ASSESSING FINE-SCALE POPULATION STRUCTURE USING RAD SEQUENCING IN A PHILOPATRIC SEABIRD, THE ATLANTIC PUFFIN (*FRATERCULA ARCTICA*)

by

Amy Catherine Wilson

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science

Cognitive and Behavioural Ecology, Faculty of Science

Memorial University of Newfoundland

August 2023

St. John's, Newfoundland and Labrador

ABSTRACT

Dispersal and philopatry play key roles in population structure and gene flow for vagile species such as seabirds. Seabirds, including the Atlantic puffin (*Fratercula arctica*), are known to have strong breeding philopatry, but due to challenges of studying pre-breeding individuals, little is known about their natal philopatry or dispersal. My thesis investigated Atlantic puffins on Gull Island, Newfoundland and Labrador, using restriction site associated DNA (RAD) sequencing to assess: if fine-scale genetic population structure is detectable at the colony level, if there are differences in dispersal between the sexes, and the level of relatedness among individuals. I predicted that natal philopatry should generate genetic population structure at the colony level. Blood from Atlantic puffins in eight plots was collected, extracted, and sent for sequencing. Genetic admixture, principal component analysis (PCA), and Mantel tests were used to assess structure. The admixture and PCA found no evidence of genetic differentiation between plots or sex and the Mantel test found no significant correlations between genetic and geographic distance. I found little evidence of population structure within the colony, which suggests there may be natal dispersal in this species. Natal dispersal may be an important life history trait for maintaining gene flow in species.

GENERAL SUMMARY

Not much is known about the natal dispersal patterns in Atlantic puffins (*Fratercula arctica*) due to juveniles living at sea for the first four years and nesting in burrows in remote locations once they are mature. Adult puffins show strong breeding philopatry, returning to the same burrow annually. I investigated whether these birds also show strong natal philopatry, returning to their natal colony to breed. I did this by taking blood from puffins in eight plots on Gul Island, in the Witless Bay Ecological Reserve, Newfoundland and Labrador, extracting the DNA, and sending it for restriction site associated DNA sequencing. I used this to examine genetic population structure on the colony. I found little evidence for population structure within the colony, which means this species may disperse more regionally as juveniles and display strong philopatry only as a breeding adult.

Acknowledgments

I thank my supervisor Pierre-Paul Bitton for believing in me and bringing me on to work on this project and working tirelessly through the COVID-19 pandemic to ensure that I and all his students could continue with their fieldwork. Finishing this degree would not have been possible without the tremendous amount of support he has provided both in the field and academically.

Thank you to my committee members Dave Wilson and Dawn Marshall for offering advice through the development and implementation of this project and special thanks to Dawn for allowing me to use her lab space and supplies, without which none of my lab work would have been possible. I also want to thank Julissa Roncal who took the time to meet with me and helped me figure out a path forward with my sequencing data.

Thank you also to my fellow lab mates, K. Kochvar, R. Zabala, F. Le Taro, A. Morel, and J. Taylor, as well as the Montevecchi lab members S. Collins, R. Blackmore, and K. D'Entremont who helped me on long nights in the field and in the lab. I would also like to thank the field assistants that helped with field work L. Lake, S. Watkins, and N. O'Brien.

I would also like to thank J. Wight who spent hours of his own time teaching me, answering questions, and checking up on me in the lab. Other thanks as well to E. Langille and all of the members of Andrew Lang's lab that let me use their equipment and lean on their lab for support.

Thank you Q. Carvey and H. Major at the University of New Brunswick for collaborating with me and providing the samples for my outgroup.

Thanks to the team at IBIS and Laval University for my sequencing prep and for working with me to complete my bioinformatics, without which I would have never been able to complete this project in the time I had.

Thank you to V. Friesen, D. Sauve, L. Gervais, and Z. Farrand who took time out of their busy schedules to help me answer questions about my analysis.

I give thanks to my family and friends who supported me on this journey. Special thanks to my friends S. Collins, who has given me endless advice, support, and notes, and K. Kochvar who spent countless hours in the field with me, and of course my partner Forrest who has supported me in every way possible throughout this process.

The funding for this work was provided by NSERC Discovery Grant and MUNL start-up funds. Thank you to Environment and Climate Change Canada (ECCC) for our access to the cabin on Gull Island and to O'Brien Boat Tours for providing transportation to and from the island in 2020 during the pandemic.

TABLE OF CONTENTS

LISTS OF FIGURES

LIST OF TABLES

INTRODUCTION

Seabirds are a commonly-used example of a taxon that displays philopatry (Breton et al., 2006; Coulson, 2016; Ibarguchi et al., 2011). Long-term banding and resighting studies have identified strong breeding philopatry in seabird species such as puffins, murres, petrels, and albatrosses (Blackmer et al., 2004; Breton et al., 2006; Harris, 2012; Ibarguchi et al., 2011; Mancilla-Morales et al., 2022; Patrick et al., 2017). Even though seabirds are often referred to as philopatric, natal philopatry is far less studied than breeding philopatry (Breton et al., 2006; Sandvik et al., 2008). Some of the reasons for this include a lack of commitment to long-term studies that monitor multiple sites and band large numbers of juveniles (often due to logistical and/or financial constraints), colonies that contain large numbers of individuals in often inaccessible locations, animals that are long-lived with long maturation times that are difficult to track over many years, and difficulty teasing apart dispersal and mortality (Breton et al., 2006; Coulson, 2016; Hatchwell, 2009; Ibarguchi et al., 2011; Morrison et al., 2009; Sandvik et al., 2008). Due to these challenges, our knowledge of patterns of natal philopatry and dispersal in seabirds is poor. Assumptions about natal philopatry are often influenced by studies of breeding philopatry, but the factors that influence natal philopatry may be very different from breeding philopatry. Resolving these questions would benefit our understanding of the circumstances under which dispersal and philopatry occur. Patterns of dispersal influence the long-term demography and movement of seabird species across their range, which has consequences for management and conservation efforts globally.

Dispersal

Natal dispersal is the movement of an individual from the site of nascency to the site of reproduction (Baguette & Van Dyck, 2007; Bonte et al., 2012; Clobert et al., 2012; Genovart et al., 2013), and is the main mechanism that leads to gene flow between populations (Bonte et al., 2012; Clobert et al., 2012; Perrin & Lehmann, 2001; Ronce, 2007). This movement is often broken down into three distinct stages: departure (leave the place of origin), transience (move around for some amount of time), and settlement (take residence in a new location to breed) (Baguette & Van Dyck, 2007; Clobert et al., 2012; Ronce, 2007). Movement between successive breeding sites is considered breeding dispersal rather than natal dispersal (Baguette & Van Dyck, 2007; Clobert, Galliard, Cote, Meylan, & Massot, 2009; Matthysen, 2012).

Natal dispersal strongly influences individual fitness by reducing interactions with close kin, thereby preventing inbreeding (Ludwig & Becker, 2012), and by dispersing offspring from the same parents to other localities with variable conditions (Clobert, 2012), which impacts genetic population structure. Genetic population structure forms when there are differences in allele frequencies between subpopulations due to non-random mating, which can occur for many reasons including environmental variation (selection), genetic drift, and dispersal patterns that restrict gene flow (Hellwege et al., 2017). The main factors that drive dispersal are variations in temporal and spatial habitat quality, which influence the benefit or cost of dispersing for individuals (Baguette et al., 2013; Ronce, 2007). Dispersal can also be a conditional trait that increases individuals' fitness by enabling them to escape locally unfavorable conditions (Clobert et al., 2012).

Since natal dispersal is integral to almost all forms of life, there are many strategies used by different organisms, due to variation in selective pressures leading to a multitude of

evolutionary solutions. Dispersal strategies range from passive dispersal (e.g., seeds dispersed by wind, zooplankton through ocean currents) to more complex active dispersal (e.g., movement from natal to breeding ground, the establishment of home ranges, joining a social group) (Baguette & Van Dyck, 2007; Clobert et al., 2012; Ronce, 2007). By utilizing ocean currents to transport its seeds over long distances, species such as the coconut palm (*Cocos nucifera*) have achieved one of the widest ranges of any plant through passive dispersal (Harries & Clement, 2014). In contrast, many large carnivores such as the mountain lion *(Puma concolor)* are obligate dispersers, where males will often travel more than 100 km from their natal site before establishing a territory (Andreasen et al., 2012; Holderegger & Gugerli, 2012).

Intraspecific variation in dispersal strategies has also been demonstrated between individuals, sexes, and populations within a species. For example, in Thick-billed murres (*Uria lomvia*), females have been shown to have bimodal dispersal distribution, where approximately 80 % of females are philopatric and breed at the same location they hatched, and 20 % disperse to other areas (Steiner & Gaston, 2005). In mammals, females often display philopatry or limited dispersal while males are more often the dispersing sex (Clutton-Brock & Lukas, 2012). Orca (*Orcinus orca*) subpopulations are known to display different strategies where some are resident year-round in one area, while others migrate seasonally (Filatova et al., 2019). Within the same population, individual Salt Marsh Wolf spiders (*Pardosa purbeckensis*) will either balloon and disperse, by climbing to the top of a high object and producing a long piece of silk to catch the wind that can move them many meters to kilometers away, or build a territory where they were hatched (Bell, 2005; Bonte et al., 2007, 2012). All these strategies can increase fitness by adding variation in the ways individuals interact with their environment. This can both strengthen

existing populations or help initiate the formation of new populations when conditions are favourable.

Philopatry

Natal philopatry is where an organism either stays close to or returns to its place of nascency to breed (Perrin & Lehmann, 2001; Shields, 1982). This strategy is seen in many taxa including some salmonids (Keefer & Caudill, 2014), sharks (Mourier & Planes, 2013), amphibians (Helfer et al., 2012), mammals (Clutton-Brock & Lukas, 2012), and birds (Andersson & Waldeck, 2007; Coulson, 2016; Ibarguchi et al., 2011; Komdeur, 1994a). While natal philopatry is most common, breeding philopatry, where an animal breeds in the same location year after year, is frequently observed especially in mammals and birds (Clutton-Brock & Lukas, 2012; Coulson, 2016; Shields, 1982). Philopatry can be an important strategy for increasing individual and inclusive fitness benefits through the opportunity for social learning and interactions with kin and increasing chances to inherit limited high-quality territories, as well as the "good location" theory, that if an individual survives to first breeding, the location they were raised must have been of high enough quality to support survival (Clutton-Brock & Lukas, 2012; Groenewoud et al., 2018; Komdeur, 1992, 1994b; Perrin & Lehmann, 2001; Stacey & Ligon, 1991). This strategy restricts gene flow and can lead to increased local adaptation, in which individuals that exhibit natal philopatry may have higher fitness than dispersing individuals if the environment is variable across their range (Kawecki & Ebert, 2004; Mobley et al., 2019).

In addition to fitness benefits, philopatry can promote genetic population structuring through the formation of kin groups that lead to non-random mating, which can lead to faster rates of local adaptation and evolution (Ibarguchi et al., 2011; Kingston & Rossiter, 2004).

Population structure has been detected in many species of philopatric animals, including the Thick-billed murre where fine-scale structure was demonstrated on breeding ledges in the Canadian Arctic and Norway (Birt-Friesen, 1992; Ibarguchi et al., 2011). In their study, ten breeding ledges in six subareas were sampled over a 205 m cliff face, with ledges ranging from 5-80 m apart from the next closest ledge. Through the use of microsatellite analysis, it was found that there were detectable genetic differences between individuals on different ledges, and birds on neighbouring ledges were more genetically similar than birds on more distant ledges. This resulted in a detectable population structure from east to west along the cliff face. Population structure can also be influenced at the landscape level by physical barriers that inhibit gene flow; these can be from natural sources such as mountain ranges, canyons, or water bodies or manmade barriers such as highways, dams, and urban development (Baguette et al., 2013; Holderegger & Gugerli, 2012). These barriers to gene flow can lead to the development of source-sink relationships, divergence, speciation, or genetic bottlenecks (Andreasen et al., 2012; Baguette et al., 2013; Lee, 2012).

There can be disadvantages to philopatry as well, including increased competition for breeding sites, which can limit breeding success by forcing some individuals into more marginal habitats, inbreeding depression, and the inability to adapt to a changing environment (Clutton-Brock & Lukas, 2012; Kruuk et al., 2002; Perrin & Lehmann, 2001; Shields, 1982). Inbreeding depression, as seen in Collared flycatchers (*Ficedula albicollis*), significantly increases the expression of lethal equivalents (loci that if homozygous result in death) found in the inbred offspring which leads to higher rates of mortality in this group (Kruuk et al., 2002). The lack of genetic diversity can lead to the formation of sink populations when environmental conditions change and there is a lack of strong local adaptation, this can be particularly damaging when

habitat loss and fragmentation prevent immigration into the population or the formation of new populations (Andreasen et al., 2012; Kawecki & Ebert, 2004).

Given that there are both benefits and costs to natal philopatry, the degree to which it is expressed, from strong to weak, can vary between species, populations, sexes, and individuals (Clutton-Brock & Lukas, 2012; Coulson, 2016; Helfer et al., 2012; Shields, 1982). Variation in philopatric strategy can range from strong, where individuals may disperse very little to not at all, to weak, where individuals may disperse away from their natal site while remaining in the same region and population (Baguette et al., 2013; Clutton-Brock & Lukas, 2012; Coulson, 2016). Many mammals and some birds show strong sex differences in philopatry (Birt-Friesen, 1992; Clutton-Brock & Lukas, 2012; Ibarguchi et al., 2011; Komdeur, 1994b). For example, Bornean orangutans (*Pongo pygmaeus*) display strong sex-biased philopatry where the female offspring will set up territories directly adjacent to their mother's territory, so that a ring of sisters' territories will surround their mother's territory (Ashbury et al., 2020). Strong philopatry can be the result of lack of suitable habitat, which is the case with Seychelles warblers (*Acrocephalus sechellensis*), where females do not disperse at all, but rather become nest helpers for their mother until she dies and the next female in line takes over the territory (Groenewoud et al., 2018; Komdeur, 1994b). In both these cases, the males tend to disperse to other territories. Other animals such as salmonids (*Salmonidae* spp.), Blacktip reef sharks (*Carcharhinus melanopterus*), sea turtles (*Chelonioidae* spp.), and mammals such as Yellow-bellied marmots (*Marmota flaviventris*), will return to the locality (stream bed, coral reef, beach, etc.) where they were born, but will not breed in the exact same location (Armitage et al., 2011; Baltazar-Soares et al., 2020; Clutton-Brock & Lukas, 2012; Keefer & Caudill, 2014; Mourier & Planes, 2013). At the weak end of the philopatric scale, animals such as some passerines and ungulates will return

to the same general area or region of their birth but not at a rate that would create population structuring (Clutton-Brock & Lukas, 2012; Perrin & Lehmann, 2001; Shields, 1982).

Study Species

The Atlantic puffin (*Fratercula arctica*) is a member of the family Alcidae and is found throughout the North Atlantic where they breed colonially from the Gulf of Maine to Labrador in the western Atlantic and Western Europe northward to Russia and southward to France in the eastern Atlantic (*BirdLife International*, 2021; Harris, 2012). A recent study has shown that Atlantic puffins have complicated genetic structure throughout their range (Kersten et al., 2020). They are genetically and socially monogamous with a low divorce rate, often staying with the same partner for multiple years (Anker-Nilssen et al., 2008; Harris, 2012). They breed on colonies, which are formed on isolated marine islands or ledges with open habitat in which they dig burrows. Both adults work cooperatively to produce one chick per season, which once fledged, will not return to breed for three to five years. Outside of the breeding season, which ranges from May to September depending on location, both adults and juveniles spend the majority of their time in the open ocean (Breton et al., 2006; Harris, 2012; Sandvik et al., 2008).

Delayed return to a breeding colony for several years after fledging makes the Atlantic puffin particularly difficult to study when trying to determine natal philopatry. While breeding philopatry has been well established in this species, numerous studies have highlighted the difficulty of tracking natal philopatry, dispersal, and survival in juvenile puffins, due to the reasons mentioned above, as well as long maturation time and difficulty in resighting burrow nesting banded birds (Breton et al., 2006; Kress & Nettleship, 1988; Morrison et al., 2009; Sandvik et al., 2008). Morrison et al. (2009) resighted only 26 % of Tufted puffin (*Fratercula*

cirrhata) chicks banded between 1999 and 2000 on their natal colony between 2002-2008, even with intensive yearly monitoring of a relatively small population of ~12,000 pairs.

Due to the difficulty in monitoring the movements of juveniles and resighting them once they return to breed, it is unclear where Atlantic puffins fall on the natal philopatric scale. While it is not contested that puffins display natal philopatry to some degree, the spatial scale and percent of the population at which it occurs are unknown (Breton et al., 2006; Harris, 2012). In this study, I tested if fine-scale genetic population structuring within a single colony of Atlantic puffins could be detected, which could potentially provide evidence for natal philopatry. To do this, I collected blood samples from puffins breeding at several locations on Gull Island in the Witless Bay Ecological Reserve, Newfoundland and Labrador, Canada, and genotyped the puffins using restriction site associated DNA (RAD) sequencing to determine if genetic population structure was present.

Study Site and Population

The Witless Bay Ecological Reserve is found off the eastern coast of Newfoundland, Canada. The reserve supports over 1 million breeding seabirds from 10 species during the summer months including the largest breeding population of Atlantic puffins in the western Atlantic (Wilhelm et al., 2013; *Witless Bay Ecological Reserve Management Plan*, 1998).

The study population consisted of breeding adult Atlantic puffins on Gull Island (47.26N, 52.77W), which is one of four islands in the Witless Bay Ecological Reserve. This colony supports approximately 250,000 Atlantic puffins. It has a steep rocky coast comprised mostly of cliffs that give way to grassy slopes used by the puffins. The center of the island is covered in dense understory ground cover, ferns, shrubs, and coniferous trees. The grassy slopes form a continuous ring of habitat (20 – 75 m width) around the island.

This ring of habitat makes Gull Island uniquely suited to assess fine-scale population structure as other studies have shown that genetic structure can form within a population when suitable habitat forms a ring (Alcaide et al., 2014; Bouzid et al., 2022). This phenomenon has been documented in the Greenish warbler (*Phylloscopus trochiloides*) where speciation and genetic clustering have formed over thousands of years along a ring of continuous habitat that spans 1500 km around the Tibetan plateau (Alcaide et al., 2014). We also know from examples such as the Thick-billed murre that genetic differentiation can occur at a much smaller scale, including at the scale of 250 m (Ibarguchi et al., 2011).

Method of Determining Relatedness

When developing this project, I planned on using microsatellite analysis to determine population structure. Microsatellites are highly variable tandem repetitive segments of DNA that can be used to measure relatedness, as well as genome variation, evolutionary processes, and population structure (Queller et al., 1993). This method relies on the amplification of these markers using polymerase chain reactions (PCR), which requires very little genetic material to work effectively. Based on this initial plan, I collected a small amount of blood from each puffin on a filter paper. Further exploration, however, revealed that too few microsatellites had been developed for puffins to address my questions at the colony scale (Dawson et al., 2005). Instead genomic sequencing allows for finer scale analysis than is possible using microsatellites, which typically use only 8 to 25 markers that must be developed for each species in a tedious and time consuming process (Arthofer et al., 2018; Dawson et al., 2005; Hatchwell, 2009; Moen, 1991; Roques et al., 2019; Vendrami et al., 2017). Therefore, I used double-digest restricted site-associated DNA (ddRAD) sequencing, which is a technique that uses restriction enzymes to

isolate short segments of the genome, which are then sequenced and used to identify single nucleotide polymorphisms (SNPs) (Andrews et al., 2016; Puckett, 2017). This process can identify thousands of unique SNPs, which can then be used to evaluate population structure, genetic relatedness, and many other population genetic metrics.

Unlike traditional techniques of determining relatedness between individuals, such as band resighting and the construction of pedigrees, genomic techniques do not require years of intensive monitoring, banding, and recording of breeding histories to provide detailed individual data (Breton et al., 2006; Ludwig & Becker, 2012). ddRAD sequencing is a relatively new technique that has become increasingly popular in recent years to address biological questions in many species and areas of research (Andrews et al., 2016; Çilingir et al., 2022; Ferrer Obiol et al., 2021; Vendrami et al., 2017; Wang, 2017; Winters et al., 2019). These include determining fine-scale population structure in Threespine sticklebacks (*Gasterosteus aculeatu*), evaluating gene flow in a highly endangered species, the Aldabra giant tortoise (*Aldabrachelys gigantea*), and detecting cryptic speciation in Gentoo penguins (*Pygoscelis papua*) (Çilingir et al., 2022; Pedersen et al., 2017; Pertierra et al., 2020). ddRAD sequencing has also been shown to be more sensitive at detecting subtle genetic structure in Great scallops (*Pecten maximus*) than other methods such as microsatellite analysis, even with a robust microsatellite array (Vendrami et al., 2017). This type of sequencing is also more cost-effective and faster than whole genome sequencing, while still allowing for the analysis of thousands to millions of SNPs.

Study Overview

Previous research by Kersten et al. (2020) explored genetic structuring of the Atlantic puffin throughout its range using whole-genome sequencing. Using a few individuals (72 total) from many colonies (12) across the complete breeding range they were able to assess broad

trends in the puffin genome on a large spatial scale. Building upon this study, I explored structuring at a much finer scale. This was accomplished by collecting blood samples for genetic analysis from 353 individual puffins over two years from the Gull Island colony. These samples were collected from eight geographically distinct plots so that genetic structure within the colony could be explored. ddRAD sequencing was used to identify SNPs that were used to calculate pairwise genetic distances that were then compared to the geographic locations of the burrows to assess whether geographic distance correlated with genetic distance. The individuals were molecularly sexed to assess whether any differences in fine-scale sex dispersal could be detected.

Hypothesis and Research Questions

In this thesis, I asked the specific following questions:

- Is there genetic population structure in the puffin colony on Gull Island, NL that can be detected using ddRAD sequencing?
- Is there a relationship between geographic distance and genetic distance?
- Is there evidence for a difference in dispersal between the sexes?
- Is there a pattern of relatedness within the colony?

I hypothesize that over time natal philopatry should lead to the presence of genetic population structure in puffins on Gull Island. This should result in a positive correlation between geographic and genetic distance, and individuals within plots will be more related than individuals in more distant plots. I also hypothesize that there may be a difference in dispersal rates for females based on evidence in closely related species, resulting in genetic structure differences between the sexes.

METHODS

Site Selection

On Gull Island, eight plots were selected from puffin breeding areas around the perimeter of the colony (Figure 1). Plots were distributed as evenly as possible with an average of 350 meters (range: 250 – 475 m), straight-line measure, between adjacent plots. They represented all cardinal directions (slope aspect) as well as marine and inland exposure, which may be a factor in the formation of population structure, as seen in another study (Çilingir et al., 2022). The distance was greater between plots than the spacing used in similar studies (Cristofari et al., 2015; Garrett et al., 2020); plots were selected to maximize the distance between them to detect population structure if present. Because of the workload needed to complete my project, sampling took place from 2019 to 2021. One plot was sampled in 2019 while developing field method protocols (plot A/Landing), four plots were sampled in 2020, including resampling of the test plot (plots A/Landing, B, C, and D), and four more plots were sampled in 2021 (plots E, F, G, and H). In 2021, three of the four plots sampled in 2020 were resampled to provide a more robust dataset (plots A/Landing, B, and D).

In each plot, 24-35 burrows were marked and, when possible, both adults were sampled. In 2020, pairs were sampled at the same time, if both were present in the burrow, but in 2021, only one adult was sampled to minimize disturbance. In addition, 14 chicks with at least one presumed parent from four plots were sampled to provide a baseline comparison for relatedness coefficients.

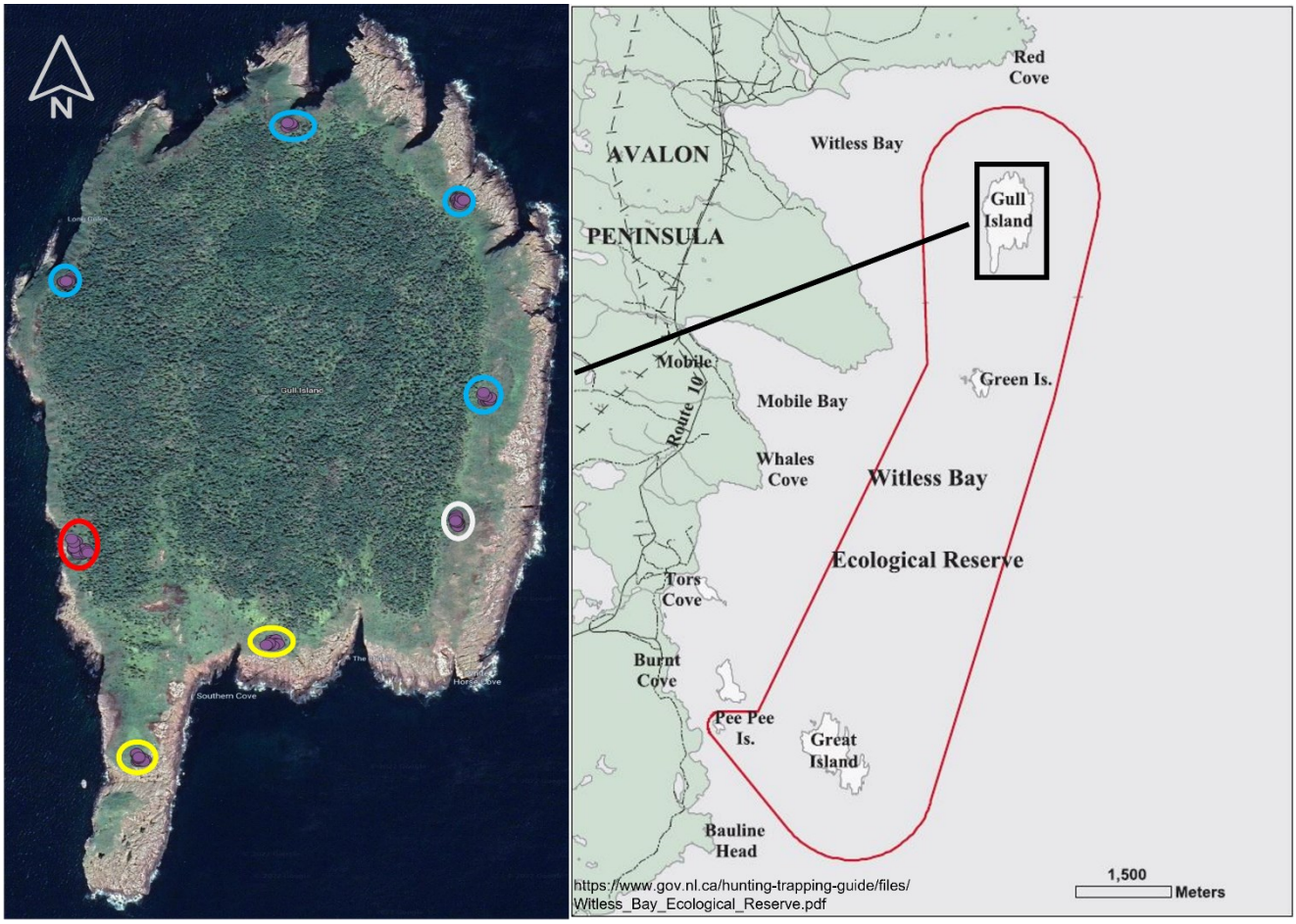

Figure 1. Gull Island is shown as part of the Witless Bay Ecological Reserve, Newfoundland and Labrador, Canada, and enlarged to show the eight sample plots colour coded by years sampled. The red plot was sampled in 2019, 2020, and 2021, yellow plots were sampled in 2020 and 2021, the white plot was sampled in 2020, and blue plots were sampled in 2021.

Disturbance Mitigation and Permits

Given that Atlantic puffins are sensitive to researcher disturbance (Kelly et al., 2015, Rodway et al., 1996), I minimized my impact during the plot selection process. Established researcher trails were used whenever possible to avoid disturbing birds and habitat. Because Atlantic puffins may be more likely to abandon their burrow when disturbed during incubation (Rodway et al., 1996), I selected only burrows with chicks for the study, which was determined based on known phenology of when chicks are typically present and confirmed by hand before

the adult was removed. At night, adult puffins are most likely to be in their burrows and very few birds are present on the slope, so sampling of adults took place at night to reduce flushing and increase efficiency. Juveniles were chosen based on size and age and sampled during the day when the adults were not present in the burrow. To minimize researcher-induced stress and potential abandonment, burrows visited more than once were never sampled on consecutive nights, with a minimum of two nights between visits. Bird handling protocols were designed to be as streamlined and efficient as possible to minimize stress on the bird, reduce the overall time they were kept outside of the burrow, and reduce the amount of time my assistants and I spent on the slopes. Captures were only conducted on nights without rain or extreme weather to not thermally tax the birds. All proper Animal Use permits including university (#19-02-DW), provincial (SERAC, WBER2019), and federal (Banding - #10926, Migratory Bird Research Permit -#SC4061) were obtained before the commencement of fieldwork.

Data Collection

Field Methods

Adults were removed one at a time from their burrows and brought to a workstation where they were banded with a Canadian Wildlife Service steel leg band. Morphometric data including weight and wing cord were measured, and a small blood sample (approximately < 70 µL) was collected from the brachial vein using a 26.5-gauge needle which was dabbed on a Whatman cellulose filter paper. The filter paper was hung up to dry and each was stored in a separate plastic bag at room temperature out of direct sunlight for several months before lab work could be completed. Following each field season, georeferencing coordinates were recorded for sampled burrows using a Trible Geo7x^{R} global positioning system (GPS) device with an accuracy of ~10 cm. These locations were then calculated into a pairwise matrix of

Euclidean geographic distances in program R (R Core Team, 2022) using the function "dist" to convert the points to Euclidean distances and the function "as.matrix" to convert it to a matrix using the base R *stats* package. Euclidean distance was used because the size of the colony (~1 km^2) allowed for the puffins to easily access any part of the island by a straight-line flight path.

DNA Isolation and Sexing

DNA was isolated from the blood samples using a DNeasy® Blood & Tissue Kit (Qiagen Inc., Toronto, ON, CA) following a protocol modified from the DNeasy® Blood & Tissue Handbook (2020), after the standard extraction protocol produced insufficient DNA concentrations to be used for RAD sequencing (Poland et al., 2012). I isolated DNA from a 5-10 mm radius circle of dried blood that was cut with sterilized scissors into very small pieces, which were collected in a 1.5 ml tube. Two hundred µL of lysis buffer was added to the tubes and allowed to incubate at room temperature for two hours. An additional 180 µL of lysis buffer and 30 µL of Proteinase K were added and the samples were incubated at 56˚ C for 40 minutes in an Eppendorf ThermoMixer®. Forty-five µL of elution buffer was used for 10 minutes of incubation. After the elution buffer was spun down through the column, a second elution step was performed by pipetting the elution buffer back into the column, allowing it to incubate for another 5 minutes, then spinning it down again. The final elution was then put into a speed vacuum for 45 to 60 minutes where it was further reduced to a total volume of 15-20 µL. These samples were then stored at -20 ˚C.

From the extracted DNA, PCR was used for sexing the individuals using standard protocols for sexing birds (Çakmak et al., 2017; Fridolfsson & Ellegren, 1999). This was done using primers 2550F and 2718R, which are used to amplify the chromo-helicase DNA 1 (CHD1) gene on the avian W and Z chromosomes, resulting in the presence of two bars for females and one

bar for males. The PCR was run on an Eppendorf Mastercycler® ep gradient S with a program of 95 °C for 5 minutes, 35 cycles of denaturing, annealing, and extension at 94 °C for 30 seconds, 50 °C for 30 seconds, 72 °C for 60 seconds, respectively. Followed by extension at 72 °C for seven minutes and a cooling period of 4 °C for 10 minutes. PCR products were separated by gel electrophoresis on a RedSafe™ or SyberSafe™ -containing 5 % agarose gel with 100 base pair reference ladders using a Thermo Scientific™ EC 300 XL apparatus for 50 minutes at 130 amps. Gels were imaged using Image Lab software (version 6.1.0). All PCR batches contained a no template control that was run on each gel.

ddRAD-Seq Library Preparation & Sequencing

The need for a modified DNA extraction protocol was the result of switching from a microsatellite-based analysis to a RAD sequencing-based analysis. The RAD sequencing required a much higher concentration of DNA (10-20 ng/µL) than the microsatellite analysis would have (~1 ng/µL). This level of concentration can be very difficult to extract from dried blood, which is why whole liquid blood is preferred for this method. This resulted in a significant loss of samples due to a lack of sufficient blood collected and DNA concentration once extracted. Of the 353 samples, 228 had enough blood from which DNA could be extracted using the modified protocol, and 152 (146 adults and 6 chicks) were of high enough quantity to send for sequencing.

Samples were tested for DNA concentration using a Thermo Fisher Scientific™ Qubit 2.0 Fluorometer. The samples were then standardized to 10 µL of 10 ng/µL concentration DNA with PCR grade (DNA, DNase, and RNase free) H_2O on two clear 96-well plates. The samples were sequenced at the Institut de Biologie Intégrative et des Systèmes (IBIS) at Laval University in Quebec City, QC, Canada.

Libraries for Illumina triple digest GbS (3D-GbS) were prepared using the procedure described by (Poland et al., 2012), with modifications suggested by IBIS. The modifications consisted of the following: three restriction enzymes (PstI/NsiI/MspI) were used instead of the PstI/MspI combination, and a Blue Pippin (SAGE sciences) was used to size libraries before PCR amplification (elution set between 50 and 65 min, on a 2 % agarose gel). Plate barcodes were added according to the procedure described in Colston-Nepali et al. (2019) for double-digest RAD sequencing.

The 152 prepared samples from Gull Island were combined prior to sequencing with 40 samples of Atlantic puffin DNA from Machias Seal Island (provided by Q. Carvey and H. Major, University of New Brunswick) to serve as an outgroup. Samples from Machias Seal Island, in the Gulf of Maine, were chosen as a suitable outgroup because they were still from the Western Atlantic region, but from far enough away (1150 km) to hopefully have detectable genetic differences from the Gull Island samples. Puffin chicks from the Witless Bay ecological reserve were used to re-establish some colonies in the Gulf of Maine during the 1970s, which is another reason using samples from this area may yield interesting results when determining if RAD sequencing can detect differences between these individuals (Kress & Nettleship, 1988). While the colony on Machias seal Island was never fully extirpated, like some of the other Gulf of Maine colonies, several of the transplanted chicks from Witless Bay did show up on the colony (Kress & Nettleship, 1988). These samples were extracted from whole liquid blood preserved in ethanol and standardized to 20 ng/μL. One hundred ninety-two total samples were sent to Genome Quebec where 3 million 150-bp, paired-end reads per sample on the NovaSeq 6000^{R} (35M read blocks) were requested for the ddRAD sequencing protocol.

Bioinformatics

Bioinformatics were performed at IBIS, including alignment to the reference genome, estimation of sequencing depth, down-sampling, genotyping, and filtering through the STACKS pipeline (version 2.59; Catchen et al., 2013). IBIS also provided analysis output files from the filtered dataset including admixture, principal components analysis (PCA), and genetic distances and relatedness estimates, which I then analyzed and interpreted.

Alignment to Reference Genome, Estimation of Sequencing Depth, and Down Sampling

The STACKs pipeline was used to demultiplex, clean, and trim sequenced reads. The individual reads were aligned with a reference genome for the Atlantic puffin and indexed from the National Center for Biotechnology Information (NCBI:txid28701, GCA_905066775.2; Kersten et al., 2020). Data imputation based on the reference genome was used to fill in missing genotypes using the program BEAGLE 5.4 (Browning et al., 2018), after which loci were assembled to be able to call SNPs and genotypes. The SNPs were filtered based on percent missingness, percent allele coverage, and percent genotype data per population. This was done using protocols outlined in the STACKs Manual (Catchen et al., 2013; Wickland et al., 2017).

Admixture

The program ADMIXTURE (version 1.3.0; (Alexander et al., 2009)) uses multilocus SNP genotype data to estimate the maximum likelihoods of individual ancestries within the dataset. To investigate clustering within the colony, an ADMIXTURE analysis was run with K = 1 to K = 7, where K represents the number of populations, or potential ancestries, within the dataset. This was done by converting variant call format (VCF) files from STACKS into PLINK (version 1.90b5.4) format (Chang et al., 2015).

PCA

A PCA was conducted on the individual genotype data that was obtained from the VCF files using PLINK software (version 1.90b5.4; (Chang et al., 2015; Purcell et al., 2007). PCA can be useful to visualize clustering within the dataset that may not be as apparent with other types of modeling. I addressed whether there was population structure found within the colony by looking for relationships among the principal components, which capture variability among genotypes in this dataset.

Genetic Distances

Euclidean pairwise genetic distances, where the distance between individuals is the number of loci by which they differ, were assessed using the R package *adegenet* (Jombart, 2008) with the function "list" from a "genind" object which was created with R package *vcfR* (Knaus & Grünwald, 2017).

Mantel Test and Correlogram

A Mantel test, was conducted in R (R Core Team, 2022) with the package *ecodist* (Goslee & Urban, 2007) using the function "mantel" to test for a correlation between genetic and geographic distance. This function calculates correlations between two symmetrical pairwise matrices for all individuals in the dataset and produces a Mantel statistic, a correlation coefficient, and a p-value. I used a one-tailed p-value to determine if there was a positive correlation between a pairwise matrix of geographic distance, derived from the latitude and longitude of the burrow locations on the colony, and a pairwise matrix of individual genetic

distances. I used scatterplots to help determine if any trends were visible in the distance correlation matrix.

A correlogram was used to further visualize the data. Correlograms are similar to spatial autocorrelation and can be useful to find subtle trends in the data that are not reflected in the overall p-value. Breaks were set in the correlogram based on distance clusters of individual pair-wise comparisons that correlate to pair-wise plot distances. The correlograms were produced using the R package *ecodist* (Goslee & Urban, 2007) using the function "mgram".

Relatedness

Relatedness between parent and offspring has a known coefficient, r=0.5. Genotypes from parent/offspring pairs were used to calibrate relatedness coefficients and estimate relatedness between adults (Ibarguchi et al., 2011; Taylor & Friesen, 2012). Pairwise comparisons of individual relatedness coefficients were generated using *vcftools* (version 0.1.17) in R using the "relatedness" output option which is based on the method of Yang et al. (2010) and calculates an unadjusted A_{jk} statistic, which produces a continuous scale relatedness coefficient where a coefficient of zero is unrelated individuals within a population, up to 0.5 for a parent-offspring or full sibling relationship, and one for an individual with itself (Danecek et al., 2011). When individuals are not genetically related at all, when they have little to no overlap in minor allele frequencies, ceiling effect lead to negative coefficients. The relatedness coefficient calculation takes into account the likelihood that differences in locus frequencies between individuals are caused by inheritance.

RESULTS

Evaluation of Molecular Data

Of the 152 samples sequenced from Gull Island, two datasets were retained based on the percentage missing data for individuals and SNPs. These were: a larger dataset of 100 individuals with up to 70 % missing data per individual, the high missingness group, and a smaller dataset of 69 individuals with up to 30 % missing data per individual, the low missingness group. The high missingness group had an average of 22 % missing data per individual while the low missingness group had an average of 8.5 % missing data per individual. The datasets were further filtered with a minimum genotype coverage of three alleles and SNPs were retained if the genotyping rate was >70 % or >80 % of the samples with a non-missing genotype for each variant for the high missingness group and the low missingness group, respectively. No individuals had to be removed due to heterozygosity issues that violated the Hardy-Weinberg equilibrium. Individuals removed due to missing data were evenly distributed among sequencing plates, plots (Figure 2), and sex (Figure 3), so these factors did not contribute very much to whether an individual had a high percentage of missing data.

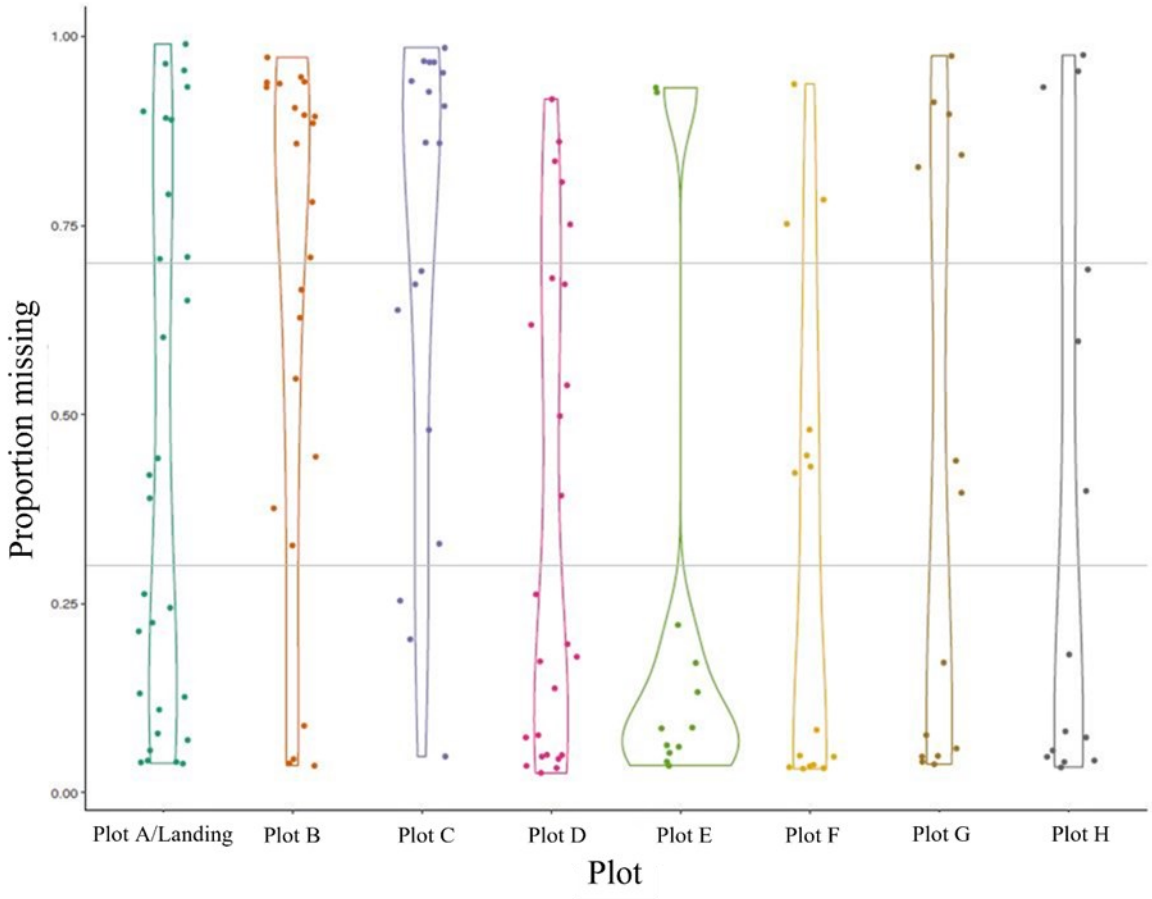

Figure 2. Distribution of all individuals sequenced (n=152) based on percent missing data per individual (Proportion missing) broken down by plot. This shows that missingness was spread fairly evenly throughout the plots (figure created by IBIS).

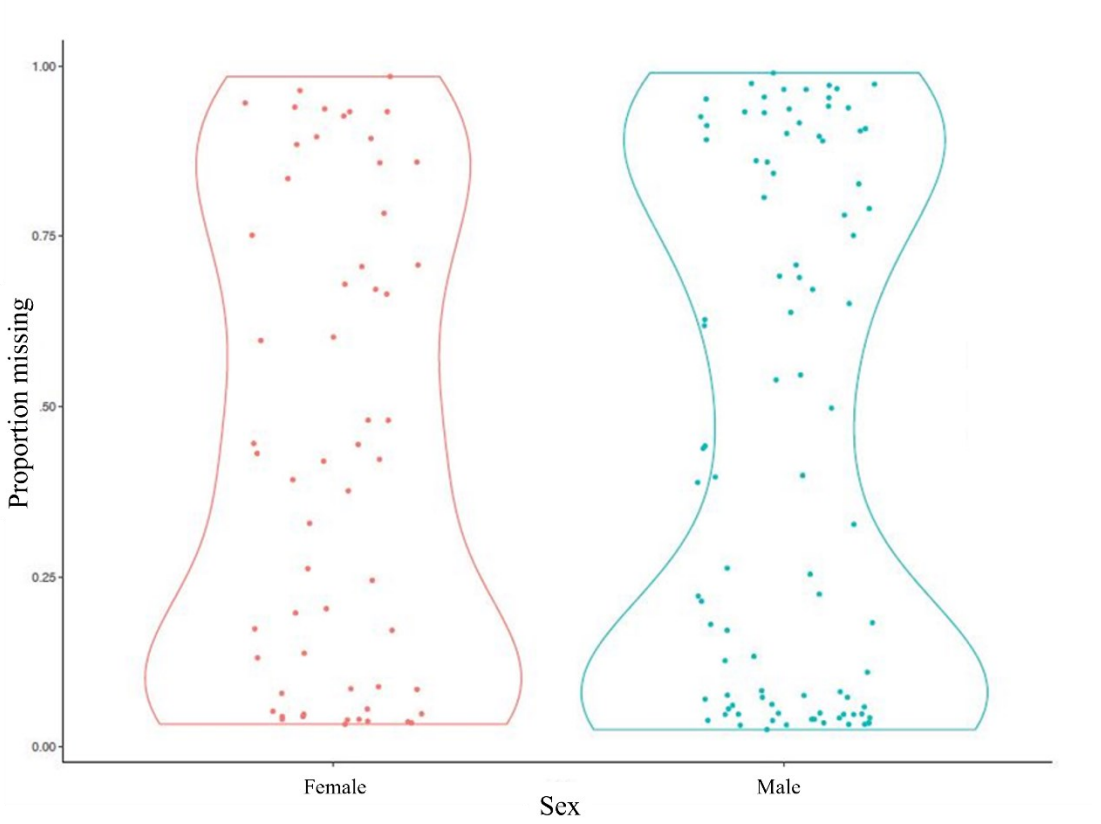

Figure 3. Distribution of all individuals sequenced (n=152) based on percent missing data per individual (Proportion missing) broken down by sex. (Figure created by IBIS)

Any duplicated and linked SNPs were removed, resulting in 2,714 shared SNPs retained for the high missingness group and 16,729 shared SNPs retained for the low missingness group. Both groups retained individual samples from each plot, including 8-21 per plot for high missingness,

and 3-15 per plot for low missingness. The high missingness group contained 59 males and 41 females. The low missingness group contained 43 males and 26 females (Figure 4).

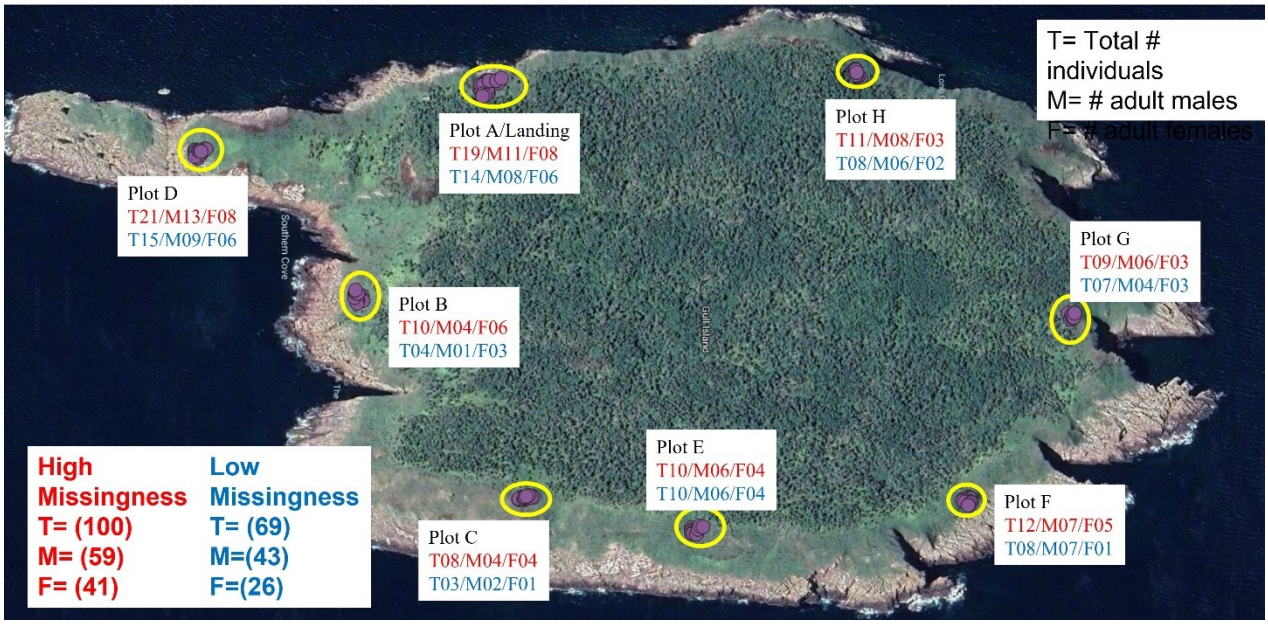

Figure 4. Distribution of individual samples that remained after filtering sequence data and removing individuals with either < 70 % missing data, high missingness group, or < 30% missing data, low missingness group. Samples are distributed by plot and sex for the high missingness group (in red) and low missingness group (in blue). T is the total per plot, M is the males, and F is the females.

To determine if any of the SNPs were sex-linked, an Admixture analysis with the number of clusters set to K = 2 was performed. Any clustering by sex would indicate sex-linked SNPs that would need to be removed to avoid bias (Figure 5). No clustering was present, which indicates that sex-linked markers do not contribute to the population structure within the dataset (Figure 5).

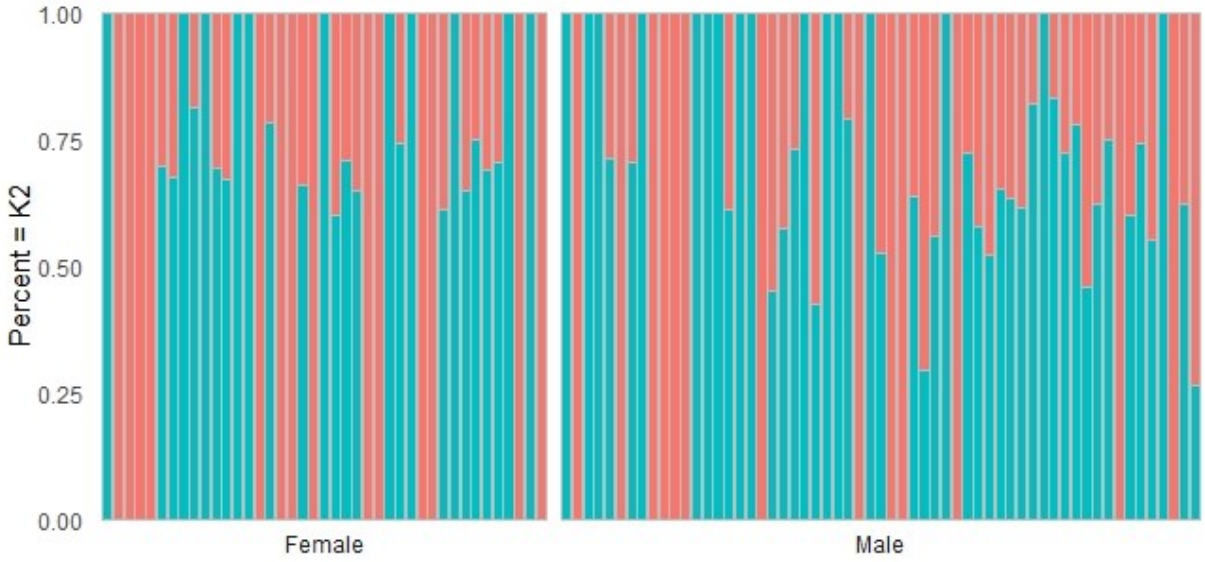

Figure 5. Admixture proportions on the high missingness group n= 100 for K=2 with cluster one in blue and cluster two in red. There is no clustering by sex, suggesting that no sex-linked markers were present in the dataset.

Analyses with Outgroup

To validate that my method could detect population structure at least at the regional level, I used data from an outgroup. I combined the 152 samples from Gull Island with 40 samples from Machias Seal Island provided by the University of New Brunswick before sequencing, and the 192 samples were assembled and filtered using the STACKS pipeline, with minimum allele coverage to keep a genotype equal to three and minimum percent of genotype data per population equal to 90. The final dataset contained 109 individuals, 71 from the study population and 38 from the outgroup, with up to 60 % missing data and 6393 shared SNPs.

Admixture

The admixture analysis showed that the study group and outgroup were two differentiated populations. This demonstrates that the ddRAD sequencing used was able to identify enough

SNPs to detect a difference in ancestry between two populations within the same geographic region (Figure 6).

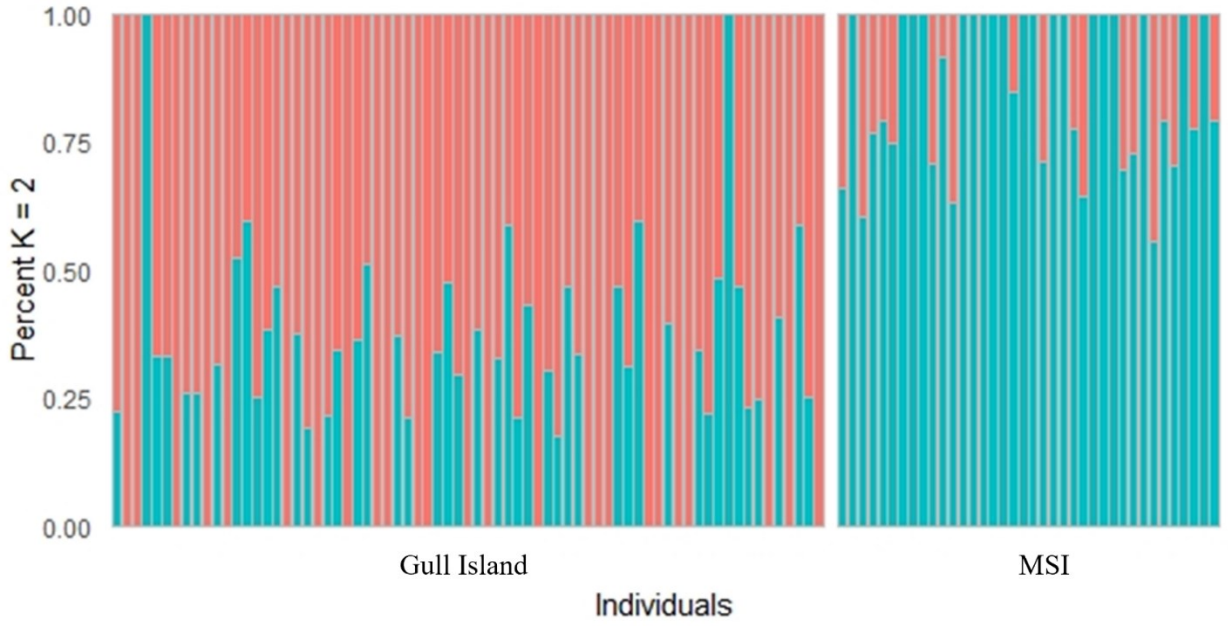

Figure 6. Admixture proportions for the outgroup (Machias Seal Island, MSI) compared to the study group (Gull Island) n= 109 for K=2 with cluster one in blue and cluster two in red. Clustering between the two colonies can be seen, indicating that the two colonies are genetically differentiated.

PCA

PCA showed two distinctive clusters separating the study group and outgroup primarily along the first PC axis (Figure 7). This finding agrees with the Admixture results that the study group and the outgroup comprised two differentiated populations of Atlantic puffin.

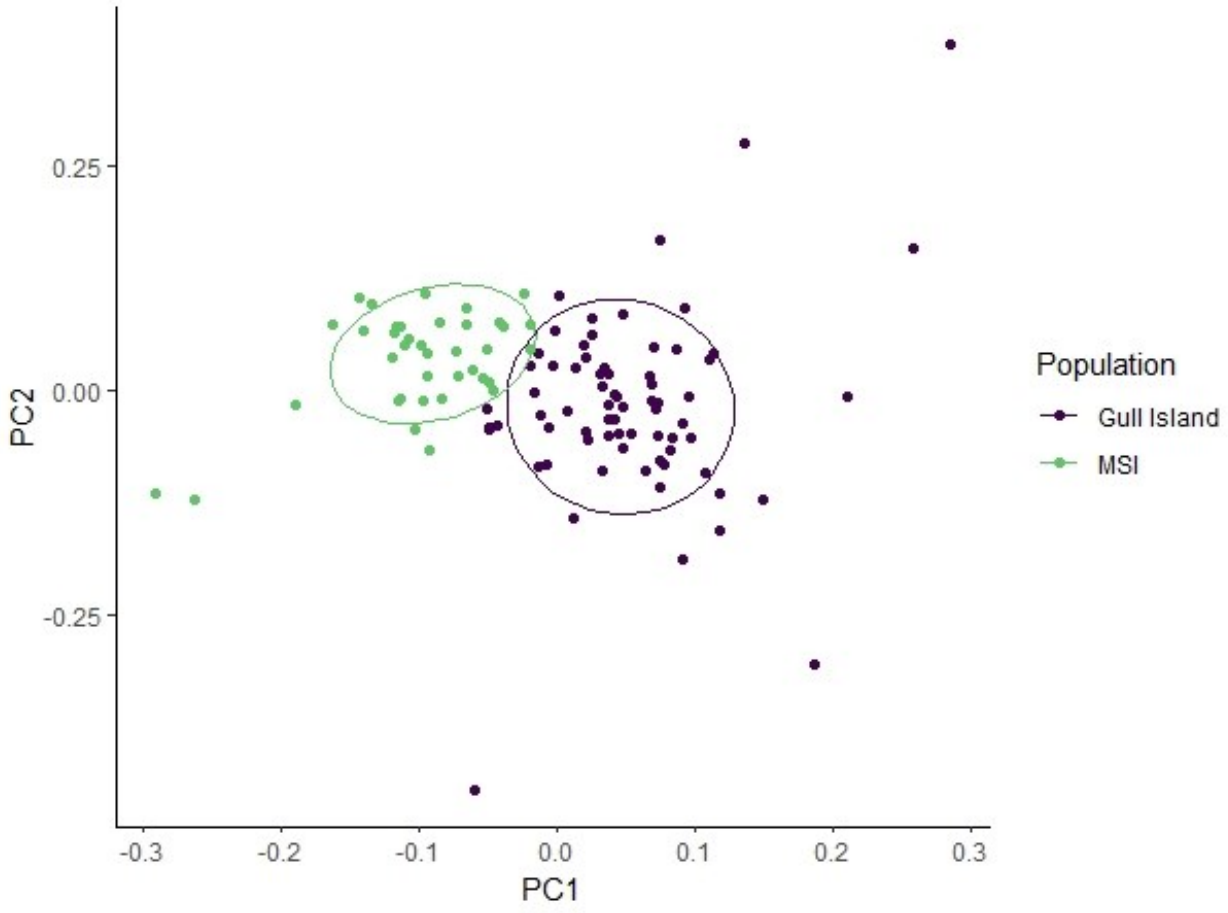

Figure 7. The Gull Island and Machias Seal Island (MSI) populations form distinct clusters based on PC1 and PC2 scores. The study group (Gull Island) is represented in purple, and the outgroup (Machias Seal Island) is represented in green. Both have an ellipsis representing 80% of the samples in each group.

Analysis Within Colony

To determine if there was genetic structure present in the Gull Island colony, I ran an Admixture analysis, a PCA, and a Mantel test based on Euclidean genetic distance. Data for each percent missingness group and sex were analyzed separately.

Admixture

The admixture analyses for the low and high missingness groups (from K = 1 to K = 7) showed that a single population best described the dataset. There were no discernible patterns in the visual analysis of admixture proportions for all individuals separated by plot from K = 2 to K = 5 for either the high or low missingness groups (Figures 8 and 9).

To include sex in the analysis, the admixture proportions were separated by plot and then by sex (Figure 10). No patterns appeared at K = 2, indicating that no sex differences could be detected. This analysis was performed on the high missingness dataset only since sample size in the low missingness dataset was too low when separated by plot and sex to be able to draw meaningful conclusions and previous analysis had shown that the results between the high and low missingness group were consistent.

PCA

Visualization of scatterplots representing the first and second principal component from a PCA on the complete SNP dataset did not show any clusters by plots or by sex (Figure 11). The PCAs showed no distinctive patterns or clustering; most samples were centered around zero.

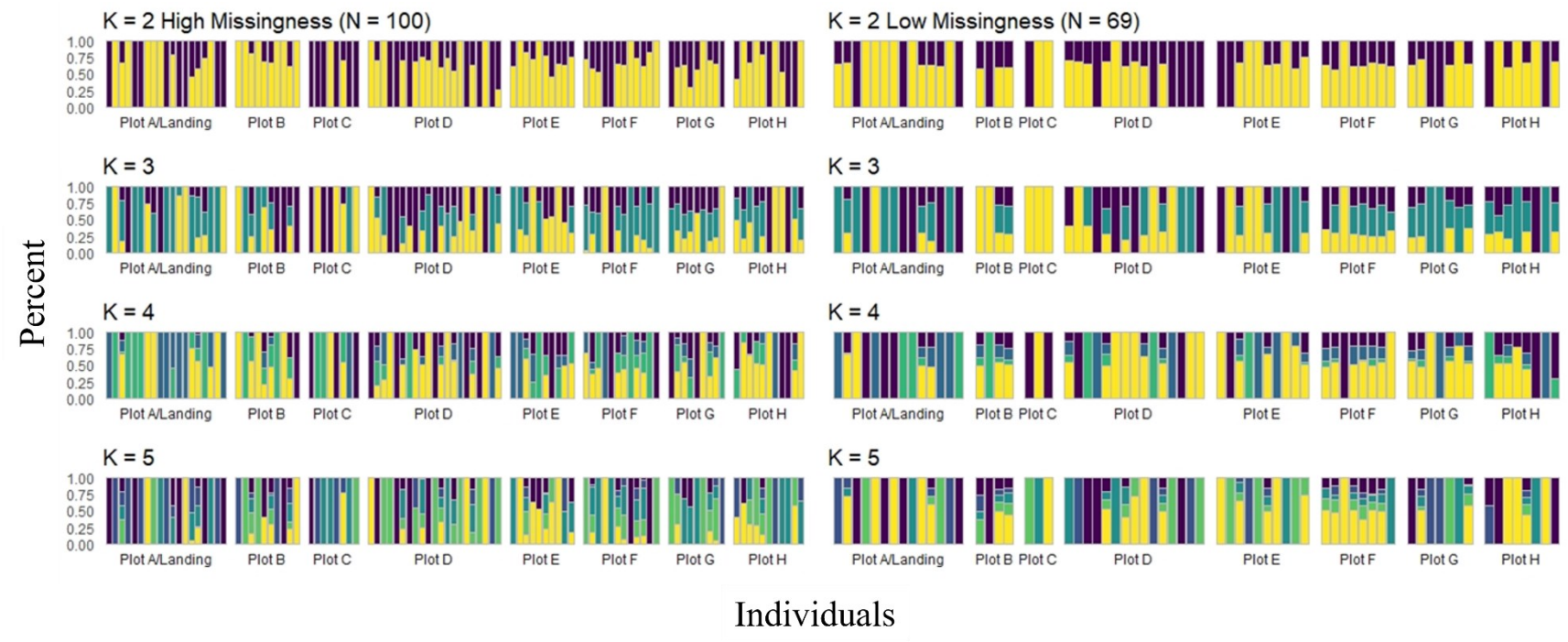

Figure 8. Admixture proportions for the study group K = 2 to K = 5 for the high (n = 100) and low (n = 69) missingness groups separated by plot. The different colours represent different clusters of K. The proportions demonstrate that clustering within plots is not present at high or low levels of K, meaning that neither simple nor complex structure can be detected. The high and low missingness groups are not visibly different from one another meaning that missingness does not interfere with the ability to detect structure.

Figure 9. Gull Island with admixture proportions for K = 2 for the high missingness group separated by plot. The individual sampling locations are shown as purple dots and the yellow circles represent the plot area, in the admixture blue represents cluster one and red represents cluster two. There are no visible differences between plots.

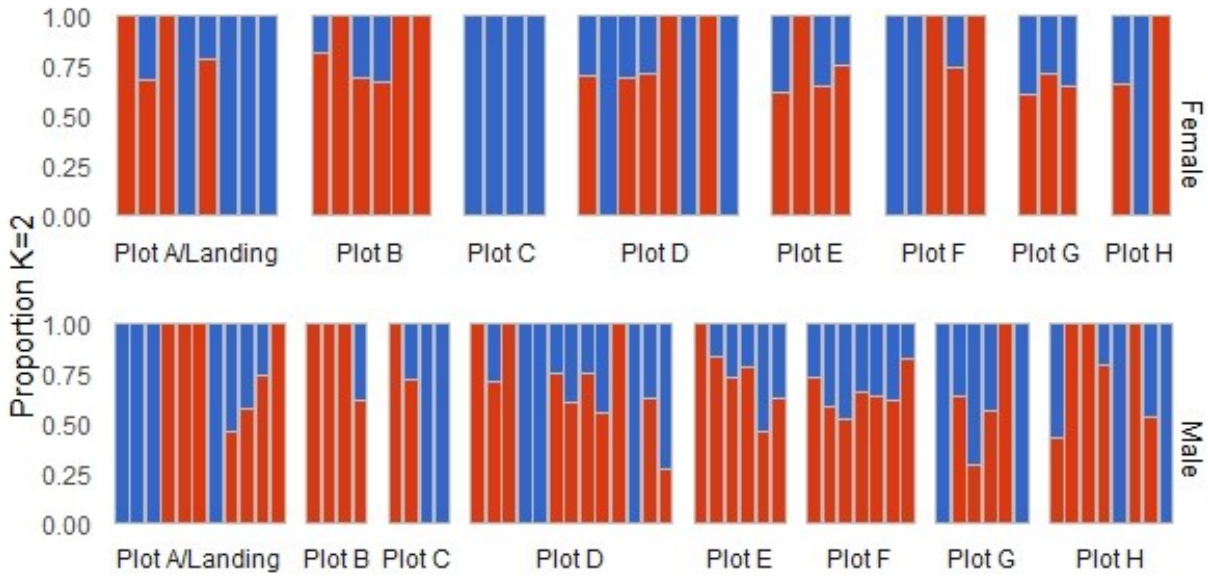

Figure 10. Admixture proportions for K = 2 in the high missingness group (n = 100) separated by plot and sex. The blue represents cluster one and the red represents cluster two. No difference in clustering between the sexes can be detected within plots.

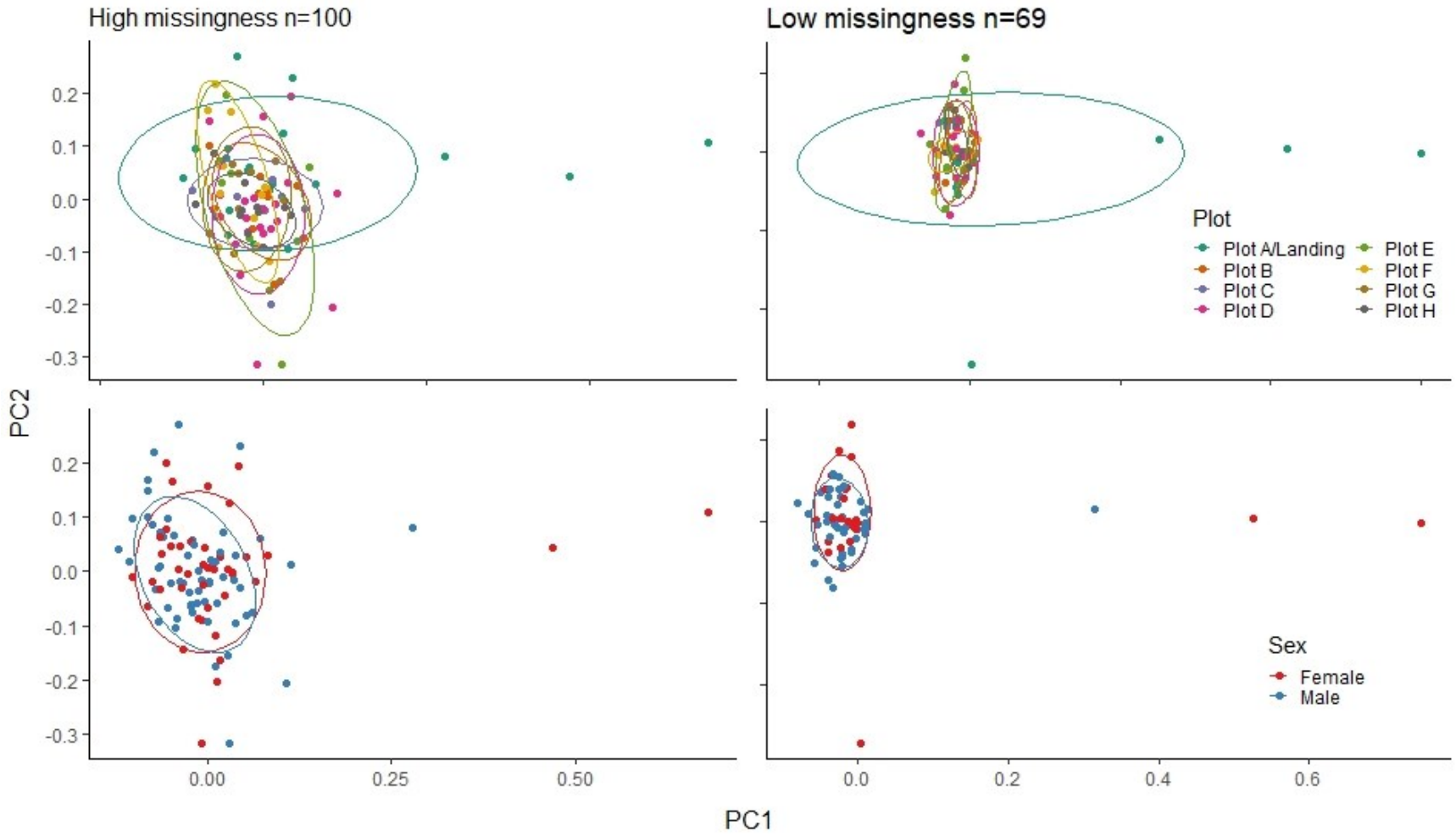

Figure 11. Individuals in different plots, as well as males and females, do not form distinct clusters based on PC1 and PC2 for the high (n = 100) on the left and low (n = 69) on the right missingness groups. The PCs on the top are broken down by plot and the PCAs on the bottom are broken down by sex. All groups have an ellipsis representing 80% of the samples. All four PCAs have the majority of samples clustered around zero, indicating that the genetic variance found in PC1 and PC2 is not explained by plot or by sex for either the high or low missingness groups

Mantel Test

A Mantel test of the relationship between genetic distance and geographic distance found no correlation when Euclidean genetic distance was compared to Euclidean geographic distances for the high missingness group (rm(100) = -0.07, p = 0.988) and the low missingness group (rm(69) = -0.072, p = 0.983). No significant correlation was found for individual sexes either (Table 1 and Figure 12).

-

		Mantel r	P-Value	95 % Cl
High Missingness	Total n=100	-0.070	0.988	-0.094 - -0.047
	Females n=41	-0.074	0.877	-0.118 - -0.023
	Males n=59	-0.068	0.973	-0.099 - -0.040
Low Missingness	Total n=69	-0.072	0.983	-0.106 - -0.041
	Females n=26	-0.191	0.975	-0.273 - -0.115
	Males n=43	-0.026	0.756	-0.061 - 0.011

Table 1. Mantel statistic, 95% confidence limits for the Mantel statistic, and one-tailed P-value, for both the total high (n = 100) and low (n = 69) missingness groups, as well as broken down by sex. No results were statistically significant.

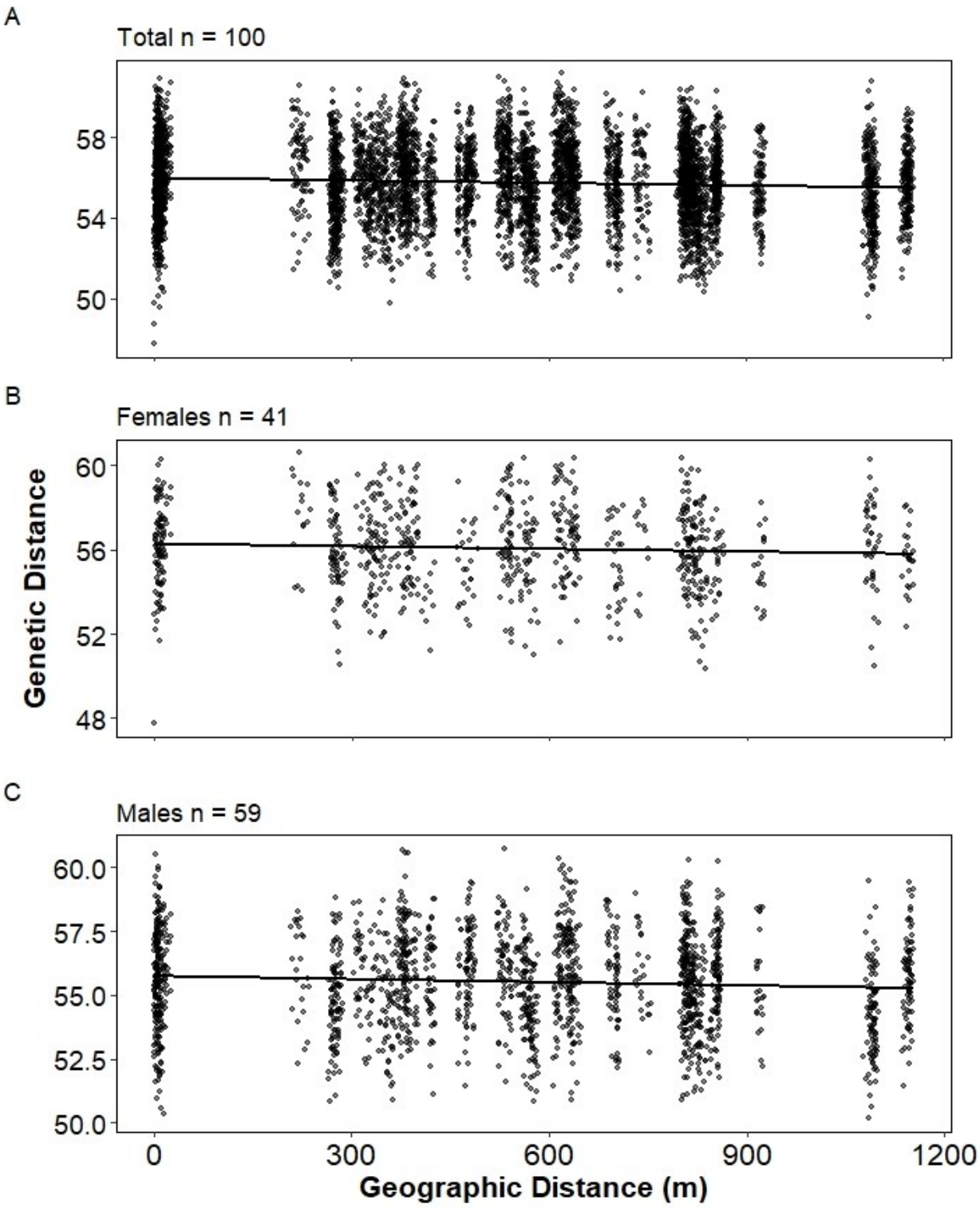

Figure 12. Relationship between geographic distance and genetic distance in Atlantic puffins. Points show pairwise comparisons of individuals, the first cluster represents all the within-plot pairwise comparisons and the other clusters represent between-plot comparisons. It is separated by A) Total n=100, B) Females n=41, and C) Males n=59.

The vertical breaks within the scatterplot that can be seen are caused by sampling being done within plots, rather than random sampling across the landscape. The first cluster represents all the within plot pairwise comparisons; the others are pairwise plot groups. There are 36 pairwise plot comparisons, but due to overlapping distances, 13 distance clusters can be seen (Figure 13).

Correlograms produced from the Mantel tests for the high and low missingness groups showed that correlation coefficients were generally negative with some variation around rm = 0. Results between the two groups and between the sexes were similar with no obvious trends (Figure 14).

Relatedness

The relatedness A_{jk} values for pairwise comparison of the imputed high missingness group ranged from 0.477 to -0.26. Coefficients skewed a bit lower than expected, most likely due to genotyping errors (e.g., overestimation of homozygosity, when half of a biallelic genotype is missing) and missing data, which can lead to an overestimation of homozygosity (Bresadola et al., 2020). Of the 4950 pairwise comparisons of different individuals, seven were in the range of first-degree relationships, parent-offspring, or full sibling, A_{jk} mean= 0.40, SD = 0.05, five of which were known adult/juvenile pairs sampled in 2021 to verify relatedness coefficients. One pair was a juvenile sampled in 2021 and the suspected parent (an adult male from the same burrow sampled in 2019) but not confirmed in 2021; one pair was two adult breeding males in different plots. Twenty-eight pairs were in the range of close relatives (first- and second-degree relatives, A_{jk} 0.28 to 0.13); 208 pairs were more distantly related (third- and fourth-degree relatives, A_{jk} 0.129 to 0.05); and 1188 were fifth degree to distant genetic relatives, A_{jk} 0.05-0).

A majority of the comparisons, 3519 pairs, were considered not genetically related (A_{jk} 0 to -0.26) (Figure 15).

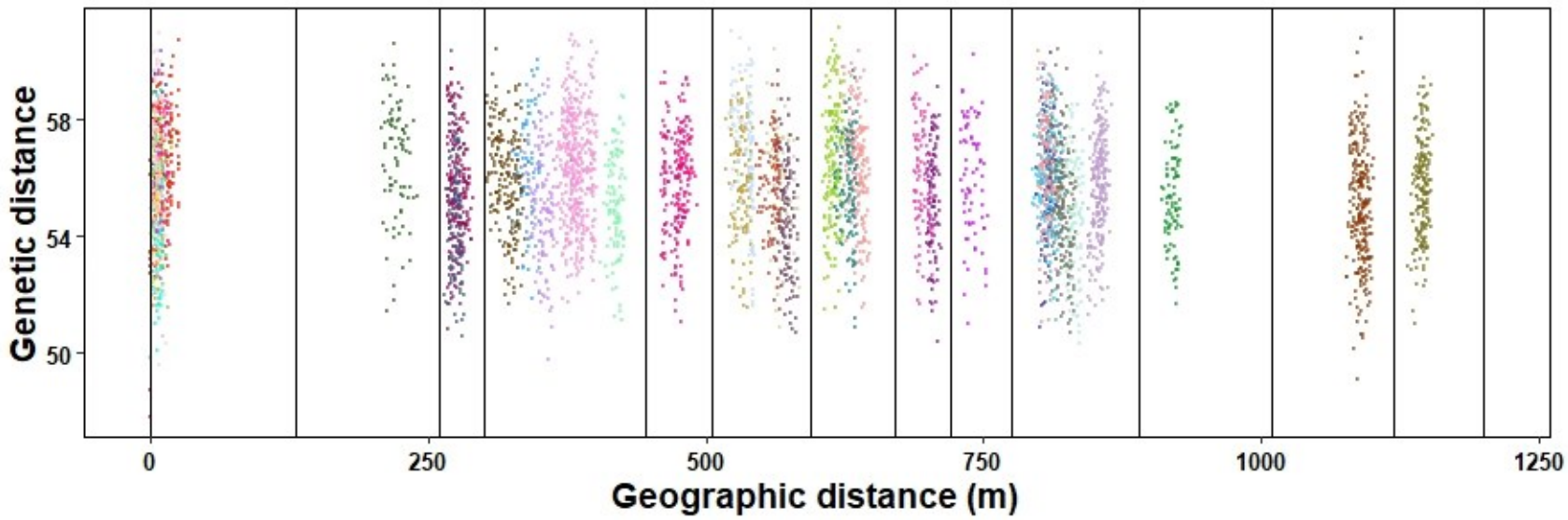

Figure 13. Relationship between geographic distance and genetic distance in Atlantic puffins. Colours show pairwise comparisons of individuals in different plots, the first cluster represents all the within-plot pairwise comparisons, and vertical lines show the midpoint breaks between distance groupings that were used for the correlogram.

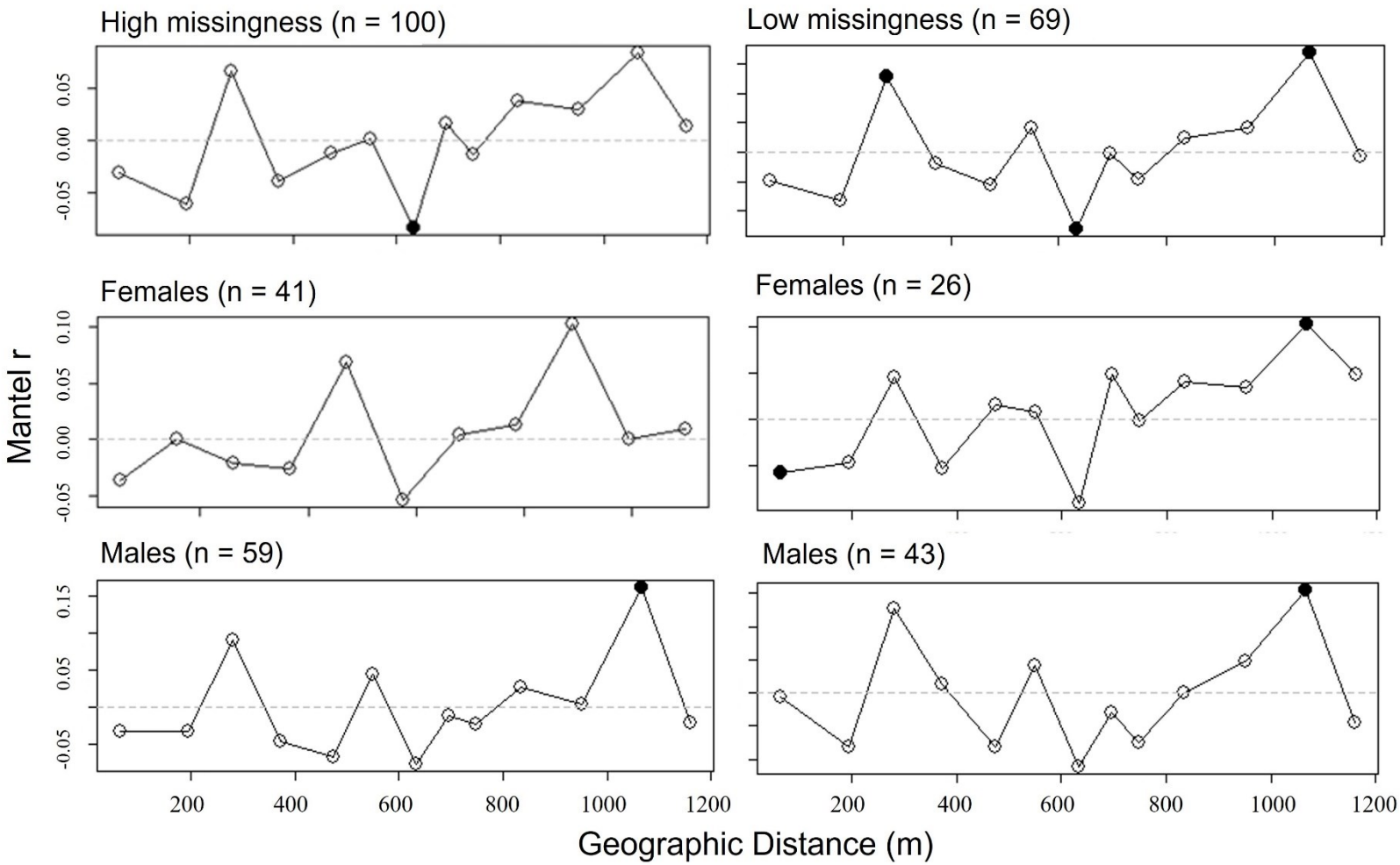

Figure 14. The line represents pairwise genetic distance comparisons at different geographic distances for the total high (n = 100), on the left, and low (n = 69), on the right, missingness groups as well as by sex. Breaks for the correlograms were set based on the midpoint between pairwise plot comparison clusters. The open circles represent non-significant breakpoints, and the filled circles represent significant breakpoints. The line shows that as distance increases, there is variation around 0 for the Mantel r which tends slightly negative, but no strong trends can be detected for either the high or low missingness groups or between the sexes.

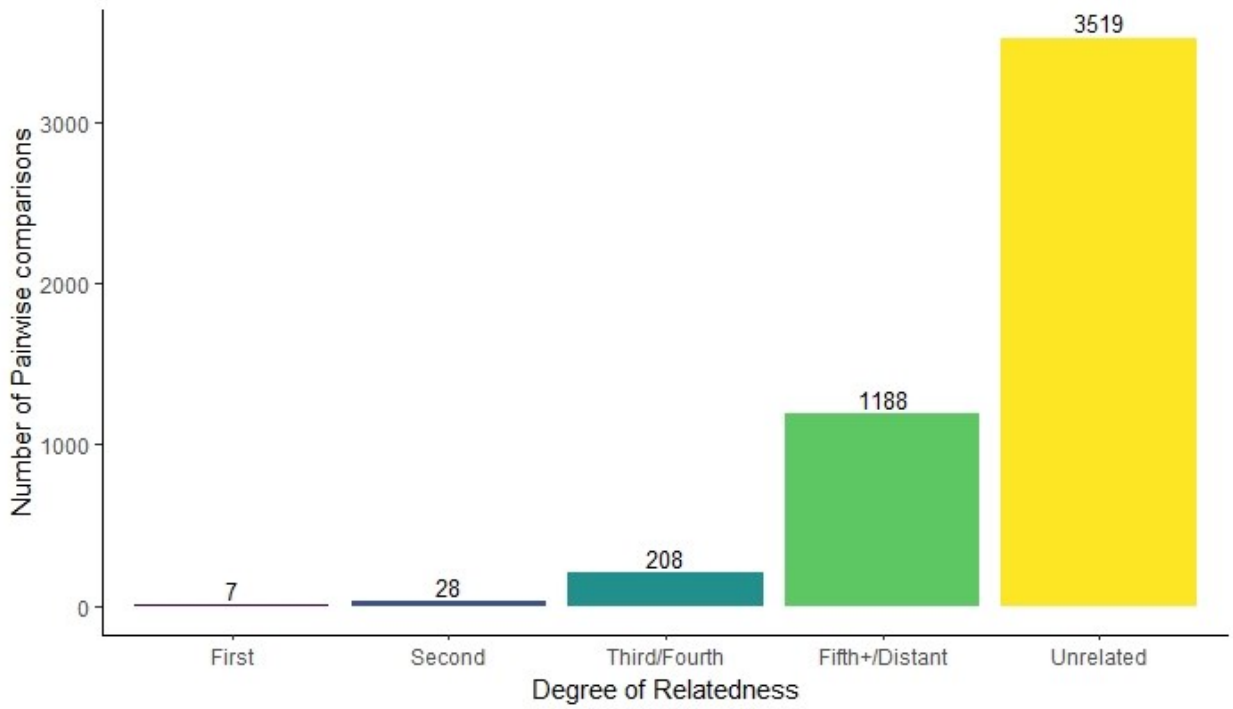

Figure 15. Degrees of relatedness found between pairwise comparisons of individuals in the high missingness group (n = 100) on Gull Island. The majority of comparisons were of unrelated individuals and only seven were first-degree relatives, of which five were known parent-offspring pairs.

A one-sided Mantel test performed on the A_{jk} coefficients and the geographic distance showed no significant p-values (rm (100) = -0.008, p = 0.717). Scatterplot and correlogram results confirm this finding with points centered around Mantel r = 0 (Figures 16 and 17). To determine whether there was sex-biased dispersal, Mantel tests for both females (rm (37) = -0.040, p = 0.852) and males (rm (57) = -0.019, p = 0.782) with juveniles removed were performed. There were no significant results, which indicates that sex-biased dispersal is unlikely in this species.

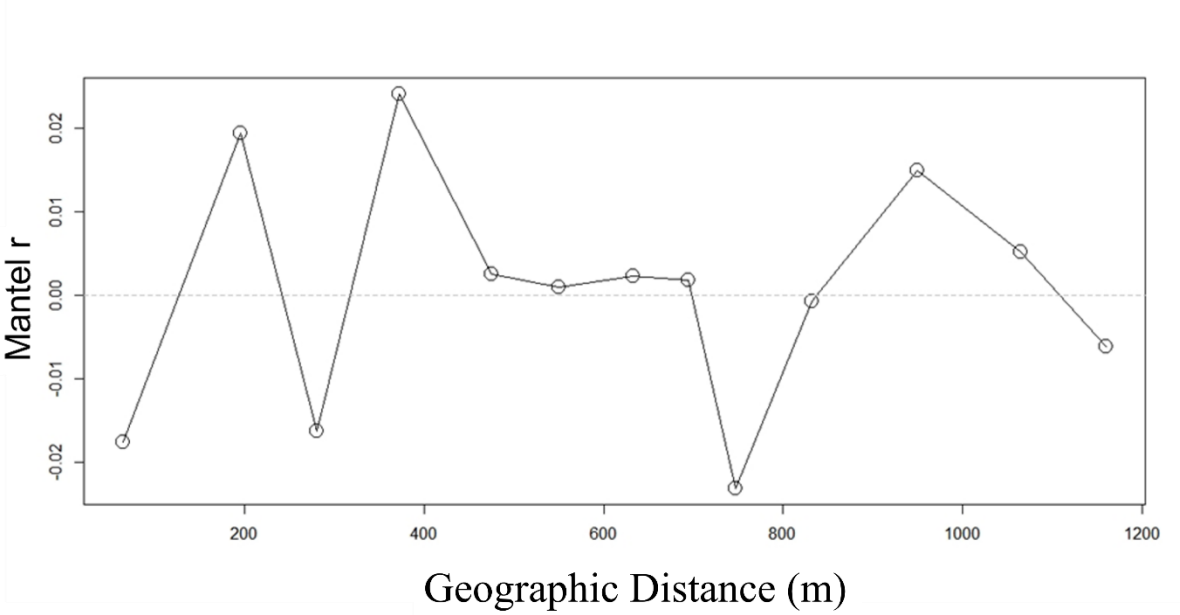

Figure 16. Pairwise relatedness coefficients (Ajk) for the high (n = 100) missingness group. Breaks for the correlograms were set based on the midpoint between pairwise plot comparison clusters. The open circles represent non-significant breakpoints. The line shows that as distance increases, there is variation around 0 for the Mantel r but no strong trends can be detected.

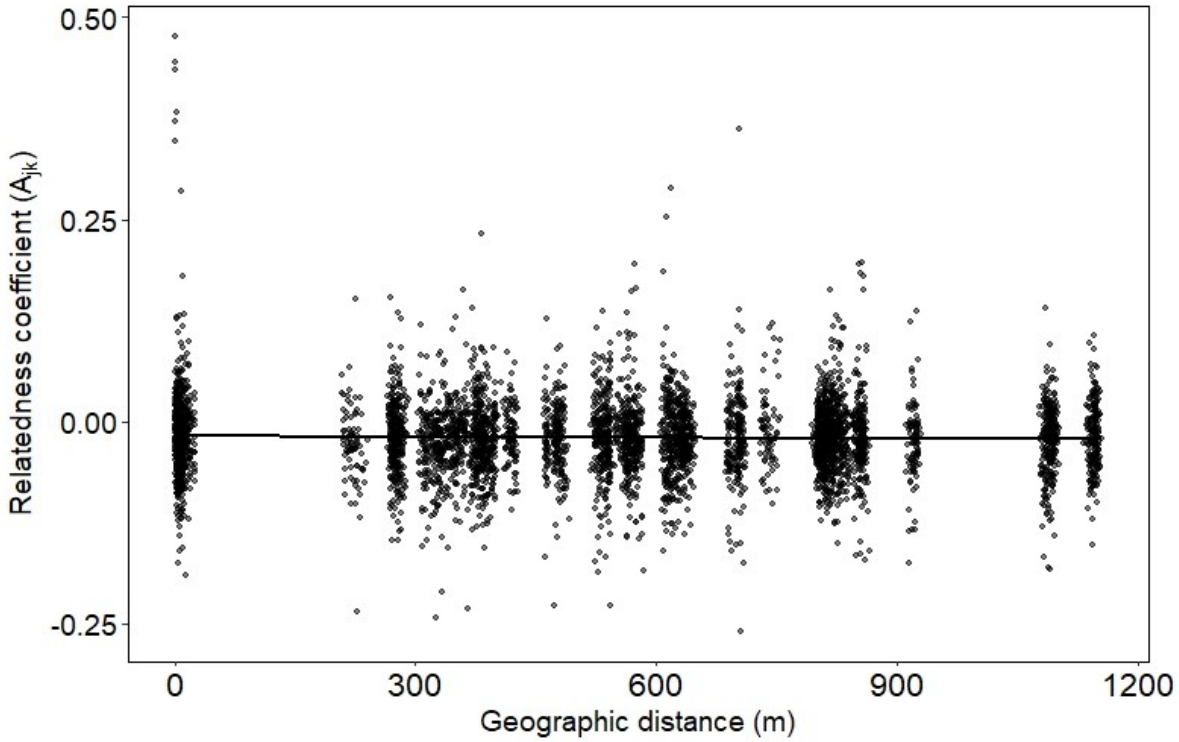

Figure 17. Pairwise relatedness coefficients (Ajk) compared to geographic distance in meters for the high missingness group (n = 100). The horizontal clustering seen represents pairs of plots. A trendline is included to highlight that there is no discernible trend within the dataset.

DISCUSSION

My main objective was to determine whether genetic population structure could be detected in a colony of Atlantic puffins. I used ddRAD sequencing on 146 breeding adults and six juveniles from Gull Island, Newfoundland and Labrador, Canada. I did not find strong evidence to support genetic structure on the colony.

Proportion Missingness does not Limit Conclusions of the Study

In this study, I explored whether having a group that retained more samples (n = 100) but had higher missingness (up to 70%) and lower numbers of SNPs (2714) was as sensitive at detecting differentiation as a group that had fewer samples (n = 69) but lower rates of missingness (up to 30%) and over six times the number of shared SNPs (16729). Within these two groups, samples were retained for all plots and both sexes within plots. Studies utilizing RAD and other types of sequencing that rely on shared SNPs can vary widely in the number of SNPs used, from tens to tens of thousands, and it can be difficult to determine how many SNPs are enough to address biological questions (Andrews et al., 2016; Puckett, 2017). Many studies utilize high quality DNA from whole liquid blood retain all of their samples, opting instead to reduce the number of SNPs used (Çilingir et al., 2022; Pedersen et al., 2017; Pertierra et al., 2020; Vendrami et al., 2017). Due to sample DNA being lower in quality which yielded a dataset with very high missingness in some samples, I chose to remove samples with a high proportion of missingness to retain a higher number of SNPs so that I would have a greater chance of being able to detect population structure at a fine scale given that all the individuals were from the same colony.

My results were consistent between the admixture, PCA analysis, and the Mantel test for both the high and low missingness groups. This finding is in accordance with other studies evaluating population structure, which were able to detect structure using a similar number of SNPs, such as Çilingir et al. (2022), which utilized 1632 SNPs, and Bouzid et al. (2022), which used 6944 SNPs. Based on the evidence, it can be inferred that more than a few thousand SNPs are not necessarily needed to detect population structure. Since ddRAD has been proven as an effective method for this type of analysis, the 2714 SNPs found in my high missingness group were sufficient to answer the questions about population structure and relatedness within my study population (Çilingir et al., 2022; Vendrami et al., 2017). While there was a lot of sample dropout due to DNA concentration being lower than ideal for RAD sequencing compared to another method such as microsatellite analysis, I am confident that the samples retained were more informative for answering questions of fine-scale population structure than the use of microsatellites would have been.

Outgroup Analysis Shows Ability to Detect Population Structure Using ddRAD Sequencing

I sequenced 40 individuals from Machias Seal Island to use as an outgroup for my study. The admixture and PCA results showed that the ddRAD sequencing could detect distinctive populations between the study group (Gull Island) and the outgroup (Machias Seal Island). This finding builds upon the work of Kersten et al. (2020) and Leigh et al. (2023), who detected broad regional differences in the genetic makeup of Atlantic puffins across their range using whole genome sequencing. Their results identify four large genetically distinct geographic regions that help highlight areas of potential gene flow as well as barriers to gene flow across the puffin range. In that study, puffins from colonies in the western Atlantic were grouped together, but

only two colonies from that region were sampled, Gull Island, NL and Gannet Island, NL. My results suggest that this region can be further divided into distinct genetic groups.

The genetic differentiation between the two colonies is interesting considering that over 700 puffin chicks from the Witless Bay Ecological Reserve in Newfoundland were translocated to Eastern Egg Rock in the Gulf of Maine between 1976 to 1981 to re-establish the colonies in this area, including Machias Seal Island (Kress & Nettleship, 1988). Unlike Eastern Egg Rock, Machias Seal Island was never fully extirpated of puffins, but some of the transplanted puffins from Witless Bay were recorded nesting there in subsequent years after their release. This could account for some of the crossover that was observed in the admixture results. As Machias Seal Island and the other colonies in the Gulf of Maine became repopulated, it is possible that they attracted individuals from other regions and those immigrants contribute to ongoing genetic differentiation in this population. More research in this area needs to be conducted to resolve differences at this scale. Differences in genetic structure among colonies in Newfoundland are unknown but would provide valuable insight into dispersal patterns in this species and to what degree natal philopatry exists.

Within-Colony Genetic Structure Opposite of Expected for Natal Philopatry

I expected to find genetic structure within the study population due to the high levels of breeding philopatry and some evidence of natal philopatry seen in this species (Breton et al., 2006; Harris, 2012). Strong natal philopatry has been shown to contribute to population structure in other seabird species such as Thick-billed murres (Ibarguchi et al., 2011). However, the admixture analysis here revealed that there were no distinctive populations within the study colony when separated by plot or sex. Thus, the null hypothesis could not be rejected and there was little detectable population structure in the colony. One consideration for this unexpected

result is habitat quality. Due to limitations of accessibility on steep, unstable ground and trying to limit researcher disturbance, some habitat was not possible to sample. Genetic structure may have been absent because although the plots were spread out, they were artificial clusters selected from a continuous ring of habitat, and it is important to consider natural variation in habitat when considering how population structure may form over time. Studies have found that natural processes including wind, sun exposure, slope aspect, and distance to shore can shape the landscape in ways where some habitats could become unsuitable due to erosion and changing ecology, preventing future generations from returning and triggering dispersal (Acker et al., 2017; Rodway, 1994). A study of King penguins (*Aptenodytes patagonicus*) found that variation in habitat quality led to more gene flow at lower quality sites because philopatry was lower and led to more genetic mixing within the population, whereas higher quality sites had higher levels of inbreeding (Garrett et al., 2020). In the present study, the quality of the habitat was not assessed, and is something that could be factored into the analysis of future studies.

The visualization of the first two components extracted from the PCA supported the lack of clear genetic structure detected by the admixture analysis and showed no clustering when separated by plot or sex. Other studies using RAD sequencing, using similar numbers of SNPs were able to detect distinctive clustering using PCA analysis in other species (Çilingir et al., 2022; Meisner & Albrechtsen, 2018; Vendrami et al., 2017). My PCA analysis using the outgroup was also able to detect differences between the two populations. This suggests that if there was any genetic structure within the colony, the PCA should be sensitive enough to detect it.

One study has found little to no genetic differentiation between colonies in another pelagic seabird, the Northern fulmar (*Fulmarus glacialis*), and suggests several reasons that this

may be the case including natal dispersal, long generation time, and historical processes that occurred during the last ice age that have prevented the population from reaching its migration-drift equilibrium (Colston-Nepali et al., 2020). Due to similar life history traits between these species, some or all of these processes could be affecting the population structure of puffins in Newfoundland as well. These factors result in a population that maintains high genetic diversity, a characteristic that has conservation and management implications. The most important of these implications is that genetic management should be considered on a larger regional scale rather than focusing on specific colonies and smaller scale management should be focused on specific local threats (Colston-Nepali et al., 2020).

The Mantel test showed no positive correlation between genetic distance and geographic distance. The lack of genetic structure within the colony suggests there may be more fine-scale natal dispersal in this species than previously thought (Breton et al., 2006). Tracking non-breeding juveniles post-fledging is notoriously difficult for a variety of reasons and return rates vary widely between studies and locations (Breton et al., 2006; Morrison et al., 2009). Some studies have attributed low return rates to the colony with mortality, but long-term studies have shown that immature survival is quite high (93 %) (Sandvik et al., 2008). Considering the high pre-breeding survival rates for this species, natal dispersal could be a driving factor in the lack of population structure seen at the colony level.

Isolation-by-distance was detected when a Mantel test was used with the relatedness coefficients. The calculation for the relatedness coefficients described in Heath et al. (2010) becomes more unstable as the values approach zero, and since a large proportion of the study population is unrelated, this creates a ceiling effect, and it may be a less reliable metric to measure isolation by distance by than the Euclidean genetic distance calculation. Isolation by

distance (Mantel test) in itself is not a very precise tool for detecting population structure, especially when few populations are being considered, and should be used as a baseline in conjunction with other measures to determine population structure (Jenkins et al., 2010). These limitations need to be considered when drawing inferences from the results and warrant exploration through further research into puffin dispersal.

I expected to find strong patterns of relatedness within the colony driven by natal philopatry. My results did not support this hypothesis and I found that only a small proportion (~5 %) of the individuals I sampled were moderately related and that the majority were only distantly related (24 %) or completely unrelated (71 %). The relatedness coefficient calculation was able to accurately identify all known relationship pairs, including confirming paternity in a suspected pair, based on previous sampling of the breeding pair and low divorce rate. This helps validate that even though there is up to 70 % missing data within the dataset, the test is a reliable metric for determining relatedness within the colony. The Mantel test for adult males and females showed no significant correlations between relatedness and geographic location on the colony, which supports a lack of sex-biased dispersal. These findings support the evidence for natal dispersal, which is surprising because many seabirds, especially puffins are well known for their strong breeding philopatry (Breton et al., 2006; Harris, 2012) and within-colony genetic structure has been demonstrated in other alcid species (Ibarguchi et al., 2011).

Due to the challenges previously outlined in collecting data on juvenile and pre-breeding puffins, not much is known about their dispersal (Morrison et al., 2009; Sandvik et al., 2008) and natal philopatry rates (Breton et al., 2006; Harris, 2012; Kersten et al., 2020). This study starts to shed light on questions about puffin dispersal and philopatry that have previously been unknown. This species may display natal philopatry on a regional rather than a fine scale, with dispersal

occurring within the islands of the Witless Bay Ecological Reserve or even further to the other colonies around Newfoundland. A recent study of Tufted puffins found that breeding colonies form their own panmictic genetic cluster, but that colonies could be differentiated from one another (Graham et al., 2023). Conducting similar regional-level studies of colonies around Newfoundland could resolve some of the uncertainty around the genetic connectivity of this species. One benefit of regional natal dispersal may be increased gene flow and a mechanism for inbreeding avoidance in this species. One study using pedigree data found that immigration limits inbreeding in another philopatric seabird, the Common tern (*Sterna hirundo*) (Ludwig & Becker, 2012).

Limitations of the Study and Improvements

As with any biological dataset, challenges can arise at any point during data collection, preparation, and analysis of the results. In this section, I will discuss some of the issues that arose during this project.

An issue that arose was the method of blood collection. For this study, blood was extracted onto filter paper, which resulted in lower quality and concentration of the extracted DNA. This resulted in an almost 2/3 reduction in sample size, as well as a loss in overall quality of the sequenced DNA which led to higher levels of genotyping error due to allele dropout.

Sequencing has come a long way in recent years, with advancements in data processing and technology that allow massive amounts of genetic material to be processed faster and cheaper than ever before, but with these improvements, challenges arise. There are many points at which errors can arise when selecting sequencing parameters such as what section of the genome to sequence, sequencing depth, and DNA concentration. Bioinformatic analysis and filtering after sequencing can also be prone to errors. Different pipelines, demultiplexing, and

where to set filter cut-offs can all have downstream implications on the quality of the results. Sequencing produces large amounts of data, thousands or millions of reads per sample which represent thousands of SNPs which can make traditional methods of validation impractical.

For future projects exploring population structure and natal philopatry in puffins, data collection could be improved by a more continuous sampling pattern within the colony and an increased overall sample size to account for downstream sample loss. In addition, the collection of whole liquid blood preserved with ethanol or using FTA™ cards that help preserve DNA, as well as storing the samples in an ultracold freezer to get a higher quality and concentration of extracted DNA to be sent for sequencing.

Another way to improve in the future would be to target specific individuals or groups. This could include breeding pair selection to look for evidence of inbreeding avoidance. More parent-offspring pairs could help determine more accurate inbreeding coefficients and paternity. Expanding the sample population to other nearby islands would be another way to investigate at what scale population structure can be detected in this species. As with all scientific studies, this study could benefit greatly from replication, which is often undervalued in science despite being one of the core principles of the scientific method.

ddRAD sequencing could be improved through the use of different restriction enzymes for selection of a section of the genome that would optimize the number of polymorphic SNPs. Biological duplicates added to the dataset would provide a better baseline of comparison to evaluate if errors occurred while sequencing, helping validate the data at the sequencing level. Another option is to sequence a few samples at a greater depth, which while more expensive, provides better coverage. Those individuals could then be compared to the shallower sequence depth samples to help correct genotyping errors within the data (Bresadola et al., 2020).

CONCLUSIONS

In addition to biological questions, consideration for future management should be taken into account and how factors that contribute to population structure and gene flow will affect the long-term survival of this species (Milot et al., 2008). As climate change affects the oceans and coastal environments, we see changes in barriers to dispersal (Friesen et al., 2007). Changes in water temperature, prey availability, fishing gear, and human population growth along the coast are all factors that contribute to marine species displacement and loss (Caplat et al., 2016). We may see changes in dispersal patterns of species affected by these factors and plasticity in how individuals respond, either with increased dispersal or philopatry.

My findings show that there is little evidence for population structure at the colony level in Atlantic puffins on Gull Island. They also suggest that natal philopatry may not be as strong, or seen at as fine a scale as breeding philopatry (Breton et al., 2006) and that there is more evidence for dispersal from the natal colony. There are still many questions surrounding this topic and more research needs to be done into what factors affect dispersal in this species. A multifaceted approach that combines longitudinal studies of chick movement and recruitment using tracking methods in conjunction with genetic testing will help us understand more fully the impacts that the pre-breeding time has on puffin population dynamics. This will ultimately help us to plan and adjust management strategies for this species in a rapidly changing environment.

REFERENCES

Acker, P., Besnard, A., Monnat, J.-Y., & Cam, E. (2017). Breeding habitat selection across spatial scales: Is the grass always greener on the other side? *Ecology (Durham)*, *98*(10), 2684–2697. https://doi.org/10.1002/ecy.1962

Alcaide, M., Scordato, E. S. C., Price, T. D., & Irwin, D. E. (2014). Genomic divergence in a ring species complex. *Nature (London)*, *511*(7507), 83–85. https://doi.org/10.1038/nature13285

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

Andersson, M., & Waldeck, P. (2007). Host-parasite kinship in a female-philopatric bird population: Evidence from relatedness trend analysis. *Molecular Ecology*, *16*(13), 2797–2806. https://doi.org/10.1111/j.1365-294X.2007.03301.x

Andreasen, A. M., Stewart, K. M., Longland, W. S., Beckmann, J. P., & Forister, M. L. (2012). Identification of source-sink dynamics in mountain lions of the Great Basin. *Molecular Ecology*, *21*(23), 5689–5701. https://doi.org/10.1111/j.1365-294X.2012.05740.x

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Anker-Nilssen, T., Kleven, O., Aarvak, T., & Lifjeld, J. T. (2008). No evidence of extra-pair paternity in the Atlantic Puffin *Fratercula arctica*. *Ibis*, *150*(3), 619–622. https://doi.org/10.1111/j.1474-919X.2008.00810.x

Armitage, K. B., Vuren, D. H. V., Ozgul, A., & Oli, M. K. (2011). Proximate causes of natal dispersal in female yellow-bellied marmots, *Marmota flaviventris*. *Ecology (Durham)*, *92*(1), 218–227. https://doi.org/10.1890/10-0109.1

Arthofer, W., Heussler, C., Krapf, P., Schlick-Steiner, B. C., & Steiner, F. M. (2018). Identifying the minimum number of microsatellite loci needed to assess population genetic structure: A case study in fly culturing. *Fly*, *12*(1), 13–22. https://doi.org/10.1080/19336934.2017.1396400

Ashbury, A. M., Willems, E. P., Utami Atmoko, S. S., Saputra, F., van Schaik, C. P., & van Noordwijk, M. A. (2020). Home range establishment and the mechanisms of philopatry among female Bornean orangutans (*Pongo pygmaeus wurmbii*) at Tuanan. *Behavioral Ecology and Sociobiology*, *74*(4). https://doi.org/10.1007/s00265-020-2818-1

Baguette, M., Blanchet, S., Legrand, D., Stevens, V. M., & Turlure, C. (2013). Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews of the Cambridge Philosophical Society*, *88*(2), 310–326. https://doi.org/10.1111/brv.12000

Baguette, M., & Van Dyck, H. (2007). Landscape connectivity and animal behavior: Functional grain as a key determinant for dispersal. *Landscape Ecology*, *22*(8), 1117–1129. https://doi.org/10.1007/s10980-007-9108-4

Baltazar-Soares, M., Klein, J. D., Correia, S. M., Reischig, T., Taxonera, A., Roque, S. M., Dos Passos, L., Durão, J., Lomba, J. P., Dinis, H., Cameron, S. J. K., Stiebens, V. A., & Eizaguirre, C. (2020). Distribution of genetic diversity reveals colonization patterns and

philopatry of the loggerhead sea turtles across geographic scales. *Scientific Reports*, *10*(1), 18001–18001. https://doi.org/10.1038/s41598-020-74141-6

Bell, J. R. (2005). The emergence of manipulative experiments in ecological spider research (1684-1973). *The Journal of Arachnology*, *33*(3), 826–849. https://doi.org/10.1043/0161-8202(2005)033[0826:TEOMEI]2.0.CO;2

BirdLife International. (2021). Species Factsheet: Fratercula Arctica. http://www.birdlife.org

Birt-Friesen, V. L. (1992). *Population differentiation and evolution among Thick-billed (Uria lomvia) and Common murres (U. aalge)*.

Blackmer, A. L., Ackerman, J. T., & Nevitt, G. A. (2004). Effects of investigator disturbance on hatching success and nest-site fidelity in a long-lived seabird, Leach's storm-petrel. *Biological Conservation*, *116*(1), 141–148. https://doi.org/10.1016/S0006-3207(03)00185-X

Bonte, D., Bossuyt, B., & Lens, L. (2007). Aerial dispersal plasticity under different wind velocities in a salt marsh wolf spider. *Behavioral Ecology*, *18*(2), 438–443. https://doi.org/10.1093/beheco/arl103

Bonte, D., Van Dyck, H., Bullock, J. M., Coulon, A., Delgado, M., Gibbs, M., Lehouck, V., Matthysen, E., Mustin, K., Saastamoinen, M., Schtickzelle, N., Stevens, V. M., Vandewoestijne, S., Baguette, M., Barton, K., Benton, T. G., Chaput-Bardy, A., Clobert, J., Dytham, C., … Travis, J. M. J. (2012). Costs of dispersal. *Biological Reviews of the Cambridge Philosophical Society*, *87*(2), 290–312. https://doi.org/10.1111/j.1469-185X.2011.00201.x

Bouzid, N. M., Archie, J. W., Anderson, R. A., Grummer, J. A., & Leaché, A. D. (2022). Evidence for ephemeral ring species formation during the diversification history of

western fence lizards (*Sceloporus occidentalis*). *Molecular Ecology*, *31*(2), 620–631. https://doi.org/10.1111/mec.15836

Bresadola, L., Link, V., Buerkle, C. A., Lexer, C., & Wegmann, D. (2020). Estimating and accounting for genotyping errors in RAD-seq experiments. *Molecular Ecology Resources*, *20*(4), 856–870. https://doi.org/10.1111/1755-0998.13153

Breton, A. R., Diamond, A. W., & Kress, S. W. (2006). Encounter, survival, and movement probabilities from an atlantic puffin (Fratercula arctica) metapopulation. *Ecological Monographs*, *76*(1), 133–149. https://doi.org/10.1890/05-0704

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *American Journal of Human Genetics*, *103*(3), 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Çakmak, E., Akın Pekşen, Ç., & Bilgin, C. C. (2017). Comparison of three different primer sets for sexing birds. *Journal of Veterinary Diagnostic Investigation*, *29*(1), 59–63. https://doi.org/10.1177/1040638716675197

Caplat, P., Pim Edelaar, Rachael Y Dudaniec, Andy J Green, Beth Okamura, Julien Cote, Johan Ekroos, Per R Jonsson, Jakob LÃ¶ndahl, Sylvie VM Tesson, & Eric J Petit. (2016). Looking beyond the mountain: Dispersal barriers in a changing world. *Frontiers in Ecology and the Environment*, *14*(5), 261–268. https://doi.org/10.1002/fee.1280

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics: Genotyping by sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, *4*(1), 7–7. https://doi.org/10.1186/s13742-015-0047-8

Çilingir, F. G., Hansen, D., Bunbury, N., Postma, E., Baxter, R., Turnbull, L., Ozgul, A., & Grossen, C. (2022). Low-coverage reduced representation sequencing reveals subtle within-island genetic structure in Aldabra giant tortoises. *Ecology and Evolution*, *12*(3), 1–13. https://doi.org/10.1002/ece3.8739

Clobert, J., Baguette, M., Benton, T. G., & Bullock, J. M. (2012). *Dispersal Ecology and Evolution: Vol. First edit*. OUP Oxford. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid&db=e000xna&AN=489629&site=ehost-live&scope=site

Clutton-Brock, T. H., & Lukas, D. (2012). The evolution of social philopatry and dispersal in female mammals. *Molecular Ecology*, *21*(3), 472–492. https://doi.org/10.1111/j.1365-294X.2011.05232.x

Colston-Nepali, L., Provencher, J. F., Mallory, M. L., Franckowiak, R. P., Sun, Z., Robertson, G. J., & Friesen, V. L. (2020). Using genomic tools to inform management of the Atlantic northern fulmar. *Conservation Genetics*, *21*(6), 1037–1050. https://doi.org/10.1007/s10592-020-01309-y

Colston-Nepali, L., Tigano, A., Boyle, B., & Friesen, V. (2019). Hybridization does not currently pose conservation concerns to murres in the Atlantic. *Conservation Genetics*, *20*(6), 1465–1470. https://doi.org/10.1007/s10592-019-01223-y

Coulson, J. C. (2016). A Review of Philopatry in Seabirds and Comparisons with Other Waterbird Species. *Waterbirds*, *39*(3), 229–240. https://doi.org/10.1675/063.039.0302

Cristofari, R., Trucchi, E., Whittington, J. D., Vigetta, S., Gachot-Neveu, H., Stenseth, N. C., Le Maho, Y., & Le Bohec, C. (2015). Spatial heterogeneity as a genetic mixing mechanism in highly philopatric colonial seabirds. *PLoS ONE*, *10*(2), 1–21. https://doi.org/10.1371/journal.pone.0117981

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dawson, D. A., Hunter, F. M., Pandhal, J., Buckland, R., Parham, A., Jones, I. L., Bradshaw, M., Jehle, R., & Burke, T. (2005). Assessment of 17 new whiskered auklet (*Aethia pygmaea*) microsatellite loci in 42 seabirds identifies 5-15 polymorphic markers for each of nine Alcinae species. *Molecular Ecology Notes*, *5*(2), 289–297. https://doi.org/10.1111/j.1471-8286.2005.00906.x

Ferrer Obiol, J., James, H. F., Chesser, R. T., Bretagnolle, V., González-Solís, J., Rozas, J., Riutort, M., & Welch, A. J. (2021). Integrating sequence capture and restriction site-associated DNA sequencing to resolve recent radiations of pelagic seabirds. *Systematic Biology*, *70*(5), 976–996. https://doi.org/10.1093/sysbio/syaa101

Filatova, O. A., Shpak, O. V., Ivkovich, T. V., Volkova, E. V., Fedutin, I. D., Ovsyanikova, E. N., Burdin, A. M., & Hoyt, E. (2019). Large-scale habitat segregation of fish-eating and mammal-eating killer whales (*Orcinus orca*) in the western North Pacific. *Polar Biology*, *42*(5), 931–941. https://doi.org/10.1007/s00300-019-02484-6

Fridolfsson, A.-K., & Ellegren, H. (1999). A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology*, *30*(1), 116–121. https://doi.org/10.2307/3677252

Friesen, V. L., Burg, T. M., & McCoy, K. D. (2007). Mechanisms of population differentiation in seabirds: Invited review. *Molecular Ecology*, *16*(9), 1765–1785. https://doi.org/10.1111/j.1365-294X.2006.03197.x

Garrett, L. J. H., Myatt, J. P., Sadler, J. P., Dawson, D. A., Hipperson, H., Colbourne, J. K., Dickey, R. C., Weber, S. B., & Reynolds, S. J. (2020). Spatio-temporal processes drive fine-scale genetic structure in an otherwise panmictic seabird population. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-77517-w

Genovart, M., Thibault, J. C., Igual, J. M., Bauzà-Ribot, M. del M., Rabouam, C., & Bretagnolle, V. (2013). Population structure and dispersal patterns within and between Atlantic and Mediterranean populations of a large-range pelagic seabird. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0070711

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*(7), 1–19. https://doi.org/10.18637/jss.v022.i07

Graham, B. A., Hipfner, J. M., Rojek, N. A., Stephensen, S. W., & Burg, T. M. (2023). Tufted Puffins exhibit low levels of genetic differentiation among breeding colonies in North America. *Ornithological Applications*. https://doi.org/10.1093/ornithapp/duad023

Groenewoud, F., Kingma, S. A., Hammers, M., Dugdale, H. L., Burke, T., Richardson, D. S., & Komdeur, J. (2018). Subordinate females in the cooperatively breeding Seychelles

warbler obtain direct benefits by joining unrelated groups. *Journal of Animal Ecology*, *87*(5), 1251–1263. https://doi.org/10.1111/1365-2656.12849

Harries, H. C., & Clement, C. R. (2014). Long-distance dispersal of the coconut palm by migration within the coral atoll ecosystem. *Annals of Botany*, *113*(4), 565–570. https://doi.org/10.1093/aob/mct293

Harris, M. P. (Michael P. (2012). *The puffin / Mike P. Harris and Sarah Wanless* (S. (Sarah) Wanless, Ed.; [Updated e). New Haven : Yale University Press.

Hatchwell, B. J. (2009). The evolution of cooperative breeding in birds: Kinship, dispersal and life history. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3217–3227. https://doi.org/10.1098/rstb.2009.0109

Heath, A. C., Benyamin, B., Nyholt, D. R., Goddard, M. E., Gordon, S., Martin, N. G., Henders, A. K., Madden, P. A., Visscher, P. M., McEvoy, B. P., Yang, J., & Montgomery, G. W. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608

Helfer, V., Broquet, T., & Fumagalli, L. (2012). Sex-specific estimates of dispersal show female philopatry and male dispersal in a promiscuous amphibian, the alpine salamander (*Salamandra atra*). *Molecular Ecology*, *21*(19), 4706–4720. https://doi.org/10.1111/j.1365-294X.2012.05742.x

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population stratification in genetic association studies. *Current Protocols in Human Genetics*, *95*, 1.22.1-1.22.23. https://doi.org/10.1002/cphg.48

Holderegger, R., & Gugerli, F. (2012). Where do you come from, where do you go? Directional migration rates in landscape genetics. *Molecular Ecology*, *21*(23), 5640–5642. https://doi.org/10.1111/mec.12032

Ibarguchi, G., Gaston, A. J., & Friesen, V. L. (2011). Philopatry, morphological divergence, and kin groups: Structuring in thick-billed murres Uria lomvia within a colony in Arctic Canada. *Journal of Avian Biology*, *42*(2), 134–150. https://doi.org/10.1111/j.1600-048X.2010.05023.x

Jenkins, D. G., Carey, M., Czerniewska, J., Fletcher, J., Hether, T., Jones, A., Knight, S., Knox, J., Long, T., Mannino, M., McGuire, M., Riffle, A., Segelsky, S., Shappell, L., Sterner, A., Strickler, T., & Tursi, R. (2010). A meta-analysis of isolation by distance: Relic or reference standard for landscape genetics? *Ecography (Copenhagen)*, *33*(2), 315–320. https://doi.org/10.1111/j.1600-0587.2010.06285.x

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Keefer, M. L., & Caudill, C. C. (2014). Homing and straying by anadromous salmonids: A review of mechanisms and rates. *Reviews in Fish Biology and Fisheries*, *24*(1), 333–368. https://doi.org/10.1007/s11160-013-9334-6

Kersten, O., Star, B., Leigh, D. M., Anker-Nilssen, T., Strøm, H., Danielsen, J., Descamps, S., Erikstad, K. E., Fitzsimmons, M. G., Fort, J., Hansen, E. S., Harris, M. P., Irestedt, M., Kleven, O., Mallory, M. L., Jakobsen, K. S., & Boessenkool, S. (2020). Complex

population structure of the Atlantic puffin revealed by whole genome analyses. *BioRxiv*, 2020.11.05.351874. https://doi.org/10.1101/2020.11.05.351874

Kingston, T., & Rossiter, S. J. (2004). Harmonic-hopping in Wallacea's bats. *Nature*, *429*(6992), 654–657. https://doi.org/10.1038/nature02487

Knaus, B. J., & Grünwald, N. J. (2017). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*(1), 44–53. https://doi.org/10.1111/1755-0998.12549

Komdeur, J. (1992). Erratum: Importance of habitat saturation and territory quality for evolution of cooperative breeding in the Seychelles warbler (Nature (1992) 358 (493-495)). *Nature*, *360*(6406), 768. https://doi.org/10.1038/360768c0

Komdeur, J. (1994a). *Experimental evidence for helping and hindering by previous offspring in the cooperative-breeding Seychelles warbler* acrocephalus sechellensis>. *34*(3), 175–186.

Komdeur, J. (1994b). The effect of kinship on helping in the cooperative breeding Seychelles warbler (*Acrocephalus sechellensis*). *Proceedings of the Royal Society B: Biological Sciences*, *256*(1345), 47–52. https://doi.org/10.1098/rspb.1994.0047

Kress, S. W., & Nettleship, D. N. (1988). Re-establishment of Atlantic puffins (*Fratercula arctica*) at a former breeding site in the Gulf of Maine. *Journal of Field Ornithology*, *59*(2), 161–170.

Kruuk, L. E. B., Sheldon, B. C., & Merilä, J. (2002). Severe inbreeding depression in collared flycatchers (Ficedula albicollis). *Proceedings of the Royal Society B: Biological Sciences*, *269*(1500), 1581–1589. https://doi.org/10.1098/rspb.2002.2049

Lee, R. E., Jr. (2012). Schowalter, Timothy D. Insect ecology: An ecosystem approach. *CHOICE: Current Reviews for Academic Libraries*, *49*(6), 1089.

Leigh, D. M., Kersten, O., Star, B., Anker-Nilssen, T., Burnham, K., Johnson, J., Provencher, J., & Boessenkool, S. (2023). Sympatry of genetically distinct Atlantic Puffins (Fratercula arctica) in the High Arctic. *Ibis (London, England)*, *165*(3), 1022–1030. https://doi.org/10.1111/ibi.13153

Ludwig, S. C., & Becker, P. H. (2012). Immigration prevents inbreeding in a growing colony of a long-lived and philopatric seabird. *Ibis*, *154*(1), 74–84. https://doi.org/10.1111/j.1474-919X.2011.01199.x

Mancilla-Morales, M. D., Velarde, E., Aguilar, A., Contreras-Rodríguez, A., Ezcurra, E., Rosas-Rodríguez, J. A., Soñanez-Organis, J. G., & Ruiz, E. A. (2022). Strong philopatry, isolation by distance, and local habitat have promoted genetic structure in Heermann's gull. *Diversity (Basel)*, *14*(2), 108. https://doi.org/10.3390/d14020108

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics (Austin)*, *210*(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Milot, E., Weimerskirch, H., & Bernatchez, L. (2008). The seabird paradox: Dispersal, genetic structure and population dynamics in a highly mobile, but philopatric albatross species. *Molecular Ecology*, *17*(7), 1658–1673. https://doi.org/10.1111/j.1365-294X.2008.03700.x

Mobley, K. B., Granroth-Wilding, H., Ellmen, M., Vähä, J.-P., Aykanat, T., Johnston, S. E., Orell, P., Erkinaro, J., & Primmer, C. R. (2019). Home ground advantage: Local Atlantic salmon have higher reproductive fitness than dispersers in the wild. *Science Advances*, *5*(2), eaav1112–eaav1112. https://doi.org/10.1126/sciadv.aav1112

Moen, S. (1991). Morphologic and genetic variation among breeding colonies of the Atlantic puffin (*Fratercula arctica*). *The Auk*, *108*(4), 755–763. https://doi.org/10.1093/auk/108.4.755

Morrison, K. W., Hipfner, J. M., Gjerdrum, C., & Green, D. J. (2009). Wing length and mass at fledging predict local juvenile survival and age at first return in Tufted puffins. *The Condor (Los Angeles, Calif.)*, *111*(3), 433–441. https://doi.org/10.1525/cond.2009.080099

Mourier, J., & Planes, S. (2013). Direct genetic evidence for reproductive philopatry and associated fine-scale migrations in female blacktip reef sharks (*Carcharhinus melanopterus*) in French Polynesia. *Molecular Ecology*, *22*(1), 201–214. https://doi.org/10.1111/mec.12103

Patrick, S. C., Weimerskirch, H., & Aubry, L. (2017). Reproductive success is driven by local site fidelity despite stronger specialisation by individuals for large-scale habitat preference. *The Journal of Animal Ecology*, *86*(3), 674–682. https://doi.org/10.1111/1365-2656.12636

Pedersen, S. H., Ferchaud, A.-L., Bertelsen, M. S., Bekkevold, D., & Hansen, M. M. (2017). Low genetic and phenotypic divergence in a contact zone between freshwater and marine sticklebacks: Gene flow constrains adaptation. *BMC Evolutionary Biology*, *17*(1), 130–130. https://doi.org/10.1186/s12862-017-0982-3

Perrin, N., & Lehmann, L. (2001). Is sociality driven by the costs of dispersal or the benefits of philopatry? A role for kin-discrimination mechanisms. *American Naturalist*, *158*(5), 471–483. https://doi.org/10.1086/323114

Pertierra, L. R., Segovia, N. I., Noll, D., Martinez, P. A., Pliscoff, P., Barbosa, A., Aragón, P., Raya Rey, A., Pistorius, P., Trathan, P., Polanowski, A., Bonadonna, F., Le Bohec, C., Bi, K., Wang-Claypool, C. Y., González-Acuña, D., Dantas, G. P. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2020). Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: Unforeseen vulnerabilities to global change. *Diversity and Distributions*, *26*(8), 958–975. https://doi.org/10.1111/ddi.13072

Poland, J. A., Brown, P. J., Sorrells, M. E., Jannink, J.-L., & Yin, T. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One*, *7*(2), e32253–e32253. https://doi.org/10.1371/journal.pone.0032253

Puckett, E. E. (2017). Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions. *Conservation Genetics Resources*, *9*(2), 289–304. Biological Science Collection; ProQuest Central; ProQuest One Academic. https://doi.org/10.1007/s12686-016-0643-7

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Queller, D. C., Strassmann, J. E., & Hughes, C. R. (1993). Microsatellites and kinship. *Trends in Ecology and Evolution*, *8*(8), 285–288. https://doi.org/10.1016/0169-5347(93)90256-O

Rodway, M. S. (1994). *Intra-colony variation in breeding success of Atlantic puffins: An application of habitat selection theory*.

Rodway, M. S., Montevecchi, W. A., & Chardine, J. W. (1996). Effects of investigator disturbance on breeding success of Atlantic puffins Fratercula arctica. *Biological Conservation*, *76*(3), 311–319. https://doi.org/10.1016/0006-3207(94)00118-9

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 231–253. https://doi.org/10.1146/annurev.ecolsys.38.091206.095611

Roques, S., Chancerel, E., Boury, C., Pierre, M., & Acolas, M. L. (2019). From microsatellites to single nucleotide polymorphisms for the genetic monitoring of a critically endangered sturgeon. *Ecology and Evolution*, *9*(12), 7017–7029. https://doi.org/10.1002/ece3.5268

Sandvik, H., Erikstad, K. E., Fauchald, P., & Tveraa, T. (2008). High survival of immatures in a long-lived seabird: Insights from a long-term study of the Atlantic puffin (*Fratercula arctica*). *The Auk*, *125*(3), 723–730. https://doi.org/10.1525/auk.2008.07059

Shields, W. M. (1982). *Philopatry, inbreeding, and the evolution of sex*. State University of New York Press.

Stacey, P. B., & Ligon, J. D. (1991). The Benefits-of-Philopatry Hypothesis for the Evolution of Cooperative Breeding: Variation in Territory Quality and Group Size Effects. *The American Naturalist*, *137*(6), 831–846. https://doi.org/10.1086/285196

Steiner, U. K., & Gaston, A. J. (2005). Reproductive consequences of natal dispersal in a highly philopatric seabird. *Behavioral Ecology*, *16*(3), 634–639. https://doi.org/10.1093/beheco/ari035

Taylor, S. A., & Friesen, V. L. (2012). Use of molecular genetics for understanding seabird evolution, ecology and conservation. *Marine Ecology Progress Series*, *451*, 285–304. https://doi.org/10.3354/meps09694

Vendrami, D. L. J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M. S., Leese, F., McMinn, C., Moore, H., & Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: Implications for understanding phenotypic plasticity. *Royal Society Open Science*, *4*(2). https://doi.org/10.1098/rsos.160548

Wang, J. (2017). Estimating pairwise relatedness in a small sample of individuals. *Heredity*, *119*(5), 302–313. https://doi.org/10.1038/hdy.2017.52

Wickland, D. P., Battu, G., Hudson, K. A., Diers, B. W., & Hudson, M. E. (2017). A comparison of genotyping-by-sequencing analysis methods on low-coverage crop datasets shows advantages of a new workflow, GB-eaSy. *BMC Bioinformatics*, *18*(1), 586. https://doi.org/10.1186/s12859-017-2000-6

Wilhelm, S. I., Schau, J. J., Schau, E., Dooley, S. M., Wiseman, D. L., & Hogan, H. A. (2013). Atlantic puffins are attracted to coastal communities in eastern newfoundland. *Northeastern Naturalist*, *20*(4), 624–630. https://doi.org/10.1656/045.020.0409

Winters, S., Allen, W. L., & Higham, J. P. (2019). The structure of species discrimination signals across a primate radiation. *BioRxiv*, 1–28. https://doi.org/10.1101/574558

Witless Bay Ecological Reserve management plan. (1998). St. John's, Nfld. : Parks and Natural Areas Division, Dept. of Tourism, Culture and Recreation, Government of Newfoundland and Labrador.