{ROH \geq 10Mb}$ estimates than wolves in the southeast or northwest groups. The highest mean $F_{ROH \geq 10Mb}$ was in IRNP (mean = 0.18, range = 0.10 – 0.31), followed by the POW group (mean = 0.16, range = 0.01 – 0.28), followed by the southeast group (mean = 0.07, range = 0 – 0.27), then the northwest group (mean = 0.04, range = 0 – 0.13). There was strong evidence that $F_{ROH \geq 10Mb}$ is higher in IRNP and POW group wolves than in wolves from the southeast and northwest (ANOVA df = 3, F = 13.5, $p = 5.62 \times 10^{-7}$; Figure 7E; Tukey multiple comparisons of means *p*-values in Appendix 5). The three Alaska wolves with the highest $F_{ROH \geq 10Mb}$ estimates were from Dall Island (POW group; $F_{ROH \geq 10Mb} = 0.28$), Mitkof Island (southeast group; $F_{ROH \geq 10Mb} = 0.27$), and POW ($F_{ROH \geq 10Mb} = 0.26$), and only one IRNP wolf had a higher $F_{ROH \geq 10Mb}$ estimate than those individuals ($F_{ROH \geq 10Mb} = 0.31$).

Discussion

Method Development, Applications, & Limitations

Our findings offer an in-depth view of inbreeding by characterizing runs of homozygosity genome-wide in wolves. This method facilitates a much more precise and accurate measure of individual genomic inbreeding, via F_{ROH}, than traditional measures like heterozygosity and F_H. Estimates of F_{ROH} allow us to understand the timing of inbreeding events to better understand how ecological changes increase or decrease inbreeding in wild populations. Understanding the links between changes in habitat and increases or decreases in inbreeding events allows for proactive management (e.g. modify harvest quotas, management of critical habitat, etc.) to ensure long-term persistence of populations.

We present a novel capture approach capable of generating tens of thousands of SNP genotypes and hundreds of thousands of SNP genotype likelihoods to enable accurate and precise characterization of individual inbreeding in Alexander Archipelago wolves. This capture design is also compatible with wolves from other regions of the species' range. This approach was highly successful in generating data suitable for inferring inbreeding, but there are important limitations to consider before applying this method to additional sample types and research questions. This approach is not yet optimized for use with non-invasive samples like hairs and scats; however, capture has been successfully used to sequence nuclear genomes from ancient DNA samples (e.g. Burbano et al. 2010; Ávila-Arcos et al. 2011; Carpenter et al. 2013; Enk et al. 2014) and non-invasively collected DNA samples (Hernandez-Rodriguez et al. 2017; White et al. 2019). As currently designed, our capture targets only neutral SNPs and is not optimally designed to detect variable sites influenced by natural selection because it does not target exons. However, our work demonstrates the utility of the capture approach for addressing a wide variety of conservation questions, and it would be straightforward to add a panel of putatively adaptive loci or to design a new capture to target exon regions of interest (e.g. Jones and Good 2016).

Heterozygosity, Inbreeding, & Population Structure

We found differences in heterozygosity between the populations, with wolves in the POW group and the southeast group having lower heterozygosity than wolves in the northwest group. We found are no differences in F_H between the three populations that we analyzed, and that measures of $F_{ROH \geq 100kb}$, $F_{ROH \geq 1Mb}$, and $F_{ROH \geq 10Mb}$ revealed patterns of inbreeding that were undetected by measures of heterozygosity or F_H. The most striking result in our F_{ROH} analyses was that POW wolves have experienced very high levels of inbreeding in the recent past, and are comparable to a population of wolves on IRNP that was founded by just two to three individuals, despite the POW population being substantially larger.

We found no evidence of a difference in F_H between the three populations that we examined. This result was inconsistent with F_H estimates generated by Breed (2007), who showed that wolves from the southwest region of our study area (GMUs 1A and 1B) had highest inbreeding coefficients, followed by wolves on POW. As stated by Breed, the heightened F_H estimate for wolves in GMUs 1A and 1B in his study is likely due in part to population substructuring and not

inbreeding alone, a phenomenon known as the Wahlund effect (Breed 2007). However, Breed's F_H results are in concordance with our F_{ROH} results, as discussed below.

When we examined F_{ROH} for varying minimum ROH lengths, there were marked differences in total genomic inbreeding (represented by $F_{ROH \geq 100kb}$) and recent inbreeding (represented by $F_{ROH \geq 10Mb}$) among these populations, and these patterns were not apparent in measures of F_H. Estimates of F_{ROH} offer more accurate and precise measures of inbreeding than F_H, and allow for inference about the relative timing of inbreeding events. The inbreeding coefficient F_{ROH} is estimated on an individual basis using that individual's heterozygous or homozygous state at each SNP across the genome at hundreds of thousands of loci to identify tracts of that genome that are identical by descent. This is in contrast with F_H, which is also a measure of individual inbreeding, but is dependent upon the allele frequencies of an individual's subpopulation, and therefore influenced by sampling schemes and population genetic substructuring, known as the Wahlund effect (Wahlund 1928; Waples and Allendorf 2015). Inbreeding estimates using F_{ROH} offer a direct measure of the genomic patterns of homozygosity that result in inbreeding depression. Inbreeding events produce long stretches of the genome that are identical by descent and, if inbreeding is not sustained thereafter, long ROH are broken up by meiotic recombination in subsequent generations. Longer ROH (≥10 Mb) from recent inbreeding events contribute more to inbreeding depression and decreased fitness because they contain disproportionately higher fractions of deleterious homozygous variants and have disproportionately strong and negative effects on fitness (Szpiech et al. 2013), so the ability to characterize individual and cumulative lengths of runs of homozygosity is important for understanding and ensuring the long-term sustainability of small, isolated populations of conservation concern.

Wolves on POW exhibit intermediate levels of total genomic inbreeding ($F_{ROH \geq 100kb}$) relative to the other two Alaska populations that we investigated, and higher $F_{ROH \geq 100kb}$ than the IRNP wolves, which we included to provide context of a highly inbred population that was founded by two to three individuals and is known to have exhibited severe inbreeding depression (Räikkönen et al. 2009; Robinson et al. 2019). Wolves in the southeast group exhibited the highest levels of total genomic inbreeding ($F_{ROH \geq 100kb}$). This result is in concordance with previous work that used microsatellites to estimate F_H for wolves in this region (Breed 2007). It is unclear what might be driving heightened total inbreeding in this region, but there is potential historic geographic isolation may be driving this pattern. Geographic barriers like large inlets and fjords may reduce connectivity between packs in this region, resulting in decreased opportunities for wolves to mate with unrelated individuals. Wolves in the northwest group had the lowest F_{ROH} estimates, indicating that this population is relatively large and/or genetically connected with nearby populations.

When we examined medium-length ROH ($F_{ROH \geq 1Mb}$), we found that wolves on POW have the highest genomic inbreeding estimates, even when compared to IRNP wolves. This result, in conjunction with the result that POW wolves have higher $F_{ROH \geq 100kb}$ than wolves on IRNP, suggests that inbreeding has been occurring on POW at higher rates and for a longer period of time than in IRNP, which was founded by two or three individuals in the late 1940s (Wayne et al. 1991). We also found that POW wolves had similar $F_{ROH \geq 10Mb}$ estimates when compared to the IRNP wolves, and both POW and IRNP wolves had $F_{ROH \geq 10Mb}$ estimates above the southeast and northwest groups. This is notable because the IRNP population was founded by just two or three

individuals in the late 1940s (Wayne et al. 1991), peaked at 50 individuals in 1980 (Peterson et al. 2014), and received one immigrant in 1997 (Adams et al. 2011) before successful reproduction stopped in 2014 as the result of severe inbreeding depression (Peterson and Vucetich 2016). The similarities in $F_{ROH \geq 1Mb}$ and $F_{ROH \geq 10Mb}$ between POW and IRNP indicate that POW wolves have experienced substantial inbreeding in recent years and may be at high risk for exhibiting inbreeding depression.

Wolves have never been observed moving between POW and mainland Southeast Alaska. Movement between POW and the mainland would likely require one long swim (~6.2 km) to the Cleveland Peninsula through strong ocean currents via the Clarence Strait or at least five swims (longest straight-line swim measured at 2.7 km) to cross Kashevarof Islands (commonly referred to as "Snow Pass") and Zarembo Island at the northeast end of POW (Appendix 2). Wolves have been observed on Bushy and Shrubby Islands (ADF&G, unpublished data), which are part of the Kashevarof Islands/Snow Pass and lie between Prince of Wales and Zarembo Islands, indicating that wolves are capable of dispersing across islands in this region. Based on mark recapture data, multiple wolves elsewhere in Southeast Alaska and coastal British Columbia have been observed on islands up to 13 km apart (U.S. Fish and Wildlife Service and citations therein). We also identified one potential immigrant or offspring of an immigrant wolf on POW using genomic data, which further supports the hypothesis that wolves can move between the mainland and POW. However, the frequency of these immigration events is unknown, and the probability of these immigrants reproducing in the POW population is also unknown. Our results show that there is sufficient isolation between the mainland and POW that the two groups are readily distinguished from one another in both PCA and Admixture analyses, and migration from the mainland population does not appear to be mitigating inbreeding on POW.

The data we present here represent wolves sampled primarily in 2015 and 2016, just one to two years after the low population estimate of 89 individuals in 2014. It is likely that the majority of these samples were collected too soon after the low population estimate to capture any potential effect this decline in abundance may have had on individual genomic inbreeding. The low population estimate of 2014 likely resulted in increased mating events between related individuals in subsequent years, and it is therefore probable that wolves currently on POW have higher inbreeding coefficients than reported in this study unless recent successful migration from the mainland has also occurred.

In context of previous studies on inbreeding and inbreeding depression in wild wolf populations, our data suggest that wolves on POW may be approaching a point at which they have already or will soon begin to exhibit signs of inbreeding depression given their geographic isolation, recent low population estimates, and evidence of high proportions of the genome being in long runs of homozygosity. From 2013 through 2017, at least three wolves in the POW island complex have been observed with notably shortened tails (ADF&G unpublished data, photos in Appendix 6). The causes of these deformities are currently unknown, and they could be the result of trauma. However, it is also possible that the shortened tails are skeletal malformations with a genetic basis and caused by inbreeding, and are perhaps similar to the vertebral defects that have been observed in the highly inbred wolves on IRNP, although the those deformities involved extra lumbar vertebra (Räikkönen et al. 2009; Hedrick et al. 2016). Tissue samples have been obtained for one of the short-tailed wolves on POW, but given the difficulties in capturing and observing wolves in this habitat, it is uncertain whether samples from the other three are likely to be obtained.

Regardless, work is currently underway to genotype and estimate F_{ROH} for the sampled short-tailed wolf. Depending on inbreeding analysis results for this individual and whether samples can be obtained for additional wolves that exhibit skeletal deformities, there could be opportunities for additional research on whether inbreeding depression is occurring in this population.

Conservation Implications

In response to the 2014 petition to list the Alexander Archipelago wolf as threatened under the ESA, the U.S. Fish and Wildlife Service (FWS) determined that "listing the Alexander Archipelago wolf [was] not warranted" and that "the Alexander Archipelago wolf population on POW does not meet the criteria of the Service's [distinct population segment (DPS)] policy, and, therefore, it does not constitute a listable entity" (Fish and Wildlife Service 2016). For a population to qualify as a DPS under the ESA, it must be "(1) **Discrete** in relation to the remainder of the taxon to which it belongs; **and** (2) biologically and ecologically **significant** to the taxon to which it belongs" (Fish and Wildlife Service 2016). The FWS found that wolves on POW are discrete in that they are morphologically and genetically distinct from mainland populations (Goldman 1944; Weckworth et al. 2005; Breed 2007; Fish and Wildlife Service 2016). However, the FWS found that POW wolves are not significant to the taxon because (1) they do not reside in a unique setting relative to the rest of the taxon, (2) the loss of the population would not result in a significant gap in the taxon, (3) the population is not the only natural occurrence of the taxon, and (4) the population is not genetically unique – the genetic discreteness recognized above can be attributed to founder effects and/or genetic drift (Weckworth et al. 2005; Breed 2007; Weckworth et al. 2010; Weckworth et al. 2011) and there is no evidence to date suggesting that the population contains unique adaptive variation (Fish and Wildlife Service 2016). Our results support the conclusion that wolves on POW are a discrete population segment, but without further analyses on population demographic history we are not yet able to determine whether this discreteness is the result of recent genetic drift or if the wolves on POW represent an older, distinct evolutionary lineage potentially originating from a glacial refugium.

The FWS also found that "inbreeding likely is not affecting the [POW] population despite its comparatively small size and insularity" based on the fact that wolves on POW had lower F_H estimates than wolves in GMU 1 (the southeast regoin of our study area; Breed 2007; Fish and Wildlife Service 2016). The data we present are strongly contrary to this conclusion. We observed that POW wolves have the highest $F_{ROH \geq 1Mb}$ and $F_{ROH \geq 10Mb}$ estimates compared to the other two Alaska populations in our study. We also observed that POW wolves have similar $F_{ROH \geq 10Mb}$ as IRNP wolves, which have demonstrated severe inbreeding depression (Räikkönen et al. 2009; Robinson et al. 2019). It has been shown previously that long runs of homozygosity are more likely to contribute to inbreeding depression and have disproportionately stronger negative effects on fitness (Szpiech et al. 2013). Of the three Alaska populations we assessed, POW wolves have the highest $F_{ROH \geq 10Mb}$ estimates and are likely at greatest risk for exhibiting signs of inbreeding depression. Inbreeding coefficient estimates for wolves on POW are not different from inbreeding coefficient estimates observed on IRNP, where wolves exhibited severe inbreeding depression and skeletal malformations (Räikkönen et al. 2009; Robinson et al. 2019). Furthermore, as many as four wolves on POW have been observed with short tails (Appendix 6). Although the tail deformities in these wolves could be from trauma and have not been conclusively attributed to a genetic basis or inbreeding depression, vertebral malformations have been observed in other inbred

wolf populations exhibiting severe inbreeding depression (Liberg et al. 2005; Räikkönen et al. 2009).

The primary concerns expressed in the original petition to list POW Alexander Archipelago wolves under the ESA were regarding habitat loss, high harvest rates, and declining population estimates (Toppenberg et al. 2015). Habitat loss via old-growth logging is a challenging issue to address and prevent, and while added extensive loss of old-growth forest habitat should be avoided to facilitate effective conservation of both Sitka black-tailed deer and Alexander Archipelago wolves, wolves are known to be resilient and flexible in both habitat and prey selection (Roffler et al. 2018; Roffler and Gregovich 2018). Wolf hunting and trapping quotas are managed directly by ADF&G and can be modified on an annual basis to control harvest and meet a target population size, and wolf abundances are known to rebound relatively quickly after harvest pressures have been relieved (e.g. see increase in annual population estimates from 2015 to 2016 in Figure 3). However, inbreeding is potentially a hidden and insidious threat to small, isolated populations, especially for populations which are difficult to monitor, like POW wolves. Inbreeding can pose significant threats to small, isolated populations, and these threats are difficult to rectify without substantial and costly management action (e.g. translocation of individuals from outside populations) to provide a genetic rescue to the inbred population. There are many challenges involved with translocating individuals, including risk of mortality during translocations and ensuring that translocated individuals are genetically compatible to the recipient population. It is therefore important to consider inbreeding when defining minimum population targets and to monitor inbreeding to avoid allowing a population to enter an extinction vortex.

Future Directions

This work represents an exciting and informative development in the field of wildlife conservation genomics. The flexibility afforded by the capture approach allows for characterization of inbreeding in species which lack a high-quality reference genome (however, a reference genome from a closely-related species is necessary for the capture design and mapping sequence reads). We are now able to monitor populations for precise and accurate measures of genomic inbreeding, as well as the potential of inbreeding depression before we begin to see evidence of reduced fitness. For small, isolated populations of conservation concern where habitat loss or other pressures may be impacting population viability, it will be beneficial to obtain an accurate and precise measure of inbreeding using genomic data. A sound understanding of inbreeding in these populations can inform whether there may be a need to modify management plans to ensure long-term sustainability.

Acknowledgements

Special thanks to Matthew R. Jones for valuable discussions and comments on methods, results, and drafts of this manuscript. Thanks also to Kris Larson for field work guidance and sample collection, Michael Kampnich for sharing field observations and photos, Colin Callahan for assistance with optimization of the library prep protocol, David Xing for laboratory and sequencing assistance, Kelton Verble for discussions about results, and Todd Cross for valued input on a draft of this manuscript. This work was funded by The Alaska Department of Fish and Game.

Table 1. Collection years for Alexander Archipelago wolf (*Canis lupus ligoni*) tissue samples in each Alaska Department of Fish & Game Game Management Unit in Southeast Alaska. *Not shown in the table is one tissue sample collected in the Yukon Territory, Canada in 2011, for a total number of 59 samples analyzed in this study.

Game Management Unit	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Unknown	Total
1A		1	3											5			9
1B		1			1									1		1	4
1C									1		1			1	1		4
1D	1	3	1														5
2			1											15			16
3						2						1	1		12		16
4															2		2
5A				2													2
Total	1	5	5	2	1	2	0	0	1	0*	1	1	1	22	15	1	58*

Table 2. Number of SNPs remaining after each filtering step when each step is applied individually (not sequentially) to the original dataset of 360,354 SNPs on autosomal chromosomes 1 through 38 in tissue samples collected from 59 Alexander Archipelago wolves (*Canis lupus ligoni*) in Southeast Alaska and Yukon Territory from 2002 through 2016.

Filter description	SNPs Remaining
No filter	**360,354**
Retain sites with minor allele frequency > 0.05	226,729
Retain sites with mean coverage ≥ 10	89,987
Retain sites with mean coverage ≤ 50	357,688
Retain biallelic loci only	351,855
Remove sites out of Hardy-Weinberg equilibrium ($p \leq 0.001$)	280,278
Remove sites with missing data in ≥15% individuals	195,536
All above filters combined	**37,082**

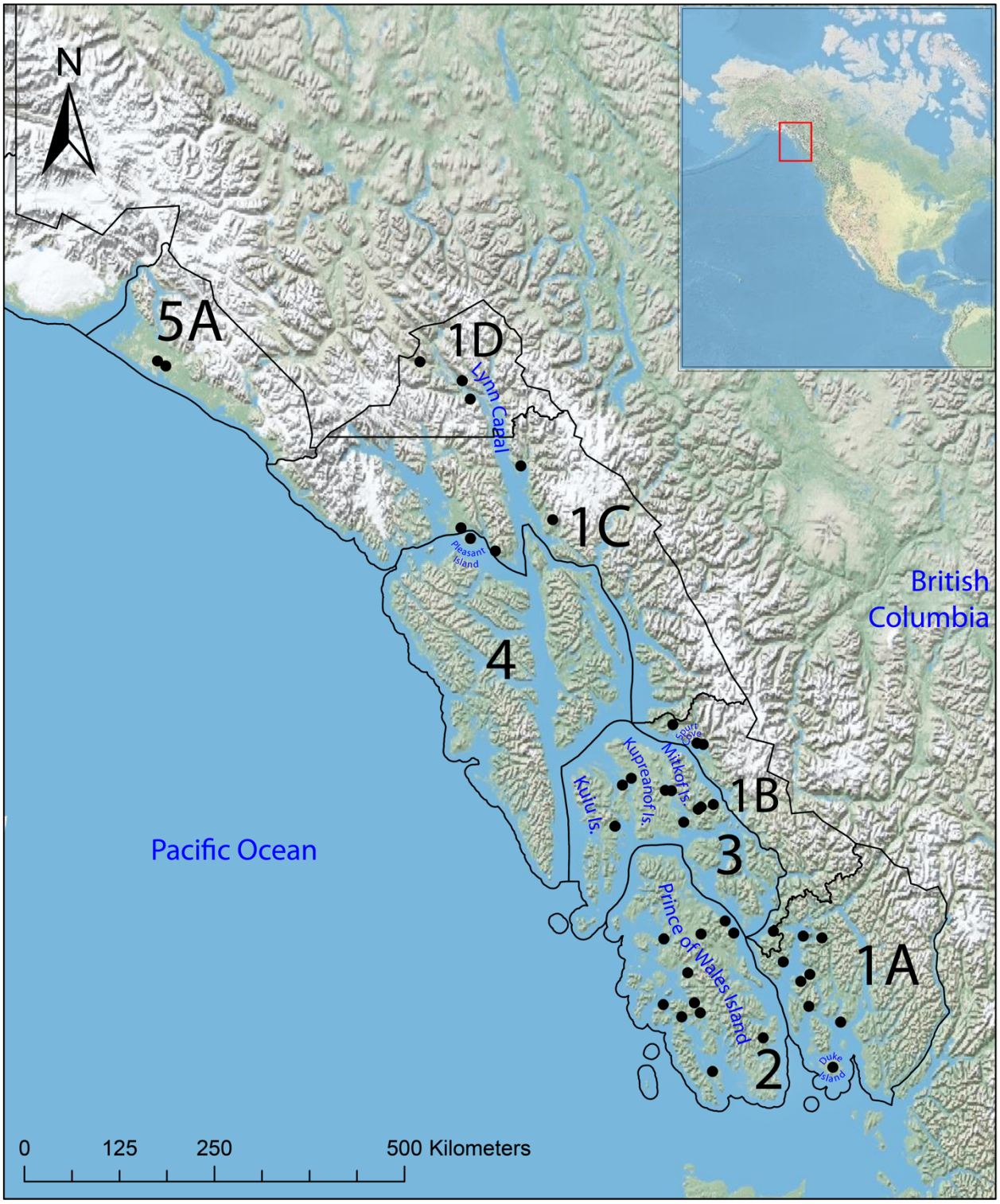


Figure 1. Map of Southeast Alaska showing locations of individual samples (●), Alaska Department of Fish and Game Game Management Units and Subunits (black text and outlines), and geographical features (blue text) referenced in main text. Note that some sampling locations overlap (especially those on Prince of Wales Island). Wolves do not inhabit large islands in Game Management Unit 4. Yukon Territory sample not shown.

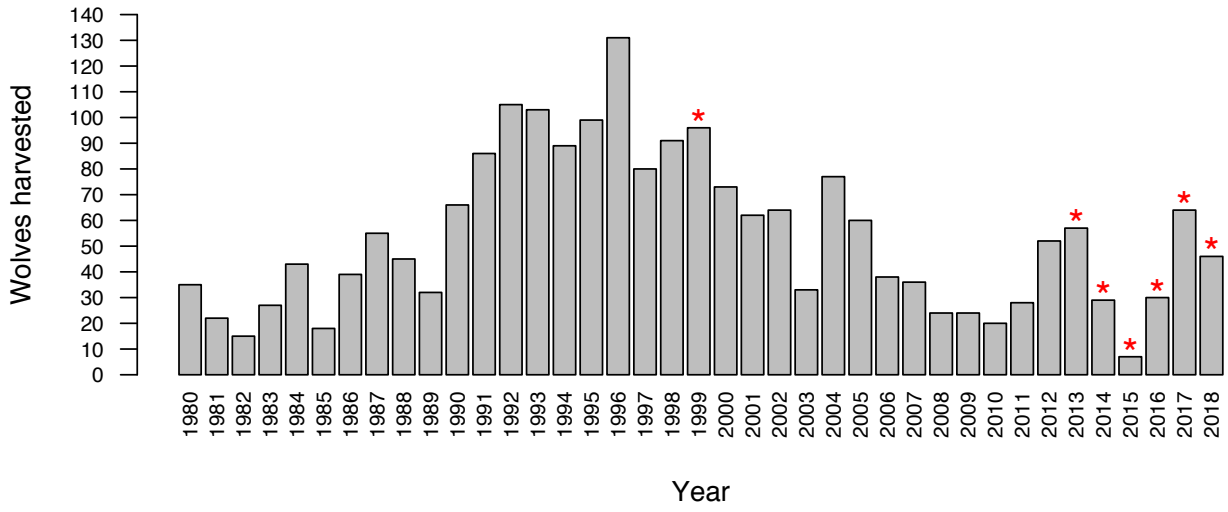


Figure 2. Number of Alexander Archipelago wolves (*Canis lupus ligoni*) harvested on an annual basis in Alaska Department of Fish & Game Game Management Unit 2 from 1980 through 2018. Red asterisks indicate years that the wolf harvest season was closed early by emergency order (Department of the Interior; Alaska Department of Fish and Game unpublished data).

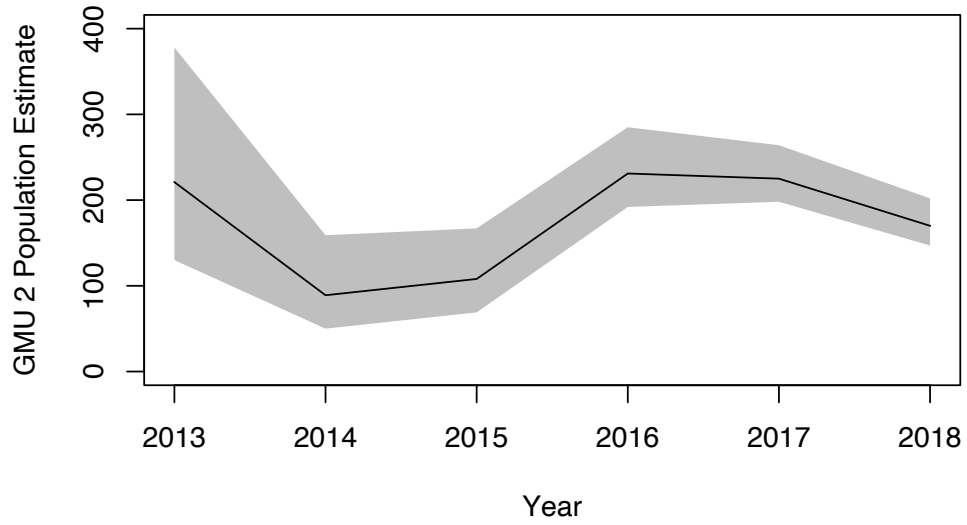


Figure 3. Annual fall (October-December) population estimates for Alexander Archipelago wolves (*Canis lupus ligoni*) in Alaska Department of Fish & Game Game Management Unit 2. Population estimates were determined using spatially-explicit capture-recapture modeling (Roffler et al. 2019). The black line denotes the population estimate calculated on an annual basis from DNA-based individual identification using hair boards surveyed from October through December. The gray shaded intervals represent the 95% confidence interval for each year's population estimate (Roffler 2015; Roffler et al. 2016; Roffler 2016; Roffler 2017; Roffler 2018; Roffler et al. 2019).

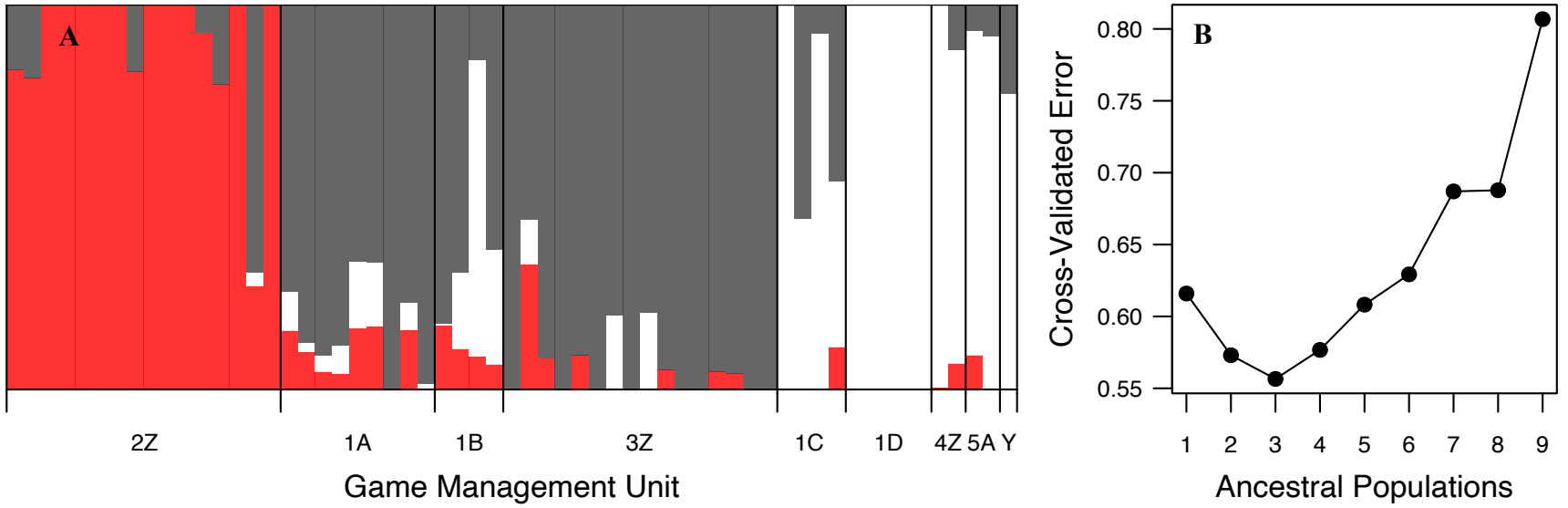


Figure 4. Proportion ancestry assigned for individual Alexander Archipelago wolves (*Canis lupus ligoni*) assuming three ancestral populations and using 37,082 genome-wide SNPs and program Admixture (Alexander et al. 2009). (A) Proportion ancestry for each sample in Alaska Game Management Units 2 (Prince of Wales and nearby islands), 1A, 1B, 3, 1C, 1D, 4, 5A, and for one sample collected in Yukon Territory, Canada (labeled "Y"). (B) Cross-validated errors for one through eight putative ancestral populations, lowest cross-validated error used to select most likely number of ancestral populations

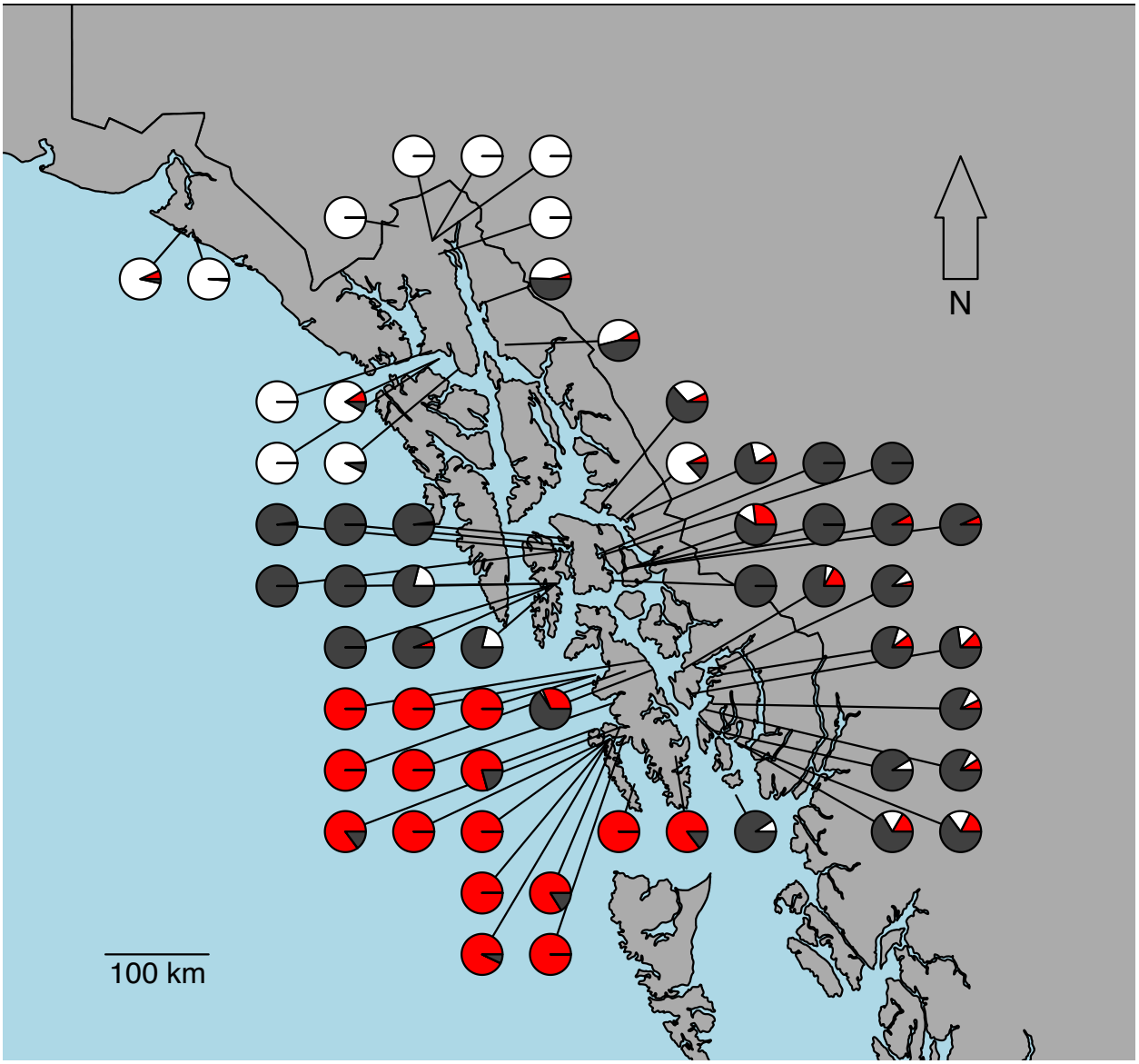


Figure 5. Proportion ancestry assigned for each Alexander Archipelago wolf (*Canis lupus ligoni*) sampled in Southeast Alaska assuming 3 ancestral populations using 37,082 genome-wide SNPs and program Admixture (Alexander et al. 2009). Proportion ancestry for the Yukon Territory wolf is not shown here, but can be viewed in Figure 4A.

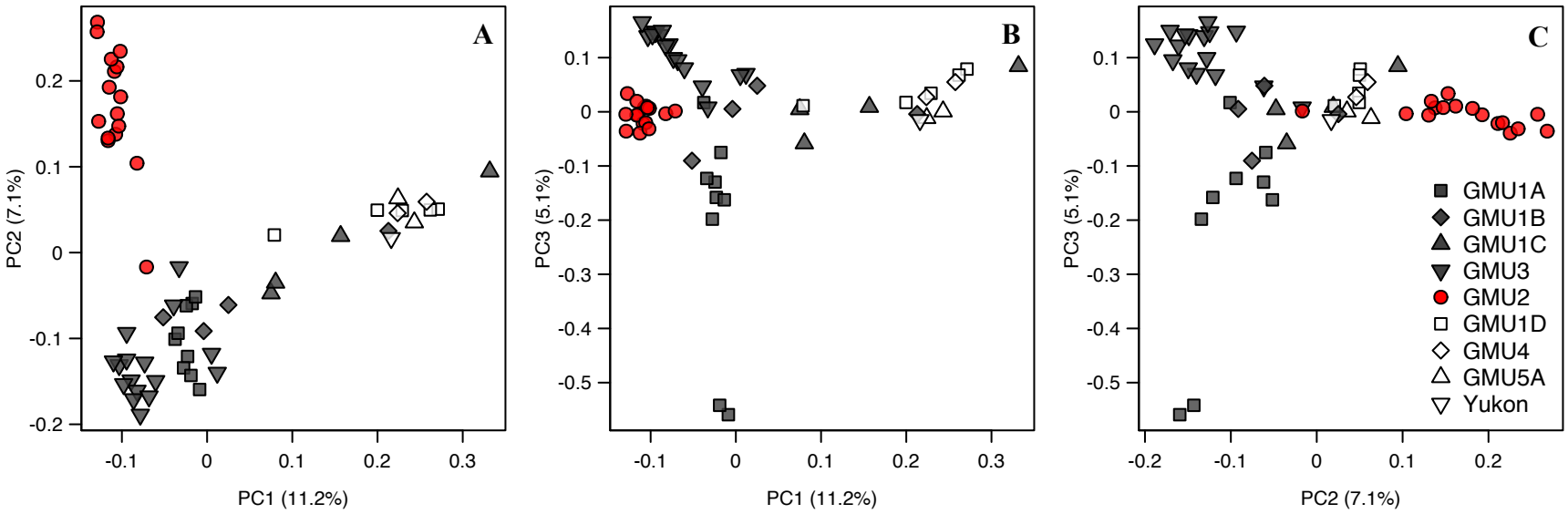


Figure 6. Principle components analysis of 37,082 genome-wide SNPs from 59 Alexander Archipelago wolves (*Canis lupus ligoni*) sampled throughout Southeast Alaska. Color and symbol for each sample corresponds with Alaska Department of Fish and Game Game Management Unit, as shown in the legend in plot C. (**A**) PC1 versus PC2, (**B**) PC1 versus PC3, (**C**) PC2 versus PC3.

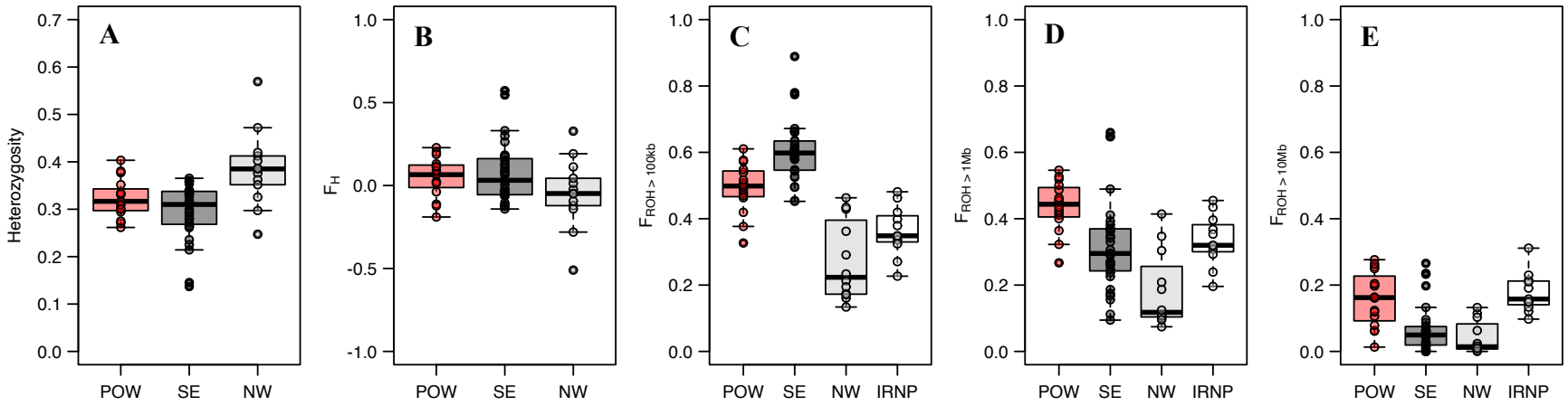


Figure 7. Boxplots showing inbreeding coefficients for Alexander Archipelago wolves (*Canis lupus ligoni*) in Admixture-defined Prince of Wales group (POW; n=15; Prince of Wales, Dall, Warm Chuck, Long, and Suemez Islands), southeast group (SE; n=31; Prince of Wales, Kuiu, Kupreanof, Mitkof, Duke Islands and mainland east of Lynn Canal), northwest group (NW; n=13; Pleasant Island and the mainland west of Lynn Canal), and for all F_{ROH} calculations, Isle Royale National Park (IRNP; n=11; Robinson et al. 2019). Calculations of F_{ROH} were generated using >540k SNP genotype likelihoods for Southeast Alaska wolves, and whole genome sequence data for IRNP wolves (Robinson et al. 2019). (**A**) Heterozygosity calculated as proportion heterozygous sites using 37,082 SNPs. (**B**) Individual inbreeding coefficients (F_H) (**C**) Proportion of the genome (autosomes only) in runs of homozygosity greater than 100 kb. (**D**) Proportion of the genome in runs of homozygosity greater than 1 Mb. (**E**) Proportion of the genome in runs of homozygosity greater than 10 Mb. Thick black line represents the mean, the lower and upper box boundaries are the 25th and 75th quartiles (Q1 and Q3), and the whiskers are Q1-[1.5×(interquartile range; IQR)] and Q3+(1.5×IQR). Dots overlaying the boxplots represent individual wolves in each group.

LITERATURE CITED

Adams JR, Vucetich LM, Hedrick PW, et al (2011) Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. Proc R Soc B-Biological Sci 278:3336–3344. doi: 10.1098/rspb.2011.0261

Albert DM, Schoen JW (2013) Use of historical logging patterns to identify disproportionately logged ecosystems within temperate rainforests of southeastern alaska. Conserv Biol 27:774–784. doi: 10.1111/cobi.12109

Albrechtsen A, Nielsen FC, Nielsen R (2010) Ascertainment biases in SNP chips affect measures of population divergence. Mol Biol Evol 27:2534–2547. doi: 10.1093/molbev/msq148

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664. doi: 10.1101/gr.094052.109

Ali OA, O'Rourke SM, Amish SJ, et al (2015) RAD Capture (Rapture): Flexible and efficient sequence-based genotyping. 64:87–95.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709. doi: 10.1038/nrg2844

Allendorf FW, Luikart G, Aitken SN (2013) Conservation and the Genetics of Populations, 2nd edn. Wiley-Blackwell, Hoboken, New Jersey

Andrews S, Krueger F, Segonds-Pichon A, et al (2012) FastQC. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ávila-Arcos MC, Cappellini E, Romero-Navarro JA, et al (2011) Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Sci Rep. doi: 10.1038/srep00074

Becker RA, Wilks AR, Brownrigg R (2018) mapdata: Extra Map Databases.

Benson G (1999) Tandem repeats finder : a program to analyze DNA sequences. Nucleic Acids Res 27:573–580.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. doi: 10.1093/bioinformatics/btu170

Breed M (2007) The tale of the dire effects of hunting on wolves in coastal Southeast Alaska: loss of genetic diversity, fragmentation and a regional sink. Uppsala University

Broad Institute (2016) HC step 4: Assigning per-sample genotypes. In: gatk User Guid. Methods Algorithms. https://software.broadinstitute.org/gatk/documentation/article.php?id=4442. Accessed 22 Jul 2019

Burbano H a, Hodges E, Green RE, et al (2010) Targeted Investigation of the Neandertal Genome by Array-Based Sequence Capture. Science (80-) 723:9–11. doi: 10.1126/science.1188046

Carpenter ML, Buenrostro JD, Valdiosera C, et al (2013) Pulling out the 1%: Whole-Genome capture for the targeted enrichment of ancient dna sequencing libraries. Am J Hum Genet 93:852–864. doi: 10.1016/j.ajhg.2013.10.002

Clark AG, Hubisz MJ, Bustamante CD, et al (2005) Ascertainment bias in studies of human genome-wide polymorphism. Genome Res 15:1496–1502. doi: 10.1101/gr.4107905

Cronin MA, Canovas A, Bannasch DL, et al (2014) Single Nucleotide Polymorphism (SNP) Variation of Wolves (Canis lupus) in Southeast Alaska and Comparison with Wolves, Dogs, and Coyotes in North America. J Hered 106:26–36. doi: 10.1093/jhered/esu075

Cronin MA, Canovas A, Bannasch DL, et al (2015) Wolf Subspecies: Reply to Weckworth et al. and Fredrickson et al. J Hered 106:1–3. doi: 10.1093/jhered/esv029

Danecek P, Auton A, Abecasis G, et al (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi: 10.1093/bioinformatics/btr330

Department of the Interior Subsistence management of wolves in the Southeast Alaska region. https://www.doi.gov/sites/doi.gov/files/uploads/4_subsistence_management_of_wolves_southeast_alaska_final.pdf. Accessed 8 Jun 2018

DePristo MA, Banks E, Poplin R, et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–501. doi: 10.1038/ng.806

Durbin K (1990) Tongass: Pulp Politics and the Fight for the Alaska Rain Forest, First. Oregon State University Press, Corvallis

Enk JM, Devault AM, Kuch M, et al (2014) Ancient whole genome enrichment using baits built from modern dna. Mol Biol Evol 31:1292–1294. doi: 10.1093/molbev/msu074

Faircloth B, Glenn T (2011) Serapure v2.2. Genome Res.

Fish and Wildlife Service (1997) Endangered and Threatened Wildlife and Plants; 12-Month Finding for a Petition to List the Alexander Archipelago Wolf as Threatened and to Designate Critical Habitat. Fed. Regist. 62:

Fish and Wildlife Service (2014) 90-day finding on a petition to list the Alexander Archipelago wolf as threatened or endangered. United States of America

Fish and Wildlife Service (2010) Endangered and Threatened Wildlife and Plants; 12-Month Finding for a Petition to List the Alexander Archipelago Wolf as Threatened.

Fish and Wildlife Service (2016) Endangered and Threatened Wildlife and Plants; 12-Month Finding on a Petition To List the Alexander Archipelago Wolf as an Endangered or Threatened Species. United States of America

Fredrickson RJ, Hedrick PW, Wayne RK, et al (2015) Mexican wolves are a valid subspecies and an appropriate conservation target. J Hered 106:415–416. doi: 10.1093/jhered/esv028

Freedman AH, Gronau I, Schweizer RM, et al (2014) Genome Sequencing Highlights the Dynamic Early History of Dogs. PLoS Genet. doi: 10.1371/journal.pgen.1004016

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. Trends Ecol Evol 27:489–496. doi: 10.1016/j.tree.2012.05.012

Gleason SL (2019) SEACC, et al. v. U.S. Forest Serv., et al. Case No. 119-cv-00006-SLG

Goldman EA (1944) The Wolves of North America, Part 2. Dover Publications Inc., New York

Hanley TA (1984) Relationships Between Sitka Black-Tailed Deer and Their Habitat. Gen Tech Rep PNW-168:2–13.

Hedrick PW, Kardos M, Peterson RO, Vucetich JA (2016) Genomic variation of inbreeding and ancestry in the remaining two Isle Royale wolves. J Hered 1–36. doi: 10.1016/j.ceca.2015.06.005

Hedrick PW, Peterson RO, Vucetich LM, et al (2014) Genetic rescue in Isle Royale wolves: genetic analysis and the collapse of the population. Conserv Genet 1111–1121. doi: 10.1007/s10592-014-0604-1

Hernandez-Rodriguez J, Arandjelovic M, Lester J, et al (2017) The impact of endogenous content, replicates and pooling on genome capture from fecal samples. Mol Ecol Resour. doi: 10.1111/ijlh.12426

Illumina (2010) CanineHD BeadChip.

Illumina (2007) CanineSNP20 BeadChip. Cust. Solut.

Jones MR, Good JM (2016) Targeted capture in evolutionary and ecological genomics. Mol Ecol 25:185–202. doi: 10.1111/mec.13304

Kardos M, Åkesson M, Fountain T, et al (2017) Genomic consequences of intensive inbreeding in an isolated wolf population. Nat Ecol Evol 1:124–131. doi: 10.1038/s41559-017-0375-4

Kardos M, Luikart G, Allendorf FW (2015) Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. Heredity (Edinb) 115:63–72. doi: 10.1038/hdy.2015.17

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17:230–241. doi: 10.1016/S0169-5347(02)02489-8

Keller MC, Visscher PM, Goddard ME (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189:237–249. doi: 10.1534/genetics.111.130922

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15:1–13. doi: 10.1186/s12859-014-0356-4

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. doi: 10.1093/bioinformatics/btr509

Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. doi: 10.1093/bioinformatics/btp352

Liberg O, Andrén H, Pedersen H-C, et al (2005) Severe inbreeding depression in a wild wolf (Canis lupus) population. Biol Lett 1:17–20. doi: 10.1098/rsbl.2004.0266

Lindblad-Toh K, Wade CM, Mikkelsen TS, et al (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803–819. doi: 10.1038/nature04338

Lowry DB, Hoban S, Kelley JL, et al (2017) Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol Ecol Resour 17:142–152. doi: 10.1111/1755-0998.12635

Manichaikul A, Mychaleckyj JC, Rich SS, et al (2010) Robust relationship inference in genome-wide association studies. Bioinformatics 26:2867–2873. doi: 10.1093/bioinformatics/btq559

McKenna A, Hanna M, Banks E, et al (2010) The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. doi: 10.1101/gr.107524.110

Mech LD, Boitani L (2007) Wolves: Behavior, Ecology, and Conservation. The University of Chicago Press, Chicago, IL

Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc. doi: 10.1101/pdb.prot5448

Morgulis A, Gertz EM, Schäffer AA, Agarwala R (2006) WindowMasker: Window-based masker for sequenced genomes. Bioinformatics 22:134–141. doi: 10.1093/bioinformatics/bti774

Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet 12:443–451. doi: 10.1038/nrg2986

Nowak RM (1995) Another Look at Wolf Taxonomy. Ecol. Conserv. Wolves a Chang. World 375–397.

Nowak RM, Paradiso JL (1983) Walker's mammals of the world. Johns Hopkins University Press, Baltimore, MD

Paijmans JLA, Fickel J, Courtiol A, et al (2016) Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. Mol Ecol Resour 16:42–55. doi: 10.1111/1755-0998.12420

Pante E, Simon-Bouhet B (2013) marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. PLoS One 8:6–9. doi: 10.1371/journal.pone.0073051

Pemberton TJ, Absher D, Feldman MW, et al (2012) Genomic patterns of homozygosity in worldwide human populations. Am J Hum Genet 91:275–292. doi: 10.1016/j.ajhg.2012.06.014

Person DK (2001) Alexander Archipelago wolves: Ecology and population viability in a disturbed, insular landscape. University of Alaska Fairbanks

Person DK, Kirchhoff M, Ballenberghe V Van, et al (1996) The Alexander Archipelago Wolf: A Conservation Assessment. Portland, OR

Person DK, Logan BD (2012) A spatial analysis of wolf harvest and harvest risk on Prince of Wales and associated islands, Southeast Alaska. Final Wildlife Research Report, WRR-2012-06. Juneau, AK

Person DK, Russell AL (2008) Correlates of mortality in an exploited wolf population. J Wildl Manage 72:1540–1549. doi: 10.2193/2007-520

Peterson RO, Vucetich JA (2016) Ecological Studies of Wolves on Isle Royale. Houghton, MI

Peterson RO, Vucetich JA, Bump JM, Smith DW (2014) Trophic Cascades in a Multicausal World: Isle Royale and Yellowstone. Annu Rev Ecol Evol Syst 45:325–345. doi: 10.1146/annurev-ecolsys-120213-091634

Price AL, Patterson NJ, Plenge RM, et al (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909. doi: 10.1038/ng1847

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/.

Räikkönen J, Vucetich JA, Peterson RO, Nelson MP (2009) Congenital bone deformities and the inbred wolves (Canis lupus) of Isle Royale. Biol Conserv 142:1025–1031. doi: 10.1016/j.biocon.2009.01.014

Robinson JA, Räikkönen J, Vucetich LM, et al (2019) Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. Sci Adv. doi: 10.1126/sciadv.aau0757

Roche (2017) SeqCap EZ HyperCap Workflow User's Guide Version 2.1. 26–36.

Roffler GH (2015) GMU 2 Wolf Population Estimate Update, Fall 2014.

Roffler GH (2017) Memorandum on GMU 2 Wolf Population Update, Autumn 2016. Douglas, AK

Roffler GH (2016) GMU 2 Wolf Population Estimate Update, Autumn 2015. Douglas, AK

Roffler GH (2018) GMU 2 Wolf Population Estimate Update, Autumn 2017. Douglas, AK

Roffler GH, Gregovich DP (2018) Wolf space use during denning season on Prince of Wales Island, Alaska. Wildlife Biol 2018:wlb.00468. doi: 10.2981/wlb.00468

Roffler GH, Gregovich DP, Larson KR (2018) Resource selection by coastal wolves reveals the seasonal importance of seral forest and suitable prey habitat. For Ecol Manage 409:190–201. doi: 10.1016/j.foreco.2017.11.025

Roffler GH, Waite JN, Flynn RW, et al (2016) Wolf population estimation on Prince of Wales Island, Southeast Alaska: a comparison of methods. Juneau, AK

Roffler GH, Waite JN, Pilgrim KL, et al (2019) Estimating abundance of a cryptic social carnivore using spatially explicit capture–recapture. Wildl Soc Bull. doi: 10.1002/wsb.953

RStudio Team (2016) RStudio: Integrated Development Environment for R. RStudio Integr. Dev. R

Schweizer RM, VonHoldt BM, Harrigan R, et al (2015) Genetic subdivision and candidate genes under selection in North American gray wolves. Mol Ecol. doi: 10.1111/mec.13364

Shafer ABA, Miller JM, Kardos M (2016a) Cross-Species Application of SNP Chips is Not Suitable for Identifying Runs of Homozygosity. J Hered 107:193–195. doi: 10.1093/jhered/esv137

Shafer ABA, Peart CR, Tusso S, et al (2016b) Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Methods Ecol Evol 1–11. doi: 10.1111/2041-210X.12700

Shanley CS, Pyare S, Goldstein MI, et al (2015) Climate change implications in the northern coastal temperate rainforest of North America. Clim Change 130:155–170. doi: 10.1007/s10584-015-1355-9

Sikes RS, Gannon WL, Mammalogists AC and UC of the AS of (2011) Guidelines of the American Society of Mammalogists for the use of wild mammals in research. J Mammal 92:235–253. doi: 10.1644/10-mamm-f-355.1

Smit AFA, Hubley R, Green P RepeatMasker Open-4.0. http://www.repeatmasker.org.

Suchan T, Pitteloud C, Gerasimova NS, et al (2016) Hybridization Capture Using RAD Probes (hyRAD), a New Tool for Performing Genomic Analyses on Collection Specimens. PLoS One 11:22. doi: 10.1371/journal.pone.0151651

Supple MA, Shapiro B (2018) Conservation of biodiversity in the genomics era. Genome Biol 19:1–12. doi: 10.1186/s13059-018-1520-3

Szpiech ZA, Xu J, Pemberton TJ, et al (2013) Long runs of homozygosity are enriched for deleterious variation. Am J Hum Genet 93:90–102. doi: 10.1016/j.ajhg.2013.05.003

Templeton AR, Read B (1994) Inbreeding: one word, several meanings, much confusion. EXS 68:91–105. doi: 10.1007/978-3-0348-8510-2_9

Tengvall K (2015) Genetic Studies in Dogs Implicate Novel Genes Involved in Atopic Dermatitis and IgA Deficiency.

Toppenberg J, Scott G, Noblin R, et al (2015) Petition to List on an Emergency Basis the Alexander Archipelago Wolf (Canis Lupus Ligoni) as Threatened or Endangered Under the Endangered Species Act.

U.S. Fish and Wildlife Service Species status assessment for the Alexander Archipelago Wolf (Canis lupus ligoni). Anchorage, AK

USDA Forest Service (2016) Tongass National Forest Land and Resource Management Plan.

Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinforma 43:11.10.1-11.10.33. doi: 10.1002/0471250953.bi1110s43.From

VonHoldt BM, Pollinger JP, Earl DA, et al (2011) A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Res 21:1294–1305. doi: 10.1101/gr.116301.110

Wahlund S (1928) Zusammensetzung von Populationen und Korrelationerscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas 11:65–106.

Wang S, Haynes C, Barany F, Ott J (2009) Genome-Wide Autozygosity Mapping in Human Populations. Genet Epidemiol 33:172–180. doi: 10.1002/gepi.20344

Waples RS, Allendorf F (2015) Testing for hardy-weinberg proportions: Have we lost the plot? J Hered 106:1–19. doi: 10.1093/jhered/esu062

Wayne RK, Lehman N, Girman D, et al (1991) Conservation Genetics of the Endangered Isle Royale Gray Wolf. Soc Conserv Biol 5:41–51.

Weckworth B, Talbot S, Cook J (2010) Phylogeography of wolves (Canis lupus) in the Pacific Northwest. J Mammal 91:363–375. doi: 10.1644/09-MAMM-A-036.1.Key

Weckworth B V., Dawson NG, Talbot SL, et al (2011) Going coastal: Shared evolutionary history between coastal British Columbia and Southeast Alaska wolves (canis lupus). PLoS One. doi: 10.1371/journal.pone.0019582

Weckworth B V., Dawson NG, Talbot SL, Cook J a. (2015) Genetic Distinctiveness of Alexander Archipelago Wolves (Canis lupus ligoni): Reply to Cronin et al. (2015). J Hered 1–3. doi: 10.1093/jhered/esv026

Weckworth B V., Talbot S, Sage GK, et al (2005) A signal for independent coastal and continental histories among North American wolves. Mol Ecol 14:917–931. doi: 10.1111/j.1365-294X.2005.02461.x

White LC, Fontsere C, Lizano E, et al (2019) A roadmap for high-throughput sequencing studies of wild animal populations using noninvasive samples and hybridization capture. Mol Ecol Resour 19:609–622. doi: 10.1111/1755-0998.12993

Wickham H, RStudio (2017) tidyverse: Easily Install and Load the "Tidyverse."

Wickham H, RStudio (2018) scales: Scale Functions for Visualization.

Wright S (1922) Coefficients of Inbreeding and Relationship. Am Nat 56:330–338.

Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: A fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30:614–620. doi: 10.1093/bioinformatics/btt593

Appendices

Appendix 1: Rationalization for Capture Approach

A major requirement for this method was high compatibility with DNA from non-invasively collected samples like hairs and scats. Because capturing wolves is costly, time-intensive, and causes stress and potential injury to the animals (Roffler et al. 2019), many hair and scat samples across Southeast Alaska have been collected opportunistically to fill in geographic gaps in sampling for the ADF&G-led population genetics study. Hair and scat samples are some of the only samples available for hypothesized migration corridors in Southeast Alaska, so the ability to incorporate those sample types into our analyses is essential in gaining an accurate understanding of wolf population connectivity in this region.

We considered several preexisting SNP genotyping methods to identify the most appropriate methodological approach for our small- and large-scale questions and the requirement for compatibility with DNA samples spanning a wide range of qualities and quantities. Recent SNP-based studies on wolves have utilized commercially available SNP chips or microarrays designed to target variable loci in the domestic dog genome (e.g. VonHoldt et al. 2011; Cronin et al. 2014; Schweizer et al. 2015). We were concerned that in the context of the larger-scale population genetic structure questions posed by ADF&G, the ascertainment bias incurred by these methods might influence our ability to accurately detect wolf population genetic structure and identify unique population segments across Southeast Alaska. Previous work has shown that the application of a commercially-available 800k Bovid SNP Chip (designed to target variable loci in domestic cattle) to two species of wild cattle and two species of domestic cattle greatly exaggerated genetic differentiation between the two domestic cattle species: this approach also artificially diminished genetic differentiation both within and between the two wild cattle species in a PCA relative to results for the same samples genotyped using RAD-seq, which is substantially less prone to ascertainment bias (Michael R. Miller, personal communication, September 6, 2015). Commercially available domestic dog SNP chips and microarrays are often designed to inform genetic disease risk factors and breed association studies (Illumina 2007; Illumina 2010). Although average genome-wide divergence between dogs and wolves is very low (~0.11%, Freedman et al. 2014), domestic dogs have experienced thousands of years of artificial selection by humans for a variety of highly specialized physical and behavioral traits, and as a result the genetic variation present among dog breeds is greatly differentiated from the variation we expect to see in wild wolves. As such, SNP chips and microarray designed to inform breed association and genetic risk factor studies in domestic dogs cannot offer an unbiased sample of genomic variation in wolves to address population genomics questions. Furthermore, previous work has also shown that cross-species applications of SNP chips is also inappropriate for identifying runs of homozygosity (Shafer et al. 2016a). Additional work has shown that the ascertainment bias incurred in the initial selection of polymorphic loci can have strong effects on F_{ST} (Clark et al. 2005) and PCAs (Albrechtsen et al. 2010). For these reasons, commercially available SNP chips and microarrays were deemed unsuitable to address our research questions.

RAD-seq and similar methods (e.g. double digest RAD-seq, genotyping by sequencing, nextRAD, ezRAD, etc.) are very well-suited for exploring population genetics questions and can genotype up to hundreds of thousands of SNPs (Lowry et al. 2017). However, they require large amounts of

high-quality, non-contaminated DNA as input (Jones and Good 2016). This precludes RAD-seq and similar methods from compatibility with DNA extracted from hair and scat, which yield small amounts of wolf DNA that are likely to be mixed with DNA from bacteria, fungus, and, in the case of scat samples, also mixed with potentially large amounts of DNA from prey like deer, beaver, fish, and black bear. Other methods combining RAD-based approaches with more flexible capture methods, like Rapture (Ali et al. 2015) or hyRAD (Suchan et al. 2016), may be more compatible with DNA from non-invasive samples, but data generated by these methods tend to have low allele call frequencies (Shafer et al. 2016b), risking obtaining a small number of SNPs that would be unsuitable for inbreeding analyses, and an initial RAD-seq effort would be required to define targeted RAD loci. Because fine-scale population genetic structure of wolves across Southeast Alaska is unknown, an approach like RAD-seq would have required a substantial and costly initial library preparation and sequencing effort to identify a large number of loci to accurately represent the population genetic structure in the region, making this approach infeasible from a per-sample cost perspective.

A custom-designed hybridization capture met all four requirements posed by the research efforts outlined in this thesis and by the ADF&G, and was determined to be the least biased, most time-efficient and cost-effective method to genotype SNPs across many samples from an unknown number of populations. Because hybridization capture can enrich targeted regions with up to 40% divergence (though capture efficiency begins to decline at 4-10% divergence, Paijmans et al. 2016) and wolves and domestic dogs are approximately 0.11% divergent genome-wide (Freedman et al. 2014), we designed a wolf-compatible capture using sequence data from the annotated domestic dog genome (canFam3.1, Lindblad-Toh et al. 2005). A significant advantage of using a *de novo* capture design was that we were able to exclude exons to ensure that we genotyped tens of thousands of neutral SNPs, which are more appropriate for addressing questions about population genetic structure. We were also able to evenly space the probes throughout the genome to ensure that we captured as many unlinked SNPs (i.e. SNPs unaffected by recombination, thus co-inherited) as possible. Our approach was robust to ascertainment bias in that it provided anonymous sampling of genome-wide SNPs without the need for a preliminary sequencing and genotyping effort on a smaller subset of samples to identify variable loci. The genetic variation identified and characterized using our approach was representative of the unbiased genome-wide genetic variation in wolves across Southeast Alaska, and our approach can also be easily applied to samples collected in the future, and to samples collected across the species' worldwide range.

Appendix 2: Map of Prince of Wales Island and nearby islands.

Appendix 3: Genotype Likelihoods & Identifying Runs of Homozygosity

To obtain the genome-wide SNP genotypes required for accurate identification and characterization of runs of homozygosity, we generated SNP genotype likelihoods for samples within each group. We used this approach to minimize the likelihood of falsely identifying variable sites in each population (i.e. sites that are fixed between two populations would appear polymorphic when individuals from those populations are treated as a single population), which could erroneously inflate heterozygosity at each site and falsely separate long runs of homozygosity into shorter runs.

To determine genotype likelihoods for all samples we used GATK, which employs the approach described by Li (2011). This approach uses a Bayesian framework informed by base quality, mapping quality, and depth of coverage each individual locus, and by comparison to DNA sequence data from other samples at the same locus, to calculate the most probable genotype for an individual (Broad Institute 2016). The simplified formula for this framework, provided by the Broad Institute (2016), is

$$P(G|D) = \frac{P(G)P(D|G)}{\sum_i P(G_i)P(D|G_i)}$$

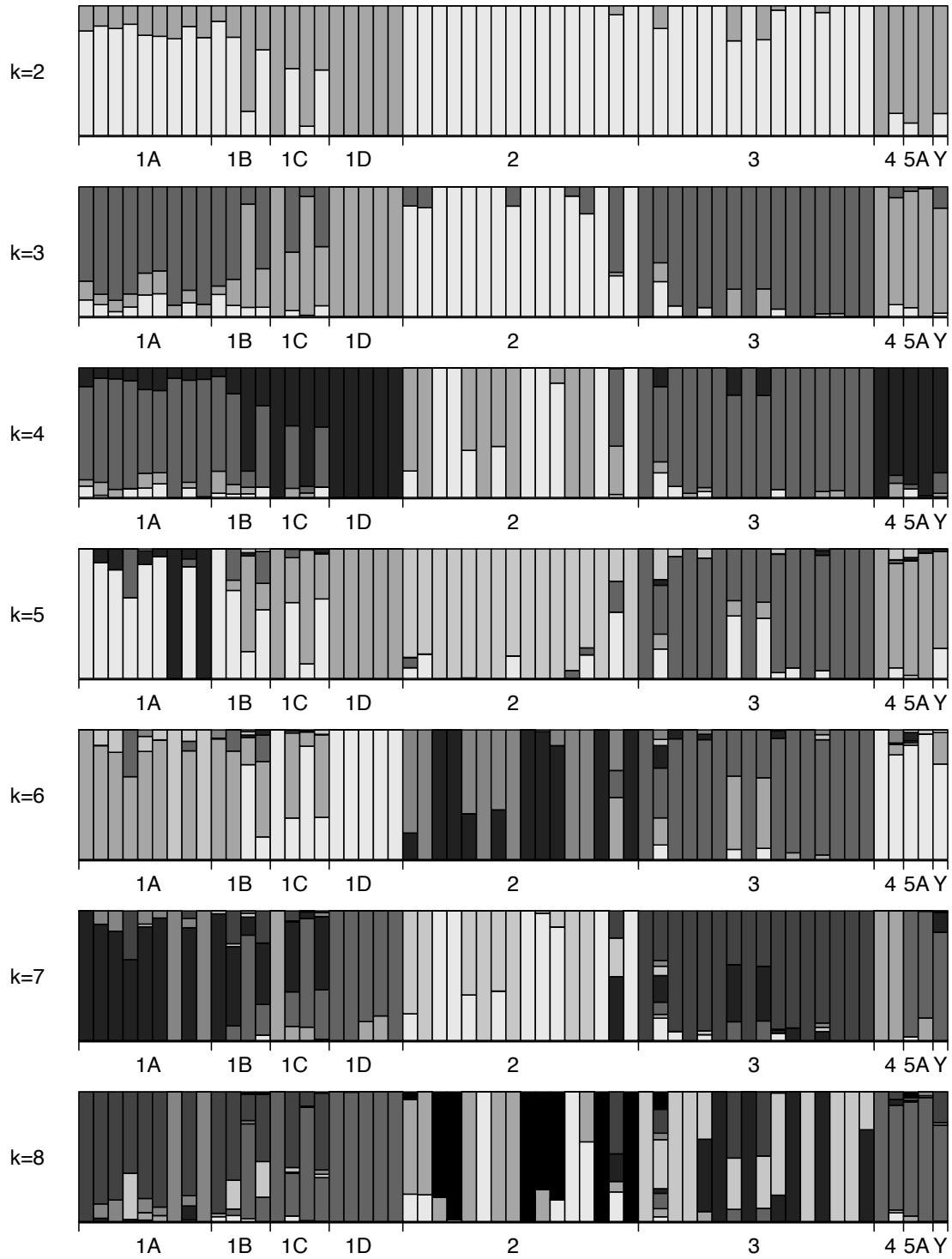
where $P(G|D)$ is the conditional probability of the genotype *G* given the observed data *D*. In the numerator, $P(G)$ is the probability that we expect to see genotype *G* based on previous observations (in GATK, this is a flat value with a default of $1/(2 * N + 1)$ where *N* is the number of samples), and $P(D|G)$ is the conditional probability of the observed data *D* given that genotype *G* is the true genotype (Broad Institute 2016). The denominator is the same for all samples, so the numerator is what determines the genotype likelihood for each sample (Broad Institute 2016). Using genotype likelihoods increased the number of called genome-wide SNP genotypes tenfold ensuring high resolution data that was better suited for characterizing runs of homozygosity.

To identify the locations and lengths of runs of homozygosity using genotype likelihoods, we separated samples into Admixture-defined groups (*Genotype Likelihoods & Identification of* section in Materials & Methods). We took this approach to minimize the number of false-positive detections of homozygous loci. By splitting samples into Admixture-defined groups, we avoided identifying sites that were fixed within but variable among populations as variable sites *within* populations, which would have artificially inflated the number of variant sites within each population, and could subsequently cause false identifications of runs of homozygosity when loci were actually fixed within a population.

To identify ROH, we used a custom R script adapted from (Kardos et al. 2017) for compatibility with genotype likelihoods. This script used a sliding window approach to evaluate the probability of each observed genotype under the assumption that the locus in question was within a ROH, and the probability of the observed genotype under the assumption that the locus is not in a ROH (probabilities calculated using equations from Wang et al. 2009). The script then calculated a "logarithm of the odds (LOD) score for each window of SNPs by summing the $\log_{10}$ of the ratio of these probabilities across all loci within the window" (Kardos et al. 2017). The script utilized a Gaussian kernel density plot (following Pemberton et al. 2012) to visualize LOD scores for all

individuals across all genomic windows. The resulting curve was bimodal, with loci in runs of homozygosity having higher LOD scores and clustering to the right of the local minimum, and loci not in runs of homozygosity having lower LOD scores and clustering to the left of the local minimum. By identifying the local minimum, the script then classified each window as either being in a ROH or not, and concatenated adjacent ROH. The resulting output file contained the locations and lengths of ROH across the genome for all individuals analyzed.

Appendix 4. Seven Admixture plots showing proportion ancestry assignment for each Alexander Archipelago wolf (*Canis lupus ligoni*) assuming two through eight ancestral populations. Each vertical bar represents one individual, and colored sections of each bar represent proportion ancestry for that individual as identified by Admixture. ADF&G Game Management Units labeled at base of each Admixture plot, with the exception of "Y" which represents one sample collected in Yukon Territory, Canada.



Appendix 5: Pairwise p-values for Tukey multiple comparisons of means tests for individual heterozygosity and F_{ROH} measured in three Alexander Archipelago wolf (*Canis lupus ligoni*) groups in Southeast Alaska and eleven wolves from Isle Royale National Park (Robinson et al. 2019). Three Southeast Alaska groups determined via majority ancestry assignment using Admixture and k = 3 (Alexander et al. 2009). POW group = Prince of Wales Island, Dall Island, Warm Chuck Island, Long Island, and Suemez Island (n = 15); Southeast group = Kuiu, Kupreanof, Mitkof, Prince of Wales, and Duke Islands and the mainland east of Lynn Canal (n = 31); Northwest group = Pleasant Island and the mainland west of Lynn Canal (n = 13); IRNP = Isle Royale National Park (n = 11).

Heterozygosity

	POW	Southeast
Southeast	0.334	
Northwest	0.021	9.51×10^{-5}

$F_{ROH \geq 100kb}$

	POW	Southeast	Northwest
Southeast	1.08×10^{-3}		
Northwest	1.36×10^{-7}	$<2.22\times10^{-16}$	
IRNP	3.00×10^{-3}	1.13×10^{-9}	0.0871

$F_{ROH \geq 1Mb}$

	POW	Southeast	Northwest
Southeast	9.54×10^{-3}		
Northwest	1.52×10^{-6}	4.48×10^{-3}	
IRNP	0.121	0.986	0.013

$F_{ROH \geq 10Mb}$

	POW	Southeast	Northwest
Southeast	4.08×10^{-4}		
Northwest	2.17×10^{-4}	0.675	
IRNP	0.883	1.43×10^{-4}	7.06×10^{-5}

Appendix 6: Photos of Alexander Archipelago wolves (*Canis lupus ligoni*) with short tails. (**A**) Short tailed suspected male wolf photographed on Lulu Island (west of Prince of Wales Island), photo by Mike Douville of Craig, AK. (**B**) Short tailed radio-collared female (left) with mate (right) photographed on Prince of Wales Island by Kris Larson (Alaska Department of Fish & Game). (**C**) Short tailed male (right) and lactating female (left) photographed at Hessa Inlet on Prince of Wales Island by Michael Kampnich (The Nature Conservancy). (**D**) Short tailed wolf (sex unknown) photographed on Dall Island (west of Prince of Wales Island) by Michael Kampnich (The Nature Conservancy). Note that the Lulu Island wolf (**A**) and Dall Island wolf (**D**) have similar tail deformities and could be the same wolf, but both have been observed several times by Michael Kampnich and Mike Douville and the Lulu Island wolf is brown and tan, while the Dall Island wolf is gray (Michael Kampnich, personal communication October 30, 2019).


2017-05-19 6:46:00 PM M 1/3 55°F
B
HC600 HYPERFIRE
RECONYX


2017-06-12 6:35:33 PM M 3/3 62°F
C
13 TNC
RECONYX


2017-09-04 8:53:04 AM M 1/3
62°F
D
16 TNC
RECONYX