Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

Hamish Andrew Burnett

Genomic consequences of anthropogenic reintroduction in high-arctic Svalbard reindeer

Master's thesis in Biology

Supervisor: Henrik Jensen, Brage Bremset Hansen, Michael D. Martin, Vanessa C. Bieker

June 2020

NTNU
Norwegian University of Science and Technology

Abstract

Anthropogenic reintroductions are a commonly used conservation strategy following local extirpations. However, establishing new populations with a small number of founders can result in population bottlenecks causing reduced genetic diversity, increased inbreeding, and altered genetic structuring of sub-populations. Following large-scale extirpations due to severe overharvesting in the 19th - 20th Century, many current populations of the endemic high-arctic Svalbard reindeer subspecies, *Rangifer tarandus platyrhynchus*, originate from recent reintroduction programs or natural recolonisations. This study uses whole-genome shotgun sequencing of 100 reindeer to investigate the population genomic consequences of both anthropogenic reintroductions and natural recolonisation. Genetic structure analyses indicate significant genetic drift resulting from reintroductions occurred, with reintroduced populations forming two distinct genetic clusters corresponding to the two reintroductions. There was little evidence for gene flow between reintroduced and natural populations, with very high genetic divergence between a reintroduced and a naturally recolonised population only separated by ~15km. However, reintroduced populations showed no significant decrease in heterozygosity or increase in inbreeding compared to the source population, and no signature of the reintroduction event could be observed in Runs of Homozygosity (RoH) length distributions. In contrast, some naturally recolonised populations showed high inbreeding (F= >0.3), longer RoH segments, and reduced heterozygosity. These results indicate that some naturally recolonised sub-populations were likely founded by a small number of individuals, and that anthropogenic reintroductions can be more effective for establishing populations with high genetic diversity than some cases of natural recolonisation. Populations naturally recolonised from remnant populations may be particularly vulnerable to the accumulation of inbreeding and drift-load, and therefore warrant particular attention in future research. These findings have relevance for the conservation and management of other species with similar life histories where natural recolonisation from fragmented populations is expected to play a key role in population recovery, or where reintroductions are being considered.

Acknowledgements

There are numerous people who have made this project possible (and very much enjoyable!). First and foremost, I must thank my supervisors; Henrik Jensen, Brage Hansen, Mike Martin, and Vanessa Bieker, who have provided invaluable assistance and encouragement through every stage of the project. I am also particularly grateful to Bart Peeters and Mathilde Le Moullec for invaluable advice and discussions regarding the project. Discussions with and advice from members of the NTNU Museum Holomuseomics group were also helpful for the bioinformatics and data analysis aspects of this project.

This project could not have happened without extensive previous field and lab work. I thank Mathilde Le Moullec, the Norwegian Polar Institute (particularly Åshild Ø. Pedersen) and numerous other field assistants for collecting the DNA samples used in this project. I am also grateful to Karoline Bælum and the Farmhamna trapping station for providing tissue samples to NP and offering my use of them. Bart Peeters and Knut H for labwork involved in extracting DNA and generously allowing my use of these samples, which formed the basis of this whole project.

Finally, the support and company of my friends and family for making my time working on this project thoroughly tolerable!

This study was financed by the Research Council Norway (Arctic Field Grant 295908), FRIMEDBIO 276080, Centres of Excellence SFF-III 223257, and the Svalbard Environmental Protection Fund (14/137 and 15/105).

Contents

Introduction

Species reintroductions involve the artificial establishment of new populations to areas in which they once occurred but have since become locally extirpated, and are commonly used in ecological restoration and biodiversity conservation programmes (Armstrong and Seddon 2008; Weeks et al. 2011). Most reintroductions involve translocation of a limited (often very small) number of individuals to establish new populations that are, to some degree, geographically isolated from the species' current range (Frankham 2010). Together, these characteristics may impart genetic consequences on the newly founded populations. Founding populations are often characterised by a small effective population size, contain only a subset of the genetic variation that exists in their source populations, and have limited or no migration between other populations (Lynch and Gabriel 1990; Frankham 2010). This can impact the short- and long-term viability of populations (Frankham, 2005; Weeks *et al.*, 2011).

Founder effects are effectively a bottleneck of the source population and have the short-term effect of an immediate reduction in genetic diversity, because the founder population typically only represents a sample of the genetic variation found in the source population (Nei et al. 1975; Willis and Willis 2010). As a consequence, reintroduced populations commonly have lower heterozygosity or allelic richness than their source populations (e.g. Williams et al. 2002; Jamieson 2011; Grossen et al. 2018). Founding events also influence genetic diversity in the medium- and long-term because bottlenecked populations have reduced population sizes, which lead to the erosion of genetic diversity as alleles become lost due to genetic drift over time (Nei et al. 1975; Allendorf 1986). The magnitude of loss of genetic diversity is highly dependent on the effective population size, and the length (in generations) of the bottleneck. Effective population sizes depends not only on the number of individuals but also factors such as mating system (variance in reproductive success), sex ratios, and generational overlap, as well as the demographic history (such as past bottlenecks) because these influence the probability that all alleles in the population will be passed on at equal rates (Frankham 1995). A key factor determining the long-term genetic effects of a bottleneck is the growth rate of the population and overlap of generations, which influence the effective length of the population bottleneck. A high degree of generational overlap and fast population growth reduce the length of the bottleneck and reduce the probability that alleles will be lost or change in frequency due to genetic drift (Nei et al. 1975; Allendorf 1986). Conversely, in populations which remain small for long periods of time, genetic drift will reduce genetic variation by leading to loss of alleles by random chance, and as a result, fixation of deleterious recessive alleles that reduce

population fitness, known as drift-load (Whitlock et al. 2000; Willi et al. 2013). In small populations the rate of inbreeding also increases because there is a higher probability that two individuals share recent common ancestors, which causes a decrease in genetic diversity at the individual level by increasing homozygosity. This leads to an increase of homozygous genotypes of segregating deleterious recessive alleles which can reduce individual fitness, known as inbreeding depression (Charlesworth and Willis 2009).

Gene flow can also act to maintain or counteract the loss of genetic diversity in small populations by introducing novel genetic material and replenishing genetic variation lost to drift and bottleneck effects (Vucetich and Waite 2000; Latch and Rhodes 2005). Furthermore, it can reduce rates of inbreeding because new migrants from genetically differentiated populations reduce the average relatedness between individuals in a population, hence population connectivity may be the most efficient way in which to maintain or restore genetic variation in small populations (Frankham et al. 2017). Inbreeding and drift-load accumulate over generations at a rate that depends on the population size and the rate of immigration (Whitlock et al. 2000; Willi et al. 2013). Therefore population growth rate and structure are key factors in determining the genetic diversity of a reintroduced population, the effect of mutational load on fitness, and thus population viability (Latch and Rhodes 2005; Biebach and Keller 2010, 2012).

Genetic diversity is important for population viability in the medium to long-term because it influences the evolutionary potential of a population and its capacity to adapt to changing environmental conditions (Frankham 2005a), Currently, climate change and the fragmentation and degradation of habitats occur at unprecedented rates across the globe (Bellard et al. 2012; Hooper et al. 2012). Such changes in the environment can lead to maladaptation and reduced fitness of populations (Hendry 2017), and evolutionary responses may be required for a population to persist (Bell and Gonzalez 2009). Small, isolated populations with low levels of genetic diversity and elevated mutational load can be particularly vulnerable to environmental change (Frankham 2005b). The fitness consequences of inbreeding and drift load may also be higher in stressful environments, such as when a population is maladapted to its environment (Hedrick 2000; Frankham 2005a). Standing genetic variation (i.e. the variation that currently exists in the population) is particularly important for enabling a population to respond to rapid environmental changes; populations with reduced genetic variation that natural selection can act upon have a reduced evolutionary potential and thus may be slower or less likely to adapt to the changes in their environment (Orr and Unckless 2008). Furthermore, if population sizes

are small, there is a greater probability of extinction before such adaptation can occur (Bell 2013; Carlson et al. 2014).

A key goal in the management of reintroduced or fragmented populations is thus to maximise the genetic diversity and minimise inbreeding to increase their long-term viability (Frankham et al. 2017). To this end, genetic studies of reintroduced populations have become a focus of conservation biology in the last decade to better understand the genetic consequences of reintroductions (Frankham 2010). Several studies have shown that genetic diversity in reintroduced populations is often higher in those which receive gene flow from other populations (Latch and Rhodes 2005; Biebach and Keller 2012), which use multiple source populations to buffer against founder effects (Williams et al. 2000; Huff et al. 2010; Williams and Scribner 2010; Sasmal et al. 2013), and those which use augmentation, or multiple waves of reintroduction (Drauch and Rhodes 2007; Cullingham and Moehrenschlager 2013). Other studies have shown that inbreeding depression can occur in reintroduced populations (Marshall and Spalton 2000; Jamieson et al. 2007). Erosion of genetic diversity in reintroduced populations most often occurs in those which are isolated and do not experience rapid population growth (Williams et al. 2002; Hundertmark and van Daele 2010). This can potentially have detrimental consequences for fitness related traits (Wisely et al. 2008).

Based on this body of work, a number of management practice recommendations have been put forward, such as "rules of thumb" that give general guidelines for the number of individuals that need to be reintroduced and the amount of gene flow required to maintain genetic diversity. For example, 20 effective founders (Willis and Willis 2010) and one effective migrant per generation (Vucetich and Waite 2000). These recommendations come with the caveat that they are highly dependent on the ecology and life history of the species in question (Willis and Willis 2010).

Typically, microsatellite markers have been used to quantify the genetic diversity and structure of reintroduced populations (Frankham 2010). Such genetic analyses are fundamental for understanding the genetic outcome of reintroductions (Hicks et al. 2007; Taylor and Jamieson 2008; Wright et al. 2014). In recent years genomic data has shown increased power and utility for providing insights into both the "traditional" population structure and genetic diversity analyses, but also into the demographic processes related to reintroductions, such as quantifying inbreeding (Grossen et al. 2018). One such advantage of genomic data is its utility for detecting runs of homozygosity. Inbreeding between individuals that share common ancestors results in offspring with stretches of contiguous homozygous loci along segments of their chromosomes that both parents inherited from a common ancestor, known as Identical by

descent (IBD) tracts or Runs of Homozygosity (RoH) (Kardos et al. 2015). These can be quantified to provide an estimate of individual inbreeding coefficients (i.e. the proportion of the genome that is in RoH) similar to other molecular methods of estimating inbreeding (Kardos et al. 2015). However, analysis of RoH can also provide insights into the demographic history of a population by providing indications of how many generations prior the shared ancestors of the parents were, based on the length distribution of RoH (Kardos et al. 2016; Druet and Gautier 2017; Brüniche-Olsen et al. 2018). RoH-based inbreeding estimates used in combination with population structure analyses (to infer/confirm admixture and population origins) can provide important insights into the contribution of historical demography to current levels of genetic diversity and inbreeding, for example past population bottlenecks associated with founder events (e.g. Grossen et al. 2018).

The Svalbard reindeer (*Rangifer tarandus platyrhynchus* Vrolik, 1829) subspecies and its strong metapopulation structure provides an ideal system in which these novel methods can now be applied to address the genetic consequences of local extirpations and subsequent anthropogenic reintroductions. Furthermore, this particular study system allows comparisons to be made between the genetic consequences of reintroductions, and those related to natural recolonisation processes. On a circumarctic scale, reindeer and caribou (*Rangifer tarandus)* populations show diverse population trends in response to habitat disturbances and climate change (Uboni et al. 2016), including widespread population declines (Vors and Boyce 2009). The number of Svalbard reindeer declined drastically due to overharvesting until 1925, and the subspecies (endemic to the Svalbard archipelago) was extirpated from much of its former range (Le Moullec et al. 2019). The subspecies has since largely recovered, with natural recolonisation and anthropogenic reintroductions restoring most of its former range. Accordingly, Svalbard reindeer are now very abundant (Le Moullec et al. 2019) with many populations relatively stable or increasing in size (Hansen et al. 2019c). However, environmental conditions are rapidly changing in the arctic due to climate change (Post et al. 2009), and this may possibly be contributing to reindeer and caribou population declines at a pan-arctic scale (Vors and Boyce 2009). On Svalbard, there has been a "regime-shift" in winter weather conditions, with an increased frequency of rain-on-snow (ROS) events (Peeters et al. 2019). These ROS events result in ground ice formation that limits reindeer's access to forage, and this interacts with population density to cause food limitation in poor winters, thus occasionally destabilizing population dynamics (Kohler and Aanes 2004; but see Hansen et al. 2019a). Reindeer partly buffer against the effects of this with behavioural plasticity in foraging strategies at small spatial scales (Hansen et al. 2010b, 2019b; Loe et al. 2016). However, a

further complicating factor is that in coastal areas on Svalbard, reindeer are distributed in small patches of suitable habitat on isolated peninsulas and valleys separated by glaciers, mountains and water (Le Moullec et al. 2019). Sea and fjord ice have facilitated gene-flow between some Svalbard reindeer populations in the past (Peeters et al. 2020), but the extent and frequency of ice cover are rapidly declining, which may lead to further isolation of populations.

As environmental conditions continue to change, genetic diversity of reindeer populations may be important for their capacity to adapt to these conditions and influence their future population dynamics. Therefore, knowledge of the genetic consequences of the Svalbard reindeer reintroduction and recolonisation is important for understanding how the meta-population might respond in future, and to contribute to a broader understanding of the genetic consequences of species reintroductions. Here, I use whole-genome sequencing data to investigate the genetic consequences of Svalbard reindeer reintroductions and compare these to the natural recolonisation process in adjacent, comparable habitats with similar ecological conditions.

Specifically, I test the prediction that the single population bottleneck associated with anthropogenic reintroductions, and the subsequent rapid population growth, would result in populations that retained most of the genetic diversity of the source population, but would still exhibit a signature of this reintroduction in the form of longer runs of homozygosity. Additionally, I predicted that naturally recolonised populations that were not admixed would show lower genetic diversity and higher inbreeding coefficients due to the compounding effects of sequential founding events. To achieve this, I (1) characterised the population structure of reintroduced, remnant, and naturally recolonised populations, and identified any admixture between them; (2) quantified and compared the genetic diversity of reintroduced and source populations to determine how much of the genetic diversity of the source population has been retained in reintroduced populations; (3) compared this to levels of genetic diversity in naturally recolonised populations; and (4) quantified inbreeding and the length distributions of RoH/IBD tracts in individuals' genomes to relate current levels of autozygosity to past demographic processes, and to detect genetic signatures of reintroduction or recolonisation events.

Methods

Study area

The high-arctic Svalbard archipelago lies in arctic ocean approx. 700km north of mainland Norway (76-81° N, 10-35°E). The archipelago is characterised by vegetated peninsulas and valleys comprising only 16% of the landmass (Johansen et al. 2012), which are fragmented by glaciers and mountains that comprise the majority of the land area. Vegetation on the archipelago includes *Dryas*, snowbed, and marsh tundra communities, arctic meadows dominated by forbs and grasses, and barren, exposed graminoid communities (Johansen et al. 2012).

Study species

The Svalbard reindeer is an endemic subspecies that likely colonised the archipelago from Eurasia 6700 - 5000 ago (Kvie et al. 2016). The species is the dominant herbivore in the terrestrial ecosystem with little interspecific competition and almost non-existent predation pressure (Derocher et al. 2000). Reindeer were overharvested to near-extinction on Svalbard during the 19th and early 20th century before coming under legal protection from hunting in 1925 (Le Moullec et al. 2019). By this time, it had been extirpated from much of its former natural range, and isolated remnant populations were largely confined to northern, northeastern, and eastern extremes of the archipelago, as well as the central Spitsbergen region (Le Moullec et al. 2019). After coming under legal protection, the sub-species began to recover but was still absent from much of its range by 1978, including the west coast of Spitsbergen. In 1978, 15 individuals (with 12 surviving the first months, including three males) were translocated from Adventdalen in central Spitsbergen to Brøggerhalvøya on the west coast (Figure 1**,** Aanes et al. 2000). In 1984-85 a second translocation reintroduced 12 individuals to Daudmannsøyra, on the north-western edge of Isfjorden, to the south of Brøggerhalvøya (Gjertz 1995). The reintroduced population at Brøggerhalvøya has been the subject of an ongoing population monitoring program (Aanes et al. 2000; Hansen et al. 2019c). This has recorded population's rapid expansion after translocation (from 12 individuals in 1978 to ~360 individuals in 1993), until a combination of high population density and poor winter conditions triggered a population crash and migration to recolonise the nearby peninsulas of Sarsøyra and Kaffiøyra to the south, and Prins Karls Forland island to the west (Gjertz 1995; Aanes et al. 2000). Reindeer populations have since recolonised most of their former range, including southern Spitsbergen, the north coast of Isfjorden to the east of the reintroduced population at Daudmannsøyra, the north-west coast south to Mitrahalvøya, 15km to the North of Brøggerhalvøya (Le Moullec et al. 2019). Genetic evidence suggests the Svalbard reindeer

metapopulation has low levels of genetic diversity and shows strong population structure (Côté et al. 2002; Kvie et al. 2016; Peeters et al. 2020), reflecting a history of population bottlenecks and the largely philopatric nature of the species with no large scale migration (Hansen et al. 2010a). Populations at Mitrahalvøya, Wijdefjorden, and Southern Spitsbergen appear naturally recolonised from remnant populations, while the origins of the populations along North Isfjorden are unclear, but possibly originated via admixture between the second reintroduction and naturally recolonising individuals (Peeters et al. 2020).

Genetic sample collection

Genetic data were generated from tissue samples (ear, antler, bone, or fur) collected in 2014-18 from 100 individual reindeer originating from twelve sub-populations on the Svalbard archipelago, Norway (Figure 1, Table 1). These included six sub-populations believed to have originated from the two translocations: (1) Brøggerhalvøya (BGR, the initial reintroduction site), Sarsøyra (SAR), Kaffiøyra (KAF), and Prins Karls Forland (PKF) from the first translocation (hereafter collectively referred to as "reintroduction 1"), and (2) Daudmannsøyra (DAU, the second reintroduction site) and North Isfjorden (NIF) from the second translocation (hereafter referred to as "reintroduction 2"). Samples were collected from the source population of the reintroductions (Adventdalen, ADV) and five other remnant or naturally recolonised populations: (Mitrahalvøya (MTR), Southern Spitsbergen (STH), Wijdefjorden (WDF), Eastern Svalbard (EST), and North East Land (NE). All samples were previously used to generate microsatellite data in a study by Peeters et al. (2020), except those from Daudmannsøyra (n=8), which are new in this study.

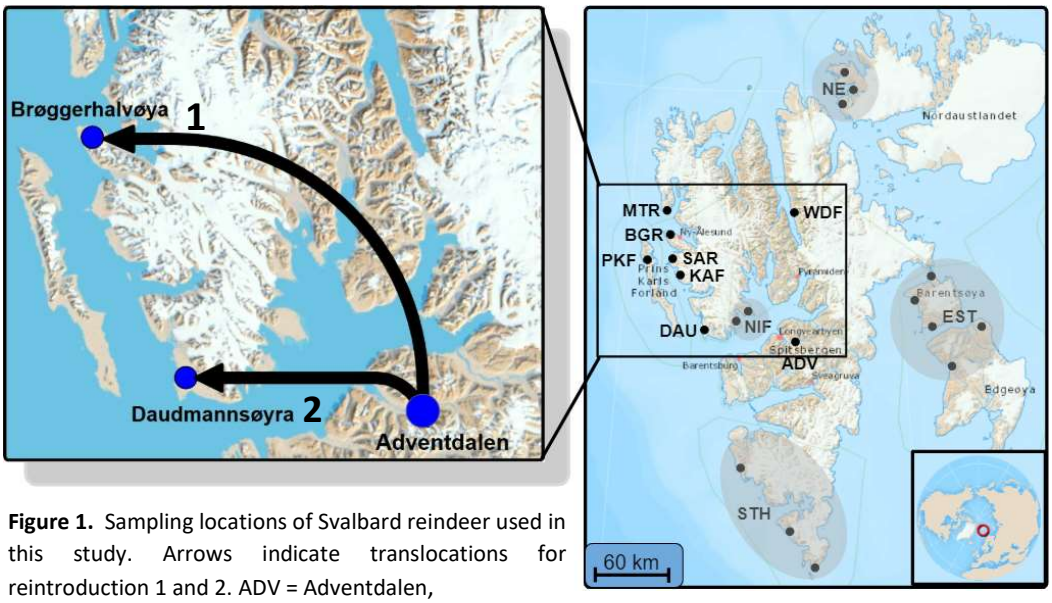


Figure 1. Sampling locations of Svalbard reindeer used in this study. Arrows indicate translocations for reintroduction 1 and 2. ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karls Forland, DAU = Daudmannsøyra, NIF = North Isfjorden, WDF = Wijdefjorden, MTR = Mitrahalvøya, STH = Southern Spitsbergen, EST = Eastern Svalbard, NE = North East Land.

Table 1. Number of samples from each sub-population sequenced, and number used in population structure analyses (NGSadmix, PCAngsd, only), inbreeding analyses (F_{IBD} and Runs of Homozygosity analyses), and heterozygosity estimates.

Sub-population	*n* Seq	*n* Struct[1]	*n* F_{IBD}[2]	*n* RoH[3]	*n* Hetero[4]
Adventdalen (ADV)[S]	17	17	14	16	16
Brøggerhalvøya (BGR)[R]	8	8	8	8	8
Sarsøyra (SAR)[R]	6	6	6	6	6
Kaffiøyra (KAF)[R]	9	9	9	9	9
Prins Karls Forland (KAF)[R]	6	5	3	3	3
Daudmannsøyra (DAU)[R]	8	6	8	8	8
North Isfjorden (NIF)[R]	9	8	8	6	6
Wijdefjorden (WDF)	8	8	8	8	7
Mitrahalvøya (MTR)	8	6	6	5	3
Southern Spitsbergen (STH)	8	6	6	4	3
Eastern Svalbard (EST)	10	9	7	7	5
North East Land (NE)	3*	3	2	2	3
Total	100	91	85	82	77

*Only two samples from NE were mapped to the caribou reference genome; [1] Samples > 0.2x coverage and individuals not closely related; [2] Only samples from low-depth sequencing run and > 0.5x coverage; [3] Samples > 1x coverage; [4] Samples > 1.3x coverage; [S] Reintroduction source population; [R] Reintroduced populations.

DNA extraction and library building

DNA was extracted from ear tissue for the eight samples from Daudmannsøyra using a Qiagen (Hilden, Germany) DNeasy Blood & Tissue extraction kit according to the manufacturer's instructions (details in SI 1.). Genomic library building was based on the BEST v 1.1 (Blunt End Single Tube) 96 well-plate format protocol (Vanessa Bieker, personal communication) based on the method presented in Carøe et al., (2018), and then sequenced to a target depth of 2-3x at the NTNU Genomics Core Facility (see Figure S1 and Table S1 for details). These sequencing data were combined with data from higher-depth (target 20x) sequencing (using the NovasSeq platform) of four of the previously sequenced 94 samples, plus an additional six samples from which libraries were built using the same method.

Bioinformatics

The Paleomix pipeline version 1.2.13.4 (Schubert et al. 2014) was used to map demultiplexed sequence reads to a reindeer reference genome. Adapters were trimmed with adapterremoval version 2 (Schubert et al. 2016) and the BWA aligner program version 0.7.15 was used with the MEM algorithm (Li 2013) without filtering for mapping quality. Reads that did not map to the reference were filtered out. Sequence data was mapped to a reindeer reference genome assembled from a female Mongolian reindeer (Li et al. 2017) as well as a caribou reference genome assembled from a North American male (Taylor et al. 2019). Svalbard reindeer are likely more closely related to Eurasian reindeer populations than North American caribou (Kvie et al. 2016), so the reindeer mapped sequence data was used in all analyses except for RoH inbreeding, for which the overall short length of the assembly scaffolds (N_{50} = 0.94 Mbp) rendered it ineffective. The caribou genome assembly has longer scaffolds (N_{50} = 11.765 Mbp) that are more appropriate for RoH inbreeding analysis, so this reference genome was used instead in these analyses.

There is considerable uncertainty in calling genotypes with low coverage data due to sequencing errors and sampling error of homologous sites. Therefore, no genotype calls are made, and instead the software package ANGSD v0.93 (Korneliussen et al. 2014) was used to generate genotype likelihood data for each individual based on the mapping and base quality scores, and coverage at each loci. This enables the uncertainty inherent in low-coverage sequence data to be properly accounted for in downstream analyses (Korneliussen et al. 2014). Genotype likelihood files were generated in beagle format inferring allele frequencies with fixed major and minor alleles using the command-line arguments *-doGlf 2, -doMajorMinor 1*, and *-doMaf 1*. Variants were called with a threshold of 1e-6, *(-SNP_pval 1e-6*) at sites for which there was sequence data in at least 50 individuals. Reads with mapping quality less than 30 and

base quality less than 20, and those with multiple mapping hits, were filtered out using *-minMapQ 30, -minQ 20,* and *-uniqueOnly 1*, and bad reads were removed with *-remove_bads 1*. Sites on scaffolds mapping to the goat X-chromosome (a list of these scaffolds was provided by Lei Chen, Center for Ecological and Environmental Sciences, NPU) were removed from the data mapped to the Mongolian reindeer reference genome. No removal of sex chromosome linked sites was necessary for the caribou mapped data because all scaffolds mapping to sex chromosomes were very short and therefore not used in the analyses (Taylor et al. 2019) (see RoH inbreeding methods below).

Ancestry/admixture analyses

To infer population structure and identify admixture between populations, the maximum likelihood based clustering analysis software package NGSadmix (Skotte et al. 2013) was used. NGSadmix uses a similar core model to more well-known software packages such as STRUCTURE (Pritchard et al. 2000) and ADMIXTURE (Alexander et al. 2009) but is appropriate for low-coverage WGS data as it uses genotype likelihoods rather than genotype data, and thus incorporates uncertainty due to coverage and sequencing error. Samples with sequencing depth less than 0.2x (n = 7) were removed prior to analysis. Closely related individuals can bias admixture results (Garcia-Erill and Albrechtsen 2020), so the ngsRelate software package (Hanghøj et al. 2019) was used to check relatedness between individuals (using population allele frequencies) for all populations with n > 5. Three individuals (sampled at the same location on the same day) were closely related in the Daudmannsøyra sub-population (coefficient of kinship > 0.7) so the two individuals with the lowest coverage were removed for all population structure analyses. Admixture models were run for the number of genetic clusters (K) ranging from 2 – 10, with 10 replicates of each. Only sites with a minimum minor allele frequency greater than 0.05 (using *-minMaf 0.05*) and that had data in at least half (46) the 92 individuals in the analysis (using *-minInd 46*) were included in the analysis. For each value of K, the replicate with the highest likelihood was chosen.

Evaluation of admixture models

Evaluating and interpreting admixture models can be problematic because they rely on many assumptions to infer ancestry and admixture of individuals, which are often not met in natural populations (Garcia-Erill and Albrechtsen 2020). One key assumption is the number of ancestral populations (that are assumed to be distinct), or K. Typically, this is not known for the study population, so admixture models are run for a range of K values and some model selection procedure is applied find the "optimum" or "true" K value, e.g. the delta-K procedure (Evanno et al. 2005). However, most of these have significant drawbacks, (the 'K conundrum',

Janes et al. 2017), and if the population history is complex, there may be no "true" *K* value (Garcia-Erill and Albrechtsen 2020). To avoid erroneous or arbitrary selection of one "optimum" model, no rule-based model selection procedure was used. Instead all *K* models from 1-10 were considered, and evaluated using the EVALadmix software package (Garcia-Erill and Albrechtsen 2020). This package allows admixture models to be evaluated by calculating and plotting the correlation of residuals between individuals and sub-populations. Positively correlated residuals within a sub-population and negatively correlated residuals among sub-populations can indicate the model underestimates genetic structure (Garcia-Erill and Albrechtsen 2020). Using this approach, admixture models were assessed at the sub-population level by identifying which populations were affected by poor model fit and accounting for this in the interpretation of the results.

Principal component analysis

Principal component analysis (PCA) was conducted by using the software package PCAngsd (Meisner and Albrechtsen 2018) to estimate a genetic covariance matrix using individual allele frequencies based on genotype likelihood data. The genetic covariance matrix was imported into R 3.6.1 (R Core Team 2019) where eigenvectors and eigenvalues were computed using the *eigen* function. Principle components (eigenvectors) were plotted and ellipses based on 95 % confidence intervals were estimated for each natural population, and for each reintroduction group (consisting of reintroduced sub-populations grouped according to their expected reintroduction of origin) using ggplot2 (Wickham 2016). Based on NGSadmix *K*=2 model results, two individuals that represented admixture between divergent populations were not included in population ellipse calculations. PCA analysis was conducted on a dataset including all individuals in the Svalbard meta-population, and then repeated only including individuals from the Adventdalen, Southern Spitsbergen, and reintroduced populations to characterise fine-scale population structure.

F_{ST} analysis

Pairwise F_{ST} values were estimated for each sub-population pair using RealSFS in ANGSD v0.93 (Korneliussen et al. 2014) based on 2D (pairwise) sub-population site frequency spectra (SFS) including all samples. Folded SFS were used because no ancestral states were available to polarize the ancestral/derived alleles. First, BAM files were used to generate unfolded per-site allele frequencies (SAF) for each sub-population using the *-dosaf 1* argument in ANGSD. The SAMtools genotype likelihood model was specified using *-gl 1*, reads with mapping quality less than 30 and base quality less than 20, and those with multiple mapping hits were filtered out using *-minMapQ 30, -minQ 20*, and *-uniqueOnly* respectively, and reads on

scaffolds mapping to the X chromosome were filtered out using the *-rf* argument. The reindeer reference genome was specified as the ancestral genome. Then using the realSFS module in ANGSD v0.93, the unfolded SAF were used to generate folded 2D SFS priors for each pair of sub-populations using the *-fold 1* option. Then both the folded SAFs and the folded 2D SFS prior were used to estimate per-site and global F_{ST} in realSFS, specifying the Hudson estimation method which is more suitable for smaller sample sizes (Bhatia et al. 2013) using *-whichFST 1*. Finally, the weighted global F_{ST} was calculated using the realSFS *fst stat* function.

Heterozygosity

Genome-wide heterozygosity was estimated for each individual using realSFS in ANGSD v0.93 based on the folded site frequency spectrum of each individual (Korneliussen et al. 2014). This followed the same procedure for filtering as described for F_{ST}, and unfolded SAF and then folded SFS were generated separately for each individual. The number of heterozygous sites was divided by the total number of non-N sites to calculate the genome-wide heterozygosity in each individual. Samples with low coverage were found to have upwardly biased heterozygosity estimates (see SI 3 for details), so only samples with coverage above 1.3x were included (n=77) because above 1.3x no relationship between heterozygosity and coverage was detected. To compare mean heterozygosity between reintroduced and source populations, a linear mixed model was performed in R using the lme4 package (Bates et al. 2014). Population group (source, reintroduction 1, reintroduction 2) was used as a fixed factor with individual subpopulation as a random intercept.

Inbreeding and runs of homozygosity

ZooRoH (Bertrand et al. 2019) was used to identify tracts of individual genomes identical by descent (IBD, inherited from a common ancestor of the parents), also known as runs of homozygosity (RoH) and to estimate inbreeding coefficients. This method utilises a hidden Markov model approach to estimate per-site probabilities of being IBD rather than a rule-based method, and can be used with genotype likelihood data, so is more appropriate for use with low-depth WGS data (Druet and Gautier 2017). This model-based approach scans the genome for IBD tracts, and can also partition the proportion of the genome IBD into inbreeding "age" classes (*A)*. IBD tracts are assigned to the most likely age classes based on their length, assuming a recombination rate of 1 cM/Mb and that lengths of IBD tracts are exponentially distributed around (1/*A*)*100 MB, where *A* is the number of generations separating the two homologous segments that are IBD (i.e. two times the number of generations to the common ancestor of the parents). The Mongolian reindeer reference genome has a scaffold N_{50} of only 0.94 (Li et al. 2017), limiting its effectiveness for detecting runs of homozygosity. Therefore,

sequence data was mapped to the more distantly related caribou reference genome with a scaffold N_{50} of 11.77 Mbp (Taylor et al. 2019). Scaffolds shorter than 10 Mbp were excluded from RoH analyses, leaving approximately 56% of the assembled genome (1.235 Gbp) covered by 4,835,080 variable sites. Prior to analysis, the beagle-format genotype likelihoods (GL) were converted to phred-scale likelihoods using the formula $-10*\log_{10}(GL)$. A model with eight predefined inbreeding classes and one non IBD class was specified: A = 8, 16, 32, 64, 128, 256, 512, 1024, corresponding to inbreeding due to common ancestors from four to 512 generations ago. Two inbreeding coefficients were estimated in addition to partitioning of IBD tracts into the different age classes: First, a total inbreeding coefficient (F_{IBD}) was estimated based on the combined proportion of the genome in all inbreeding age classes. This analysis was based on individuals with coverage > 0.5x and excluded high-depth samples from the second sequencing run (total n = 85). Estimates based on non LD pruned data are upwardly biased when including all inbreeding age classes (Bertrand et al. 2019), but LD pruning introduced upward biased estimates in F_{IBD} for high-depth samples from the second round of sequencing (data not shown), so these samples were excluded for this analysis. The genotype likelihood data for F_{IBD} analysis was LD pruned based on called genotypes in PLINK v 1.9 (Chang et al. 2015) using *--indep-pairwise 50 5 0.3* to specify a window size of 50, step size of 5, and a r^2 threshold of 0.3. Secondly, an alternative measure of inbreeding based on the proportion of the genome in ROH segments > 0.3mb (F_{ROH}) was estimated for all individuals with coverage >1x (n = 82).

RoH Simulation

A custom script was written in R 3.6.1 (R Core Team 2019) to simulate runs of homozygosity and inbreeding in a population founded by 12 outbred diploid individuals. Each individual was simulated with one diploid 1 Gbp chromosome consisting of 10,000 evenly spaced markers. The model simulated recombination and Mendelian inheritance at an individual chromosome level, with overlapping generations, random mating and equal probability to reproduce each year dependent on breeding population size (resulting in a stable maximum population size). The script did not simulate separate sexes or non-random individual differences in reproductive success. An age- (but not density) based survival probability was implemented to reduce an excess of older individuals reproducing while still allowing some long generational overlaps, and therefore reduce the chance to overestimate average RoH lengths. Age based mortality was loosely based on data for female Svalbard reindeer in Douhard et al. (2016), but with higher survival for one and two year-olds. Survival probability was arbitrarily reduced in the older age classes because males have lower survival probabilities than females (Reimers 1983), but no age-specific survival data on males was available. Age-based mortality probability under

10 years was not implemented for founding individuals to avoid unrealistically slow population growth in the starting years, accounting for the high resource availability and rapid population growth after reintroduction (Aanes et al. 2000; Kohler and Aanes 2004). A more detailed description of the simulation can be found in SI 2.

Results

Sequencing results

Whole-genome sequencing of 100 individual samples resulted in an average of 2.33x coverage (mapped to the reindeer genome) and 3.03x (mapped to the caribou genome) for the 90 samples sequenced to low target coverage, after quality filtering. The 10 deep-sequenced samples had an average coverage of 19.07x and 24.92x when mapped to the reindeer and caribou genomes respectively (see Figure S1 for distribution of sequencing coverage). Seven samples had < 0.2x coverage and were not used in downstream analyses, except for SAF/SFS estimates. Genotype likelihoods for 8,255,693 variable sites were calculated from the reindeer genome mapped sequence data (6,309,215 after removing scaffolds mapping to the X chromosome) after quality filtering.

Admixture analysis

Admixture analysis identified clear genetic structure in the Svalbard reindeer meta-population (Figure 2). On a broad scale, the K=2 and K=3 models indicated the southern/western (reintroduced populations, Adventdalen and Southern Spitsbergen), northern (Mitrahalvøya) and eastern Svalbard populations originate from different ancestral populations, with Wijdefjorden showing admixture between the Mitrahalvøya and the East / North East Land Svalbard clusters. On a finer scale, reintroduced populations were assigned either entirely or as admixed to one unique genetic cluster (models K=4 – K= 6), or to two (K=7 – K=9) or three (K=10) genetic clusters distinct from each other, the source population in Adventdalen, and the southern Spitsbergen population. Two individuals showed evidence of admixture between reintroduced and naturally recolonised populations; one individual from Kaffiøyra was assigned approximately 50% to the Mitrahalvøya cluster, and one individual in the natural Wijdefjorden population was assigned approximately 50% to the first reintroduction cluster in all K-value models (Figure 2b). Evaluation of the admixture models showed moderate or high correlation between residuals within some populations for all K models < K= 7, which indicated poor model fit (Figure S2). Residuals were positively correlated within some sub-populations and negatively correlated among them in models where reintroduced sub-populations shared a genetic cluster with the Adventdalen or Southern Spitsbergen, indicating the model

underestimated genetic structure. The *K*= 7 model was the model with the lowest number of genetic clusters where no sub-population had average correlation of residuals > 0.1 and is most compatible with PCA and F_{ST} results (see below), so was selected as the most appropriate model.

Principal component analysis

The major axis of variation in the principal component analysis (PC1) appears to explain variation due to genetic divergence between the Adventdalen/South/reintroduced populations and Mitrahalvøya, while the secondary PC axis was driven by variation explained by divergence between the Eastern population and both Mitrahalvøya and Adventdalen/South/reintroduced populations, indicating these three groups were divergent from each other (Figure 3a). On a finer scale, PCA including only the Adventdalen/South/reintroduced populations showed structuring within this group, with Adventdalen and Southern Spitsbergen segregating from each other and reintroduced sub-populations (Figure 3b). Individuals in reintroduced populations segregated according to their reintroduction of origin, and only one individual was outside the 95% confidence ellipse calculated using PC1 and PC2. No structuring among sub-populations within reintroduction groups was evident until PC4, where segregation between most individuals in Daudmannsøyra and North Isfjorden was apparent (Figure S4).

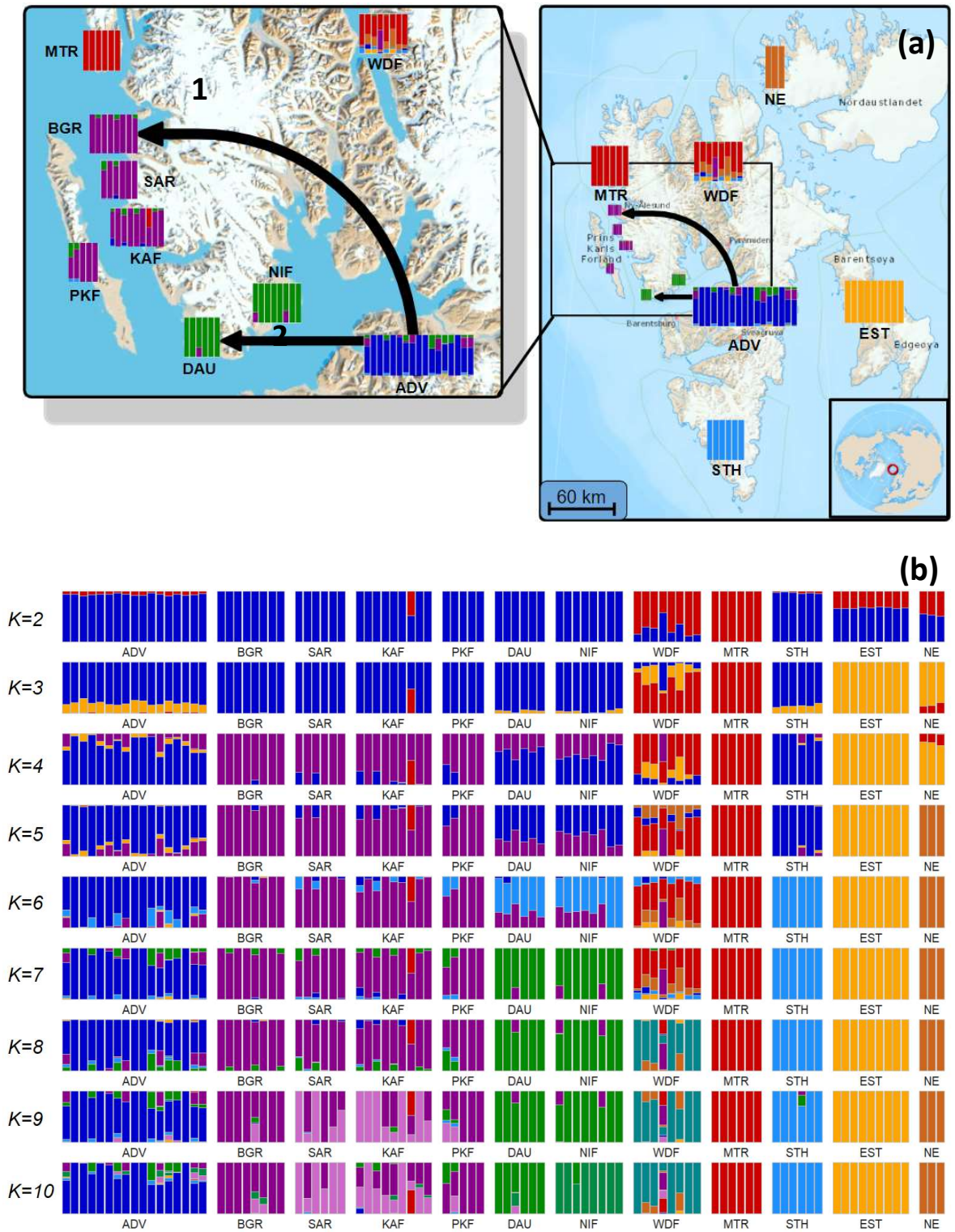


Figure 2. Admixture analysis results from NGSadmix. a) Admixture proportions for model *K*=7 shown at sub-population locations. Arrows indicate translocations for reintroduction 1 and 2; b) Admixture proportions for *K*=2 – 10. Vertical bars represent individual reindeer and colours correspond to genetic cluster assignment. ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karls Forland, DAU = Daudmannsøyra, NIF = North Isfjorden, WDF = Wijdefjorden, MTR = Mitrahalvøya, STH = Southern Spitsbergen, EST = Eastern Svalbard, NE = North East Land.

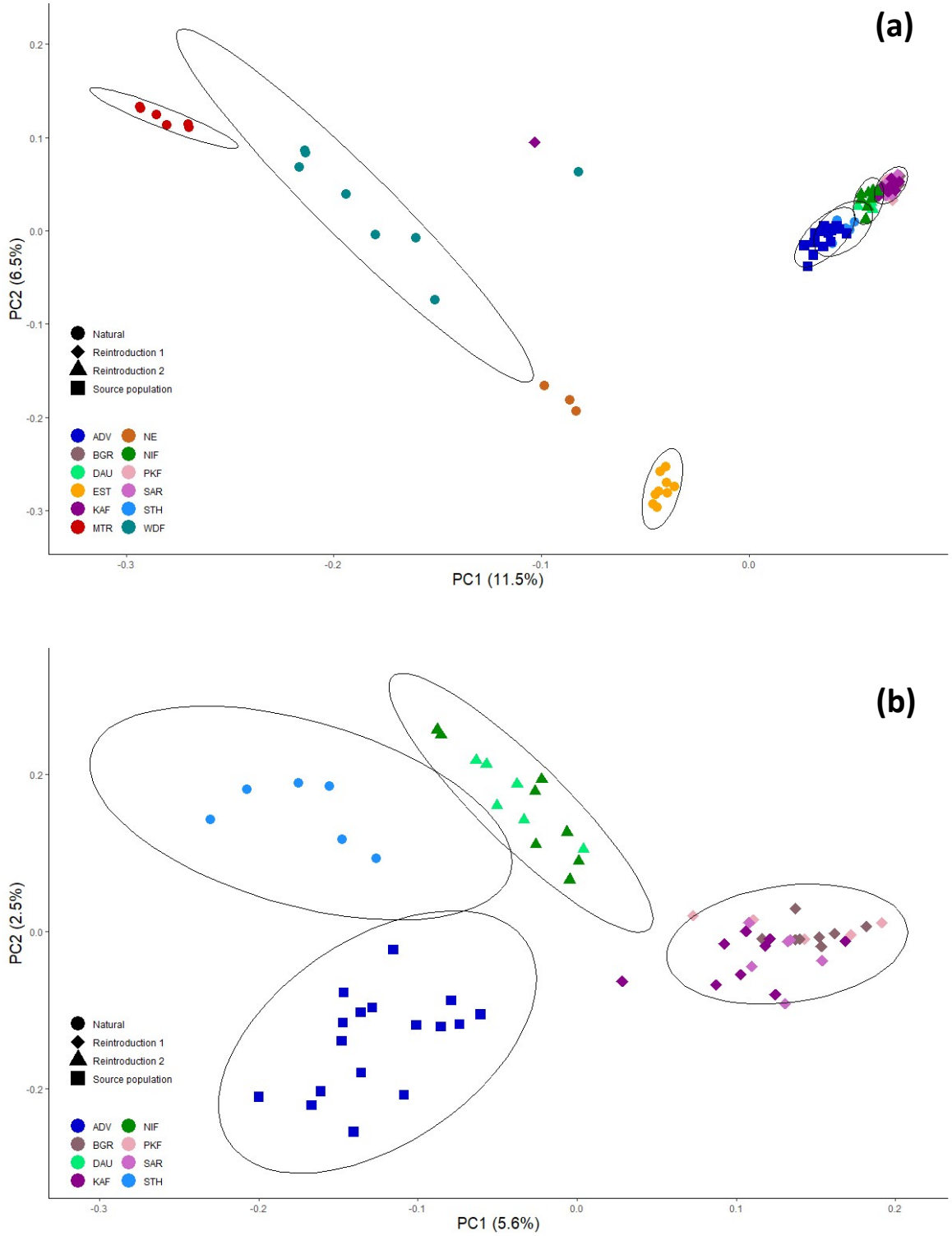


Figure 3. a) PC1 and PC2 from principal component analysis of the Svalbard reindeer meta-population. Variation explained by each PC axis indicated on axis labels. on Ellipses represent normal 95% confidence intervals of population means (Reintro 1 (BGR, SAR, KAF, PKF) and Reintro 2 (DAU, NIF) combined, and all other sub- populations individually). Two admixed individuals were not included in population ellipse estimations, and the North East Land population had too few samples for an ellipse estimate. b) PCA including only the source population, reintroduced populations, and Southern Spitsbergen population. ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karls Forland, DAU = Daudmannsøyra, NIF = North Isfjorden, WDF = Wijdefjorden, MTR = Mitrahalvøya, STH = Southern Spitsbergen, EST = Eastern Svalbard, NE = North East Land.

F_{ST} analyses

Pairwise F_{ST} estimates largely supported admixture and PCA results (Table 2). Sub-populations assumed to originate from the same reintroductions showed lower pairwise F_{ST} values between each other than other sub-populations, but populations within the first reintroduction group (BGR, SAR, KAF, and PKF) showed lower divergence among each other than did the two populations in the second reintroduction group (DDM and NIF). Similar levels of genetic differentiation were found in comparisons between the source population and both groups of reintroduced sub-populations, and between the two groups of reintroduced sub-populations. The naturally recolonised population at Mitrahalvøya appeared to be the most genetically distinct, as the highest F_{ST} values were found between this and all other populations except for Wijdefjorden.

Table 2. Pairwise weighted F_{ST} estimates based on 2DSFS between all sub-populations. ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karl Forland, DAU = Daudmannsøyra, NIF = North Isfjorden, WDF = Wijdefjorden, MTR = Mitrahalvøya, STH = Southern Spitsbergen, EST = Eastern Svalbard, NE = North East Land.

Population	ADV[S]	BG1[1]	SAR[1]	KAF[1]	PKF[1]	DAU[2]	NIF[2]	WDF	MTR	EST	STH
BGR[1]	0.047										
SAR[1]	0.044	0.016									
KAF[1]	0.032	0.009	0.011								
PKF[1]	0.051	0.015	0.014	0.018							
DAU[2]	0.041	0.053	0.051	0.042	0.051						
NIF[2]	0.046	0.048	0.060	0.046	0.058	0.035					
WDF	0.116	0.155	0.151	0.126	0.158	0.149	0.150				
MTR	0.283	0.327	0.327	0.290	0.334	0.324	0.322	0.110			
EST	0.129	0.170	0.169	0.155	0.176	0.167	0.168	0.164	0.326		
STH	0.082	0.108	0.104	0.093	0.109	0.094	0.093	0.174	0.351	0.190	
NE	0.194	0.238	0.237	0.216	0.240	0.232	0.233	0.174	0.359	0.210	0.257

[S] Reintroduction source population; [1] Reintroduction 1; [2] Reintroduction 2

Heterozygosity

The mean genome-wide heterozygosity for all individuals in this study with coverage >1.3x was 0.00090 ± 0.00014 SD, (range 0.00053 – 0.00142, n=77). The mean heterozygosity of all regions (sub-populations originating from each reintroduction are combined) was 0.00089 ± 0.00012. Sub-populations originating from the first and second reintroductions had mean heterozygosities of 0.00089 ± 0.00011 and 0.00088 ± 0.00011 respectively, representing an approximately 6% reduction in mean heterozygosity compared to the source population (Genome-wide heterozygosity 0.00094 ± 0.00008, Figure 4a) although this difference was not statistically significant (mixed model including group (source, reintro1, reintro2) as fixed factor and sub-population as random intercept, $p > 0.1$). The remnant natural population in Eastern

Svalbard had the highest mean heterozygosity (0.00109 ± 0.00021) whilst the naturally recolonised population at Mitrahalvøya had the lowest heterozygosity of all populations (0.00065 ± 0.00011), however this estimate was based on only three individuals.

Inbreeding

The total inbreeding coefficient (F_{IBD}) measured as the proportion of the genome in IBD tracts closely reflects heterozygosity scores (Figure 4b). Total Inbreeding coefficients of individuals in reintroduced populations (Reintro 1 mean F_{IBD} 0.12 ± 0.05 SD, Reintro 2 F_{IBD} 0.10 ± 0.03) were similar to those in the source population, ADV (F_{IBD} 0.11 ± 0.04), and lower than in most natural populations. The naturally recolonised population at Mitrahalvøya had the highest mean F_{IBD} of 0.27 ± 0.06. Most of the total inbreeding coefficient was attributed to RoH < 0.3 Mbp as seen by the lower F_{ROH} inbreeding coefficients (Figure 4c). F_{ROH} was also similar between the source and two reintroduced populations (ADV F_{ROH} 0.008 ± 0.007 SD, Reintro 1 F_{ROH} 0.008 ± 0.008, Reintro 2 F_{ROH} 0.011 ± 0.013), and lower than in all natural populations except North East Land. Southern Spitsbergen had the highest F_{ROH} with 0.042 ± 0.033.

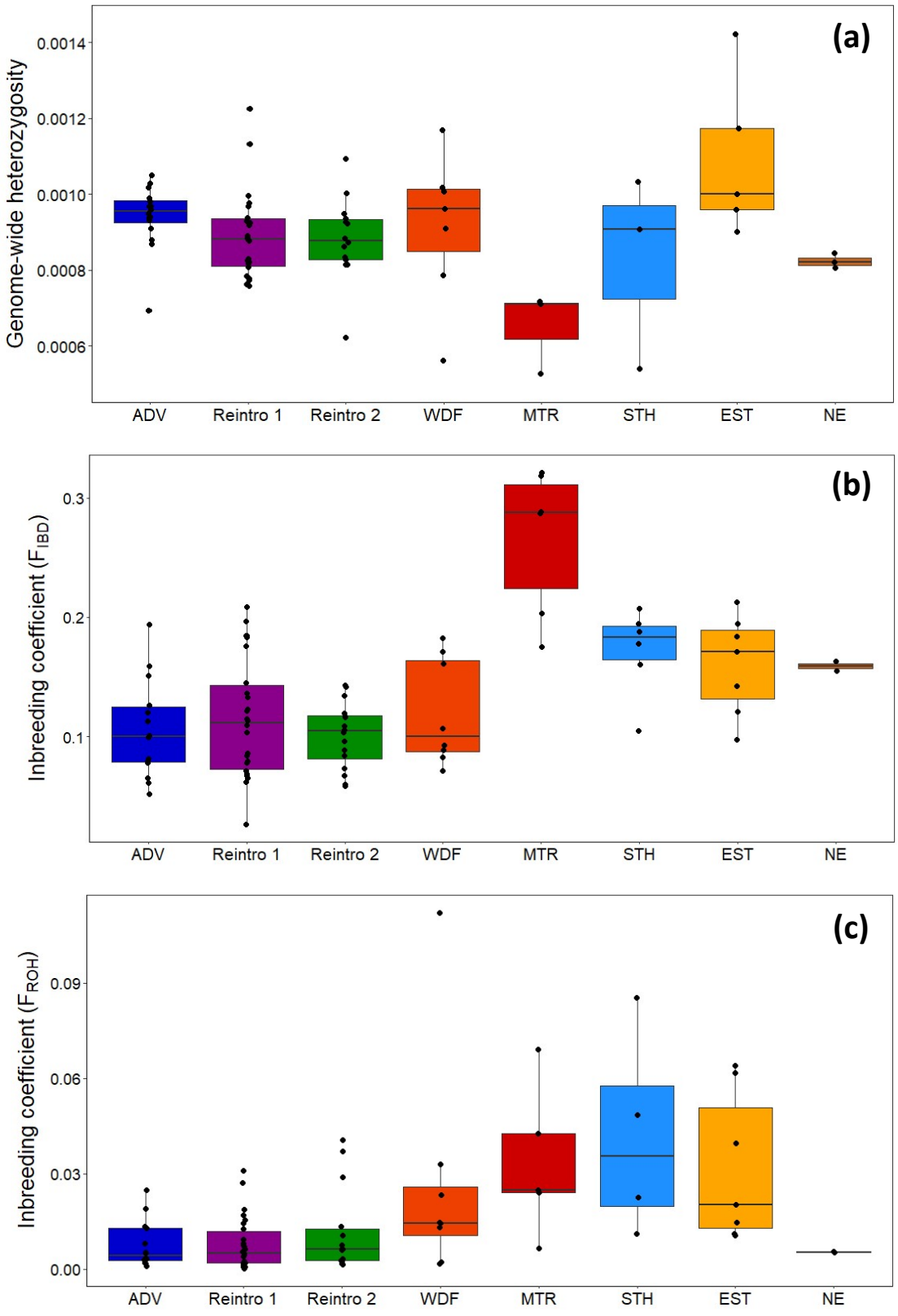


Figure 4. a) Genome-wide heterozygosity for each sub-population. b) Inbreeding coefficient based on proportion of the genome in IBD tracts using whole metapopulation allele frequencies. c) Inbreeding coefficient as proportion of genome in RoH > 0.3 Mbp, estimated using whole metapopulation allele frequencies. Colours correspond to admixture genetic cluster assignment in figure 1. ADV = Adventdalen; Reintro 1 = Brøggerhalvøya, Sarsøyra, Kaffiøyra, and P.K. Forland; Reintro 2 = Daudmannsøyra and North Isfjorden; WDF = Wijdefjorden; MTR = Mitrahalvøya; STH = Southern Spitsbergen; EST = Eastern Svalbard; NE = North East Land.

Distribution of RoH lengths and simulated inbreeding

The ZooRoH age-class inbreeding model assigned all inbreeding to common ancestors > 512 generations ago. The longest RoH detected using the ZooRoH analysis was 1.77 Mbp, found in an individual at Mitrahalvøya, and in total only 16 RoH > 1 Mbp were found in seven individuals. Only one of these individuals was in a reintroduced sub-population (Daudmannsøyra), while remaining six belonged to naturally recolonised populations. Reintroduced populations showed only a minor increase in RoH lengths compared to the source population. In contrast, simulated reintroduced populations originating from twelve outbred founders had exponentially distributed RoH lengths (Figure 5b) with mean length in individuals of 7.4 Mbp (+/- 3.2 SD) and mean maximum RoH length in individuals of 20.1 Mbp (+/-11.9 SD). These RoH resulted in a mean F_{ROH} attributable to shared common ancestry from the founding individuals of 0.059 (+/-0.037 SD), equivalent to a mean reduction in heterozygosity of 5.9%.

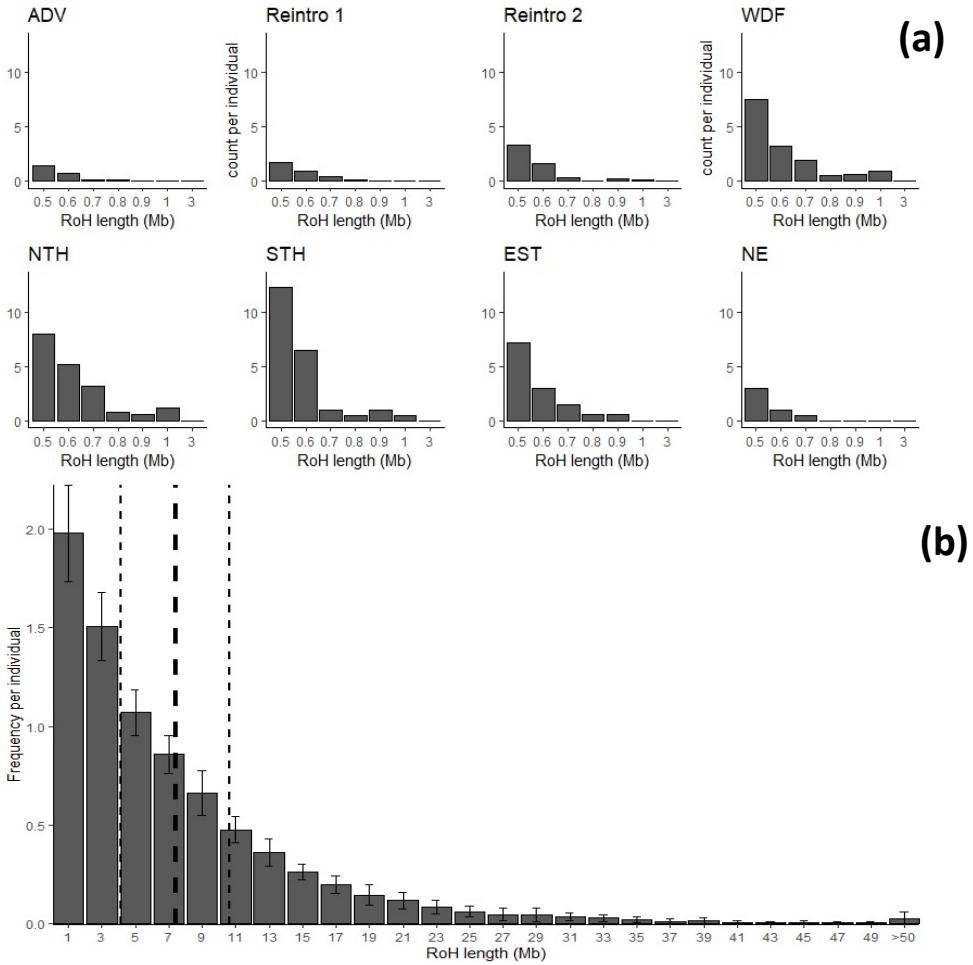


Figure 5. a) Frequency of runs of homozygosity > 0.5 Mbp long identified using ZooRoH in individuals with greater than 1x coverage. b) Simulated RoH length distributions of individuals produced in year 40 from 20 simulation runs. Error bars indicate SD of population means among runs. Dashed line indicates individual mean RoH length +/- SD. ADV = Adventdalen; Reintro 1 = Brøggerhalvøya, Sarsøyra, Kaffiøyra, and P.K. Forland; Reintro 2 = Daudmannsøyra and North Isfjorden; WDF = Wijdefjorden; MTR = Mitrahalvøya; STH = Southern Spitsbergen; EST = Eastern Svalbard; NE = North East Land.

Discussion

By using whole-genome shotgun sequencing of Svalbard reindeer, this study has demonstrated important population genomic consequences of anthropogenic reintroductions versus natural recolonizations. I found evidence of significant genetic drift resulting from reintroductions (Figures 2, 3, Table 2), with reintroduced populations forming two distinct genetic clusters corresponding to the two reintroductions (Figures 2, 3b). There was little evidence for gene flow between reintroduced and natural populations, with very high genetic divergence between a reintroduced and a naturally recolonised population only separated by ~15 km (Figure 2a, Table 2). However, reintroduced populations showed no significant decrease in heterozygosity or increase in inbreeding compared to the source population (Figure 4), and no signature of the reintroduction event could be observed in Runs of Homozygosity length distributions (Figure 5). In contrast, some naturally recolonised populations showed very high inbreeding (F= > 0.3), longer ROH segment distributions), and reduced genetic diversity (Figures 4, 5).

Admixture, principal component, and F_{ST} analyses identified a high degree of structuring in the Svalbard reindeer metapopulation**,** typical of ungulate populations with a history of population fragmentation and bottlenecks due to past harvesting pressure (Williams et al. 2002; Haanes et al. 2010). On a broad-scale, results show strong genetic differentiation between populations from Southern/Central Svalbard, Mitrahalvøya, Eastern Svalbard, and North East Land, in agreement with results based on a microsatellite analysis by Peeters et al. (2020). The close association between the source population in Adventdalen, the reintroduced populations, and the Southern Spitsbergen population is congruent with all having origins in Central Spitsbergen, and forming a broad "Central Svalbard" genetic group. Reindeer were apparently extirpated in Southern Spitsbergen until after the 1950s, but their distribution in the region is now almost continuous with that of central Svalbard (Le Moullec et al. 2019). Given the genetic evidence presented here and in Peeters et al. (2020), it appears the region has been naturally recolonised from Central Spitsbergen, resulting in genetic drift that has driven the observed divergence from the Adventdalen population. The high degree of genetic structure is likely to be a consequence of both: (1) the philopatric behaviour of Svalbard reindeer which, unlike most other reindeer and caribou populations, do not undertake large-scale migration (Hansen et al. 2010a); and (2) the geographical isolation of suitable habitats, which are separated by geographical features such as mountain ranges and glaciers that inhibit dispersal of reindeer (Peeters et al. 2020). In addition, severe overharvesting during the 19-20th centuries resulted in remaining populations becoming small and isolated (Le Moullec et al. 2019), which may have accentuated genetic drift and subsequent divergence.

On a finer scale, this study reveals population structure within the Central Svalbard group between the source and reintroduced populations, and among reintroduced sub-populations. Admixture and principal component analyses indicate the presence of two distinct genetic clusters among reintroduced populations corresponding to the two separate reintroductions, and pairwise F_{ST} estimates indicate both reintroductions have resulted in a similar degree of genetic divergence from the source population. Founder effects and subsequent genetic drift resulting in genetic divergence between reintroduced populations and their source has been documented in several reintroductions (Williams et al. 2002; Latch and Rhodes 2005; Brekke et al. 2011; Andersen et al. 2014; Grossen et al. 2018), which reflects isolation from the source population after reintroduction. However, such patterns do not always occur; a lack of population divergence between reintroduction and source populations is more typical when high levels of gene flow occur between them e.g. Latch and Rhodes (2005) and Drauch and Rhodes (2007).

Close genetic clustering of multiple sub-populations colonised by individuals from a common reintroduced population is a common characteristic of reintroduction programmes (Andersen et al. 2014; Grossen et al. 2018). The first reintroduction group of sub-populations (BGR, SARS, KAFI and PKF) appears almost homogenous, with low pairwise F_{ST} estimates between all pairs, no segregation among sub-populations in the PCA analysis, and no genetic clusters exclusive to any sub-population in any of the admixture models. The low level of genetic structure is likely due to dispersal between these sub-populations, and the relatively recent re-colonisation; after reintroduction in 1978, reindeer were restricted to BGR until poor winter conditions and high population density (Aanes et al. 2000; Kohler and Aanes 2004) triggered migration to PKF (Gjertz 1995), SAR, and KAF (Aanes et al. 2000) in 1994. Population monitoring since the initial reintroduction has identified dispersal between BGR, SAR, and KAF (Hansen et al. 2009; Stien et al. 2010) which would be expected to reduce divergence between these populations (Wright 1931). The second reintroduction group ("reintroduction 2") also shows lower pairwise F_{ST} among its two sub-populations (DAU and NIF) than between these and any other populations, although they are more differentiated than those within the first reintroduction group. A PCA examining only the central Svalbard group shows segregation between these two sub-populations on PC4 (Figure S4) and admixture models assign NIF as either admixed or in an entirely separate genetic cluster to DAU in the K=10 model.

The cause of the higher population divergence among the two sub-populations within reintroduction 2 than among those in reintroduction 1 is unclear, and no long-term population

monitoring program exists for DAU or NIF to provide additional insight. Stronger genetic drift as a result of lower migration rates, smaller populations sizes, a longer time since the colonisation of NIF from DAU compared to colonisation of sub-populations in reintroduction 1, and/or admixture between NIF and an unsampled population to the east are all potential causes for this. Reduced gene flow between DAU and NIF could be due to a greater separation by water and other landscape features between areas of suitable foraging habitat (see figure 1 in Le Moullec et al. 2019) inhibiting dispersal between the two sub-populations. Alternatively (or additionally), a low level of gene flow between NIF and unsampled populations further east along the north coast of Isfjorden is likely. Reindeer are distributed on peninsulas and in valleys along the whole north coast of Isfjorden (Le Moullec et al. 2019), and the peninsulas further east are more likely to be connected by sea ice (Peeters et al. 2020), which occurs more frequently further east (inland) in the fjord (Muckenhuber et al. 2016). Sea ice may therefore be more likely to facilitate dispersal between NIF and unsampled populations to the east than it is to bridge the barriers to gene flow between DAU and NIF.

Similar patterns of genetic structure as seen here in Svalbard reindeer have also been found in systems with multiple reintroductions in both alpine ibex (Grossen et al. 2018) and large blue butterflies (Andersen et al. 2014), where multiple secondary sub-populations founded by individuals from the same primary reintroduced populations form genetic clusters with their source population, distinct from other groups of reintroduced sub-populations. Such patterns of genetic structuring are likely driven by founder effects and a lack of gene flow between the primary reintroduced populations (Hicks et al. 2007; Andersen et al. 2014; Grossen et al. 2018).

A notable discrepancy between F_{ST} analysis, admixture, and (to a lesser extent) the principal component analyses is the relationship between DAU, NIF, Southern Spitsbergen and the source population in Adventdalen. F_{ST} analysis indicates Adventdalen, reintroduction 1 and reintroduction 2 are all approximately equally differentiated to each other, and more differentiated to Southern Spitsbergen population. However, DAU and NIF (models K=4 – K=6) are assigned as admixed between reintroduction 1 and Adventdalen and/or Southern Svalbard whereas reintroduction 1 is assigned to a separate genetic cluster. This discrepancy is likely due to sample size bias. Both PCA (McVean 2009) and admixture-type analyses (Lawson et al. 2018) with unequal sample sizes can result in primary PC axes or admixture models that separate larger populations first, even if genetic differentiation may be higher between two other smaller populations because doing so explains more of the total variance in the genetic data. In this case, DAU and NIF may be more similar to Southern Spitsbergen at loci which explain more genetic variance than the genome-wide average, and result in them

appearing more closely related in PCA and admixture analyses despite being equally differentiated when measured across all loci.

Admixture analysis indicates very little gene flow between reintroduced populations and sampled natural populations. Only one individual in a reintroduced population appears to be admixed with a natural population, with admixture proportions consistent with an F1 offspring resulting from a mating between individuals in reintroduction 1 and Mitrahalvøya genetic clusters. Mitrahalvøya and Brøggerhalvøya, the closest subpopulation in reintroduction 1, are only approximately 15km apart across the mouths of Kongsfjorden and Krossfjorden. The outer regions of these fjords rarely freeze over in winter due to warm Atlantic currents and exposure to ocean swell (Pavlova et al. 2019), which has likely contributed to the lack of gene flow and maintained a high degree of genetic differentiation between these geographically proximate populations. Admixture between the two reintroduction groups is more difficult to elucidate due to the shared recent ancestry of these populations. Some individuals in all reintroduced sub-populations show low to moderate levels of assignment to the opposite reintroduction group, however these should be interpreted with caution, as many individuals from the reintroduced populations are also assigned partially to the source population at Adventdalen. Given the geographical distance and barriers to dispersal between Adventdalen and the reintroduced populations (Peeters et al. 2020), it is highly unlikely to represent migration and is more likely an artefact of their shared ancestry. Interpreted in this light, there is possible admixture between the reintroductions, but at levels low enough that genetic differentiation between them has been maintained.

Small and isolated populations (such as reintroduced populations) are susceptible to accumulation of inbreeding and erosion of genetic diversity (Frankham 2005b; Frankham et al. 2017). Despite their isolation however, the reintroduced populations had comparable levels of heterozygosity to the source population and showed no increase in inbreeding based on F_{IBD} (which closely mirrored heterozygosity results) or F_{ROH} results. Genome-wide heterozygosity estimates showed a non-significant ~6% reduction in heterozygosity for both reintroduction 1 and 2 in this study, which differs markedly from a ~23% reduction of microsatellite heterozygosity in this population found by Peeters et al. (2020). However, allelic richness may more accurately predict genome-wide heterozygosity than microsatellite heterozygosity (Fischer et al. 2017), and a 5% reduction in allelic richness found by Peeters et al. (2020) is concordant with the results in this study. Rapid population growth immediately after reintroduction (Aanes et al. 2000), and the long generational overlap in reindeer, are both characteristics that act to reduce the loss of genetic diversity after population bottlenecks (Nei

et al. 1975). These characteristics probably contributed to the maintenance of diversity in the reintroduced populations. In the case of a single-generation population bottleneck of 12 individuals, a small loss of genetic diversity similar to that observed in these reintroductions is expected (Allendorf 1986). Similar outcomes have been observed in many other reintroductions, particularly those that experience rapid population growth after reintroduction (Hicks et al. 2007; Brekke et al. 2011; Murphy et al. 2015; White et al. 2017). Ungulate populations colonised by only a few founding individuals have been shown to retain high levels of heterozygosity as a result of overlapping generations and rapid population expansion (Kaeuffer et al. 2007; Kekkonen et al. 2012), whereas those which have remained at small population sizes for several generations after reintroduction often display a pronounced reduction in diversity (e.g. Williams et al. 2002; Wisely et al. 2008). Furthermore, populations which have experienced past bottlenecks and have low genetic diversity , as is the case for Svalbard reindeer (Kvie et al. 2016), may be less susceptible to loss of genetic diversity in future bottlenecks because they have fewer rare alleles (Taylor and Jamieson 2008; Wright et al. 2014). However, this should be more likely to reduce loss of allelic diversity than heterozygosity (Allendorf 1986). These results indicate that a founding population size of twelve individuals has been sufficient to maintain most of the heterozygosity of the source population and avoid serious accumulation of inbreeding in both reintroductions, both of which are a key concern for reintroduced populations (Weeks et al. 2011; Frankham et al. 2017).

Relatively low sample sizes in the naturally recolonised populations (Mitrahalvøya, Southern Spitsbergen, and Wijdefjorden) limit the confidence with which inferences can be made. However, there was a high degree of variation in heterozygosity and inbreeding among these populations which corresponds to variation in their origins and demographic histories. Colonising populations that are admixed between divergent lineages tend to show higher genetic diversity than those originating from a single source (Kolbe et al. 2008; Havrdová et al. 2015), and this is reflected in some naturally recolonising Svalbard reindeer. The population at Wijdefjorden appears to be admixed between highly divergent genetic clusters (primarily the northern genetic cluster [closely related to Mitrahalvøya] and the North-East land cluster) and has higher genetic diversity and lower inbreeding than Mitrahalvøya and Southern Spitsbergen, both of which appear to originate from a single source (Peeters et al. 2020).

Signatures of reintroduction in the form of long RoH have been identified in several reintroduced populations of mammal species. Druet et al. (2020) found two populations of European bison, descended from 7 and 12 founders respectively, had a high proportion of the

genome in (>0.2) in RoH segments likely tracing back to common ancestors 4 – 8 generations ago. Grossen et al. (2018) found long RoH segments (>20 Mbp) in up to 50% of individuals in reintroduced (but not source) European ibex populations. In contrast to these studies, no long RoH were detected here in any population, and the ZooRoH inbreeding model did not allocate inbreeding in any sub-population to recent inbreeding age classes. Hence, no clear signature of the reintroduction bottleneck (e.g. inbreeding from < 10 generations in the past) or bottlenecks associated with natural recolonisation since the end of hunting (< 30 generations in the past) could be detected. Indeed, all inbreeding was inferred to be due to common ancestors over 500 generations in the past. This hindered efforts to tease apart the relative contribution of recent inbreeding (that could be associated with recolonisation bottlenecks) and inbreeding due to small population sizes and low genetic diversity in populations' more distant past. Thus, it is unclear to the extent recolonisation has contributed to the extremely low level of genetic diversity and high inbreeding in the Mitrahalvøya population, because the low genetic diversity of its source population in North-Western Spitsbergen (Peeters et al. 2020) is probably at least partially responsible.

Admixture and population structure analyses suggest the Southern Spitsbergen population originates from the same (or a closely related) ancestral population as reintroduced individuals (Adventdalen) allowing a more direct comparison of reintroduction vs recolonisation. The higher levels of inbreeding in the Southern Spitsbergen population may also reflect multiple bottlenecks and isolation from a sequential recolonisation process. Sequential recolonisation of suitable peninsulas and valleys from the remnant natural source populations may have resulted in cumulative effects from multiple founder-event (Le Corre and Kremer 1998), reducing effective population sizes and eroding genetic diversity (Clegg et al. 2002; Pruett and Winker 2005). Although no long RoH were found in any populations, naturally recolonised populations still had longer RoH distributions and a greater proportion of the genome in RoH longer than 0.3 Mbp than the reintroduced and source populations. This lends support to the idea that the recolonisation process has resulted in populations with smaller effective population sizes in their recent history and thus greater inbreeding compared to reintroduced populations. Ongoing gene flow among nearby sub-populations recolonised from the initial reintroduced populations (Hansen et al. 2010a; Stien et al. 2010) may have buffered against sequential founder effects accumulating to the same extent as in the naturally recoloinsed populations.

The lack of long RoH is unexpected given the known histories of these populations, and given that much longer RoH are usually found in other species with known population bottlenecks

(Kardos et al. 2017; Sinding et al. 2018). To confirm that the distribution of RoH identified in the RoH analysis did indeed deviate from that expected given this specific (known) population history, a simplistic simulation of the reintroduction was conducted. Simulation results suggest given the limited number of founders, the parents would share all or most of the 12 reintroduced reindeer as common ancestors, and thus share many identical segments of their genome. Reindeer can reproduce from about three (occasionally two) until over 12 (and up to 19) years of age (Lee et al. 2015; Douhard et al. 2016) this suggests that any segment of genetic material probably passed through between three and 13 generations (i.e. on average between three and 13 recombination events per 100cM) from the reintroduction founders to individuals born 40 years after reintroduction (i.e., in 2018). The longest RoH segment found in any reintroduced individual was 1.7 Mbp, which (assuming a recombination rate of 1cM/1Mbp), would be the expected mean length for RoH segments formed due to a common ancestor approx. 28 generations previously. However, the length of a single RoH segment is unlikely to be a reliable indicator of its generation of origin (age of inbreeding); RoH originating from common ancestry in any given generation in the past are expected to have an exponential distribution of lengths (Druet and Gautier 2017), thus the distribution of RoH lengths (which suggests no inbreeding in an age class as recent as 28 generations) is a more appropriate indicator of inbreeding age (Druet and Gautier 2017).

The simulation study predicted the reintroduction bottleneck to have resulted in approximately 6% of the genome to be in RoH exponentially distributed around a mean of 7.4 Mbp, with each individual on average carrying at least one RoH 20Mbp long in every 1 Gbp of genome. This was in agreement with the 6% reduction in heterozygosity found in both reintroduced populations. However, results here found no reintroduced individual had even 5% of their genome in RoH of 0.3 Mbp, a mean length of which would correspond to common ancestors of the parents over 150 years ago. The ZooRoH inbreeding partitioning model allocated no inbreeding to any age class more recent than 512 generations ago, which suggested no significant inbreeding from common ancestors within hundreds of generations, and highlights a major discrepancy between expectations and the RoH analysis results.

This discrepancy could be due to a number of factors. First, recombination rates are variable which can influence RoH length (Kardos et al. 2017) and it is plausible that the recombination rate in Svalbard reindeer is higher than 1 cM/Mbp. No data has been published on reindeer recombination rates, but a related cervid (red deer) has a genome-wide average recombination rate of 1.04 cM/Mbp (Johnston et al. 2017). Another factor likely to reduce observed RoH lengths in the Svalbard reindeer data compared to the simulation is the reference genome

assembly. The simulation assumed a single chromosome 1000 Mbp in length. This is longer than real reindeer chromosomes, of which there are 34 (Gripenberg et al. 1991) in an approximately 2.76 Gbp genome (Li et al. 2017), and thus allowed unrealistically long RoH to occur. Only scaffolds with a minimum length of 10 Mbp were used for RoH analyses, the median length being 16.4 Mbp. Short scaffolds increase the likelihood of RoH being broken by scaffold edges and could be expected to reduce the mean and maximum observed RoH lengths to some extent (Brüniche-Olsen et al. 2018). However, the effects of both higher-than-simulated recombination rate and limited scaffold length should contribute to a greater number of shorter RoH, and even relatively short RoH (~1 Mbp) are absent in the reintroduced populations, suggesting these issues are not the primary factor responsible for the discrepancy in expected and observed RoH distributions.

Immigration of unrelated individuals after the reintroduction could explain the short distribution of RoH, as reindeer would share fewer common ancestors from recent generations. However, the population structure and admixture analyses show little evidence of admixture in most reintroduced sub-populations, although a low level of admixture between them cannot be ruled out. Given the known population history from long term monitoring (Aanes et al. 2000; Hansen et al. 2019c), results from admixture analysis and F_{ST} comparisons, and the simulation based on population genetics theory, the RoH inbreeding analysis results remain unexplained and should be treated with caution. Technical factors related to the analysis could also potentially confound RoH analysis, such as higher than expected sequencing error rates (Brüniche-Olsen et al. 2018), biased genotype likelihood estimates (Fuentes-Pardo and Ruzzante 2017) or errors in mapping to the reference genome assembly. No RoH inbreeding analysis on any reindeer or caribou population could be found for comparison to this study.

Svalbard reindeer face rapid environmental change due to climate change (Hansen et al. 2019c), which could lead to maladaptation in these populations. For example, the changing climate has resulted in an increase in winter precipitation or "rain-on-snow" events which increases ground ice cover during winter (Peeters et al. 2019), and lead to some coastal reindeer populations adopting novel foraging strategies by eating kelp (Hansen et al. 2019b). Climate change is also altering the abundance and distributions of pathogens and parasites in the arctic, and therefore influence host-pathogen interactions (Kutz et al. 2009; Hueffer et al. 2011). The target and strength of any selection acting on Svalbard reindeer have not been measured, however mortality and reproductive output are both highly related to basal-ice conditions as a result of winter-rainfall ("rain on snow"), and diet switching appears to be a plastic response to food shortages in winter, with coastal reindeer feeding on kelp in years with high ground ice

coverage (Hansen et al. 2019b). Furthermore, host-pathogen dynamics may play an additional role in Svalbard reindeer population dynamics (Albon et al. 2002), so selection on phenotypes that influence tolerance to these novel conditions is plausible. Any evolutionary responses to changing environmental conditions will likely depend on sufficient standing genetic variation in populations for natural selection to act upon given the long generation time of reindeer (Carlson et al. 2014).

Adaptation may be driven by selection on quantitative traits (i.e. controlled by many genes of small effect, e.g. Grant and Grant 2006), or by single mutations that have large effects (e.g. Colosimo et al. 2005; Pelz et al. 2005). Additive genetic variance in quantitative traits is less sensitive to loss of rare alleles than heterozygosity in general (Lande 1980) so reintroduced Svalbard reindeer populations may have incurred little reduction in additive genetic variance for quantitative traits, however genetically depauperate remnant or naturally recolonised populations may have reduced evolutionary potential. Rare alleles, which contribute relatively little to heterozygosity, are much more likely to be lost during population bottlenecks or founder events (Nei et al. 1975; Allendorf 1986). Simulations by (Weiser et al. 2013) showed a red deer population with a carrying capacity of 100 individuals had only a 34% probability of retaining a rare allele (frequency of 0.05) 10 generations after a bottleneck size of 40, despite this bottleneck having only a small effect on heterozygosity. Empirical studies based on microsatellites have revealed that reintroduced populations can maintain similar levels of heterozygosity as source populations but have reduced allelic richness (White et al. 2017). Therefore, both reintroduced and naturally recolonised populations may be less likely to carry rare alleles that could be beneficial for their long-term viability. For example, population bottlenecks have been found to reduce the variability in MHC genes in other species (Radwan et al. 2007; Ejsmond and Radwan 2011), which may be important for parasite and disease resistance (Radwan et al. 2010; Lee et al. 2011) and loss of MHC gene diversity has been implicated in population declines due to disease in other species (Siddle et al. 2007). Other adaptive alleles have been found to exist as standing variation at low frequency in populations of other species prior to the environmental changes (Colosimo et al. 2005; Pelz et al. 2005; Jones et al. 2013). Recolonised sub-populations with eroded genetic diversity and higher inbreeding coefficients such as the population at Mitrahalvøya may warrant particular attention in future because they are likely to be the most limited in their ability to adapt to changing environmental conditions, and stressful conditions can exacerbate the fitness consequences of inbreeding (Kristensen et al. 2008; Fox and Reed 2011).

In conclusion, the results of this study indicate that translocations of Svalbard reindeer to found two reintroduced populations has been successful in establishing populations with comparable genetic diversity to the source population, and higher than some naturally recolonising populations. This supports other research showing that in "ideal" ecological conditions, such as those likely experienced by Svalbard reindeer populations at the time of reintroduction, rapid population growth after reintroduction is sufficient to maintain the genetic diversity of even relatively small founder populations (Kekkonen et al. 2012). These findings have relevance for the conservation and management of other species with similar life histories (e.g. other ungulates) where reintroduction is being considered or where natural recolonisation from fragmented populations is expected. The lower levels of genetic diversity and higher inbreeding in naturally recolonised populations originating from a single source is likely to be related to sequential founding events, although the RoH-based inbreeding analyses in this study failed to identify any signatures of recent inbreeding that could be associated with founding events. Importantly, however, despite the successful reintroduction of Svalbard reindeer to the west coast, the meta-population as a whole still appears fragmented, and connectivity between isolated populations may degrade further due to loss of sea ice that provides corridors for dispersal between populations (Peeters et al. 2020). Some of the results in the present study are based on small sample sizes, and more data is needed on remnant and naturally recolonised populations to make robust assessments on how the genetic consequences of reintroduction compare to recolonisation. However, these results suggest populations naturally recolonised from already genetically depauperate remnant source populations may be particularly vulnerable to the accumulation of inbreeding and drift-load. These populations are therefore most likely to be limited in their capacity to adapt to the rapidly changing environment, and thus warrant particular attention in future monitoring or research programmes.

References

Aanes R, Sæther BE, Øritsland NA (2000) Fluctuations of an introduced population of Svalbard reindeer: The effects of density dependence and climatic variation. Ecography (Cop.). 23:437–443

Albon SD, Stien A, Irvine RJ, et al (2002) The role of parasites in the dynamics of a reindeer population. Proc R Soc B Biol Sci 269:1625–1632. https://doi.org/10.1098/rspb.2002.2064

Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. Zoo Biol 5:181–190. https://doi.org/10.1002/zoo.1430050212

Andersen A, Simcox DJ, Thomas JA, Nash DR (2014) Assessing reintroduction schemes by comparing genetic diversity of reintroduced and source populations: A case study of the globally threatened large blue butterfly (Maculinea arion). Biol Conserv 175:34–41. https://doi.org/10.1016/j.biocon.2014.04.009

Armstrong DP, Seddon PJ (2008) Directions in reintroduction biology. Trends Ecol Evol 23:20–25. https://doi.org/10.1016/j.tree.2007.10.003

Bates DM, Maechler M, Bolker BM, Walker S (2014) Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw arXiv:1406.5823. https://doi.org/10.18637/jss.v067.i01

Bell G (2013) Evolutionary rescue and the limits of adaptation. Philos Trans R Soc B Biol Sci 368:1–6. https://doi.org/10.1098/rstb.2012.0080

Bell G, Gonzalez A (2009) Evolutionary rescue can prevent extinction following environmental change. Ecol Lett 12:942–948. https://doi.org/10.1111/j.1461-0248.2009.01350.x

Bellard C, Bertelsmeier C, Leadley P, et al (2012) Impacts of climate change on the future of biodiversity. Ecol Lett 15:365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Bertrand AR, Kadri NK, Flori L, et al (2019) RZooRoH: An R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. Methods Ecol Evol 10:860–866. https://doi.org/10.1111/2041-210X.13167

Bhatia G, Patterson N, Sankararaman S, Price AL (2013) Estimating and interpreting FST: The impact of rare variants. Genome Res 23:1514–1521. https://doi.org/10.1101/gr.154831.113

Biebach I, Keller LF (2010) Inbreeding in reintroduced populations: The effects of early reintroduction history and contemporary processes. Conserv Genet 11:527–538. https://doi.org/10.1007/s10592-009-0019-6

Biebach I, Keller LF (2012) Genetic variation depends more on admixture than number of founders in reintroduced Alpine ibex populations. Biol Conserv 147:197–203. https://doi.org/10.1016/j.biocon.2011.12.034

Brekke P, Bennett PM, Santure AW, Ewen JG (2011) High genetic diversity in the remnant island population of hihi and the genetic consequences of re-introduction. Mol Ecol 20:29–45. https://doi.org/10.1111/j.1365-294X.2010.04923.x

Brüniche-Olsen A, Kellner KF, Anderson CJ, DeWoody JA (2018) Runs of homozygosity have utility in mammalian conservation and evolutionary studies. Conserv Genet 19:1295–1307. https://doi.org/10.1007/s10592-018-1099-y

Carlson SM, Cunningham CJ, Westley PAH (2014) Evolutionary rescue in a changing world. Trends Ecol Evol 29:521–530. https://doi.org/10.1016/j.tree.2014.06.005

Carøe C, Gopalakrishnan S, Vinner L, et al (2018) Single-tube library preparation for degraded DNA. Methods Ecol Evol 9:410–419. https://doi.org/10.1111/2041-210X.12871

Chang CC, Chow CC, Tellier LCAM, et al (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience 4:1–16. https://doi.org/10.1186/s13742-015-0047-8

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat. Rev. Genet. 10:783–796

Clegg SM, Degnan SM, Kikkawa J, et al (2002) Genetic consequences of sequential founder events by an island-colonizing bird. Proc Natl Acad Sci U S A 99:8127–8132. https://doi.org/10.1073/pnas.102583399

Colosimo PF, Hosemann KE, Balabhadra S, et al (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science (80-) 307:1928–1933. https://doi.org/10.1126/science.1107239

Côté SD, Dallas JF, Marshall F, et al (2002) Microsatellite DNA evidence for genetic drift and philopatry in Svalbard reindeer. Mol Ecol 11:1923–1930. https://doi.org/10.1046/j.1365-294X.2002.01582.x

Cullingham CI, Moehrenschlager A (2013) Temporal Analysis of Genetic Structure to Assess Population Dynamics of Reintroduced Swift Foxes. Conserv Biol 27:1389–1398. https://doi.org/10.1111/cobi.12122

Derocher AE, Wiig Ø, Bangjord G (2000) Predation of Svalbard reindeer by polar bears. Polar Biol 23:675–678. https://doi.org/10.1007/s003000000138

Douhard M, Loe LE, Stien A, et al (2016) The influence of weather conditions during gestation on life histories in a wild Arctic ungulate. Proc R Soc B Biol Sci 283:. https://doi.org/10.1098/rspb.2016.1760

Drauch AM, Rhodes OE (2007) Genetic Evaluation of the Lake Sturgeon Reintroduction Program in the Mississippi and Missouri Rivers. North Am J Fish Manag 27:434–442. https://doi.org/10.1577/m06-024.1

Druet T, Gautier M (2017) A model-based approach to characterize individual inbreeding at both global and local genomic scales. Mol Ecol 26:5820–5841. https://doi.org/10.1111/mec.14324

Druet T, Oleński K, Flori L, et al (2020) Genomic Footprints of Recovery in the European Bison. J Hered 194–203. https://doi.org/10.1093/jhered/esaa002

Ejsmond MJ, Radwan J (2011) MHC diversity in bottlenecked populations: A simulation model. Conserv Genet 12:129–137. https://doi.org/10.1007/s10592-009-9998-6

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

Fischer MC, Rellstab C, Leuzinger M, et al (2017) Estimating genomic diversity and population differentiation - an empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. BMC Genomics 18:1–15. https://doi.org/10.1186/s12864-016-3459-7

Fox CW, Reed DH (2011) Inbreeding depression increases with environmental stress: An experimental study and meta-analysis. Evolution (N Y) 65:246–258. https://doi.org/10.1111/j.1558-5646.2010.01108.x

Frankham R (2005a) Stress and adaptation in conservation genetics. J Evol Biol 18:750–755. https://doi.org/10.1111/j.1420-9101.2005.00885.x

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. Biol. Conserv. 143:1919–1927

Frankham R (2005b) Genetics and extinction. Biol Conserv 126:131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frankham R (1995) Effective population size/adult population size ratios in wildlife: A review. Genet Res (Camb) 66:95–107. https://doi.org/10.1017/S0016672308009695

Frankham R, Ballou JD, Ralls K, et al (2017) Genetic Management of Fragmented Animal and Plant Populations. OUP Oxford

Fuentes-Pardo AP, Ruzzante DE (2017) Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. Mol Ecol 26:5369–5406. https://doi.org/10.1111/mec.14264

Garcia-Erill G, Albrechtsen A (2020) Evaluation of Model Fit of Inferred Admixture Proportions. Mol Ecol Resour 1–14. https://doi.org/10.1111/1755-0998.13171

Gjertz I (1995) Reindeer on Prins Karls Forland, Svalbard. Polar Res 14:87–88. https://doi.org/10.3402/polar.v14i1.6654

Grant PR, Grant BR (2006) Evolution of character displacement in Darwin's finches. Science (80-) 313:224–226. https://doi.org/10.1126/science.1128374

Gripenberg U, Huuhtanen S, Wessman M, Nieminen M (1991) A fragile site in the X chromosome of the reindeer (Rangifer tarandus L). Genet Sel Evol 23:S135. https://doi.org/10.1186/1297-9686-23-s1-s135

Grossen C, Biebach I, Angelone-Alasaad S, et al (2018) Population genomics analyses of European ibex species show lower diversity and higher inbreeding in reintroduced populations. Evol Appl 11:123–139. https://doi.org/10.1111/eva.12490

Haanes H, Røed KH, Flagstad, Rosef O (2010) Genetic structure in an expanding cervid population after population reduction. Conserv Genet 11:11–20. https://doi.org/10.1007/s10592-008-9781-0

Hanghøj K, Moltke I, Andersen PA, et al (2019) Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. Gigascience 8:1–9. https://doi.org/10.1093/gigascience/giz034

Hansen BB, Aanes R, Herfindal I, et al (2009) Winter habitat-space use in a large arctic herbivore facing contrasting forage abundance. Polar Biol 32:971–984. https://doi.org/10.1007/s00300-009-0597-2

Hansen BB, Aanes R, Sæther B-E (2010a) Partial seasonal migration in high-arctic Svalbard reindeer (Rangifer tarandus platyrhynchus). Can J Zool 88:1202–1209. https://doi.org/10.1139/Z10-086

Hansen BB, Aanes R, Sæther BE (2010b) Feeding-crater selection by high-arctic reindeer facing ice-blocked pastures. Can J Zool 88:170–177. https://doi.org/10.1139/Z09-130

Hansen BB, Gamelon M, Albon SD, et al (2019a) More frequent extreme climate events stabilize reindeer population dynamics. Nat Commun 10:1–8. https://doi.org/10.1038/s41467-019-09332-5

Hansen BB, Lorentzen JR, Welker JM, et al (2019b) Reindeer turning maritime: Ice-locked tundra triggers changes in dietary niche utilization. Ecosphere 10:. https://doi.org/10.1002/ecs2.2672

Hansen BB, Pedersen ÅØ, Peeters B, et al (2019c) Spatial heterogeneity in climate change effects decouples the long-term dynamics of wild reindeer populations in the high Arctic. Glob Chang Biol 25:3656–3668. https://doi.org/10.1111/gcb.14761

Havrdová A, Douda J, Krak K, et al (2015) Higher genetic diversity in recolonized areas than in refugia of Alnus glutinosa triggered by continent-wide lineage admixture. Mol Ecol 24:4759–4777. https://doi.org/10.1111/mec.13348

Hedrick PW (2000) Inbreeding depression in conservation biology. Annu Rev Ecol Syst 31:139–162. https://doi.org/10.1146/annurev.ecolsys.31.1.139

Hendry AP (2017) Eco-evolutionary dynamics. Princeton university press.

Hicks JF, Rachlow JL, Rhodes OE, et al (2007) Reintroduction and Genetic Structure: Rocky Mountain Elk in Yellowstone and the Western States. J Mammal 88:129–138.

https://doi.org/10.1644/06-mamm-a-051r1.1

Hooper DU, Adair EC, Cardinale BJ, et al (2012) A global synthesis reveals biodiversity loss as a major driver of ecosystem change. Nature 486:105–108. https://doi.org/10.1038/nature11118

Hueffer K, O'Hara TM, Follmann EH (2011) Adaptation of mammalian host-pathogen interactions in a changing arctic environment. Acta Vet Scand 53:1–8. https://doi.org/10.1186/1751-0147-53-17

Huff DD, Miller LM, Vondracek B (2010) Patterns of ancestry and genetic diversity in reintroduced populations of the slimy sculpin: Implications for conservation. Conserv Genet 11:2379–2391. https://doi.org/10.1007/s10592-010-0124-6

Hundertmark KJ, van Daele LJ (2010) Founder effect and bottleneck signatures in an introduced, insular population of elk. Conserv Genet 11:139–147. https://doi.org/10.1007/s10592-009-0013-z

Jamieson IG (2011) Founder Effects, Inbreeding, and Loss of GeneticDiversity in Four Avian Reintroduction Programs. Conserv Biol 25:115–123. https://doi.org/10.1111/j.1523-1739.2010.01574.x

Jamieson IG, Tracy LN, Fletcher D, Armstrong DP (2007) Moderate inbreeding depression in a reintroduced population of North Island robins. Anim Conserv 10:95–102. https://doi.org/10.1111/j.1469-1795.2006.00078.x

Janes JK, Miller JM, Dupuis JR, et al (2017) The K = 2 conundrum. Mol Ecol 26:3594–3602. https://doi.org/10.1111/mec.14187

Johansen BE, Karlsen SR, Tømmervik H (2012) Vegetation mapping of Svalbard utilising Landsat TM/ETM+ data. Polar Rec (Gr Brit) 48:47–63. https://doi.org/10.1017/S0032247411000647

Johnston SE, Huisman J, Ellis PA, Pemberton JM (2017) A high-density linkage map reveals sexual dimorphism in recombination landscapes in red deer (Cervus elaphus). G3 Genes, Genomes, Genet 7:2859–2870. https://doi.org/10.1534/g3.117.044198

Jones BL, Raga TO, Liebert A, et al (2013) Diversity of lactase persistence alleles in ethiopia: Signature of a soft selective sweep. Am J Hum Genet 93:538–544. https://doi.org/10.1016/j.ajhg.2013.07.008

Kaeuffer R, Coltman DW, Chapuis JL, et al (2007) Unexpected heterozygosity in an island mouflon population founded by a single pair of individuals. Proc R Soc B Biol Sci 274:527–533. https://doi.org/10.1098/rspb.2006.3743

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

Kardos M, Qvarnström A, Ellegren H (2017) Inferring individual inbreeding and demographic history from segments of identity by descent in Ficedula flycatcher genome sequences. Genetics 205:1319–1334. https://doi.org/10.1534/genetics.116.198861

Kardos M, Taylor HR, Ellegren H, et al (2016) Genomics advances the study of inbreeding depression in the wild. Evol Appl 9:1205–1218. https://doi.org/10.1111/eva.12414

Kekkonen J, Wikström M, Brommer JE (2012) Heterozygosity in an Isolated Population of a Large Mammal Founded by Four Individuals Is Predicted by an Individual-Based Genetic Model. PLoS One 7:1–8. https://doi.org/10.1371/journal.pone.0043482

Kohler J, Aanes R (2004) Effect of winter snow and ground-icing on a Svalbard reindeer population: Results of a simple snowpack model. Arctic, Antarct Alp Res 36:333–341. https://doi.org/10.1657/1523-0430(2004)036[0333:EOWSAG]2.0.CO;2

Kolbe JJ, Larson A, Losos JB, De Queiroz K (2008) Admixture determines genetic diversity and

population differentiation in the biological invasion of a lizard species. Biol Lett 4:434–437. https://doi.org/10.1098/rsbl.2008.0205

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15:356

Kristensen TN, Barker JSF, Pedersen KS, Loeschcke V (2008) Extreme temperatures increase the deleterious consequences of inbreeding under laboratory and semi-natural conditions. Proc R Soc B Biol Sci 275:2055–2061. https://doi.org/10.1098/rspb.2008.0426

Kutz SJ, Jenkins EJ, Veitch AM, et al (2009) The Arctic as a model for anticipating, preventing, and mitigating climate change impacts on host-parasite interactions. Vet Parasitol 163:217–228. https://doi.org/10.1016/j.vetpar.2009.06.008

Kvie KS, Heggenes J, Anderson DG, et al (2016) Colonizing the high arctic: Mitochondrial DNA reveals common origin of Eurasian archipelagic reindeer (Rangifer tarandus). PLoS One 11:1–15. https://doi.org/10.1371/journal.pone.0165237

Lande R (1980) Genetic Variation and Phenotypic Evolution During Allopatric Speciation. Am Nat 116:463–479. https://doi.org/10.1086/283642

Latch EK, Rhodes OE (2005) The effects of gene flow and population isolation on the genetic structure of reintroduced wild turkey populations: Are genetic signatures of source populations retained? Conserv Genet 6:981–997. https://doi.org/10.1007/s10592-005-9089-2

Lawson DJ, van Dorp L, Falush D (2018) A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nat Commun 9:1–11. https://doi.org/10.1038/s41467-018-05257-7

Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. J Evol Biol 11:495–512. https://doi.org/10.1007/s000360050102

Le Moullec M, Pedersen ÅØ, Stien A, et al (2019) A century of conservation: The ongoing recovery of svalbard reindeer. J Wildl Manage 83:1676–1686. https://doi.org/10.1002/jwmg.21761

Lee AM, Bjørkvoll EM, Hansen BB, et al (2015) An integrated population model for a long-lived ungulate: More efficient data use with Bayesian methods. Oikos 124:806–816. https://doi.org/10.1111/oik.01924

Lee CY, Munyard KA, Gregg K, et al (2011) The influence of MHC and immunoglobulins A and e on host resistance to gastrointestinal nematodes in sheep. J Parasitol Res 2011:. https://doi.org/10.1155/2011/101848

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv Prepr 1–3

Li Z, Lin Z, Ba H, et al (2017) Draft genome of the reindeer (Rangifer tarandus). Gigascience 6:1–5. https://doi.org/10.1093/gigascience/gix102

Loe LE, Hansen BB, Stien A, et al (2016) Behavioral buffering of extreme weather events in a high-Arctic herbivore. Ecosphere 7:1–13. https://doi.org/10.1002/ecs2.1374

Lynch M, Gabriel W (1990) Mutation load and the survival of small populations. 44:1725–1737

Marshall TC, Spalton JA (2000) Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. Anim Conserv 3:241–248. https://doi.org/10.1017/S1367943000000962

McVean G (2009) A genealogical interpretation of principal components analysis. PLoS Genet 5:. https://doi.org/10.1371/journal.pgen.1000686

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

Muckenhuber S, Nilsen F, Korosov A, Sandven S (2016) Sea ice cover in Isfjorden and Hornsund,

Svalbard (2000-2014) from remote sensing data. Cryosphere 10:149–158. https://doi.org/10.5194/tc-10-149-2016

Murphy SM, Cox JJ, Clark JD, et al (2015) Rapid growth and genetic diversity retention in an isolated reintroduced black bear population in the central appalachians. J Wildl Manage 79:807–818. https://doi.org/10.1002/jwmg.886

Nei M, Maruyama T, Chakraborty R (1975) The Bottleneck Effect and Genetic Variability in Populations. Evolution (N Y) 29:1. https://doi.org/10.2307/2407137

Orr HA, Unckless RL (2008) Population extinction and the genetics of adaptation. Am Nat 172:160–169. https://doi.org/10.1086/589460

Pavlova O, Gerland S, Hopp H (2019) Changes in Sea-Ice Extent and Thickness in Kongsfjorden, Svalbard (2003–2016). In: The Ecosystem of Kongsfjorden, Svalbard. Advances in Polar Ecology, vol 2. Springer, Cham, pp 105–136

Peeters B, Le Moullec M, Raeymaekers JAM, et al (2020) Sea ice loss increases genetic isolation in a high Arctic ungulate metapopulation. Glob Chang Biol 26:2028–2041. https://doi.org/10.1111/gcb.14965

Peeters B, Pedersen ÅØ, Loe LE, et al (2019) Spatiotemporal patterns of rain-on-snow and basal ice in high Arctic Svalbard: Detection of a climate-cryosphere regime shift. Environ Res Lett 14:. https://doi.org/10.1088/1748-9326/aaefb3

Pelz HJ, Rost S, Hünerberg M, et al (2005) The genetic basis of resistance to anticoagulants in rodents. Genetics 170:1839–1847. https://doi.org/10.1534/genetics.104.040360

Post E, Forchhammer MC, Bret-Harte MS, et al (2009) Ecological dynamics across the arctic associated with recent climate change. Science (80-) 325:1355–1358. https://doi.org/10.1126/science.1173113

Pruett CL, Winker K (2005) Northwestern song sparrow populations show genetic effects of sequential colonization. Mol Ecol 14:1421–1434. https://doi.org/10.1111/j.1365-294X.2005.02493.x

R Core Team (2019) R: A Language and Environment for Statistical Computing

Radwan J, Biedrzycka A, Babik W (2010) Does reduced MHC diversity decrease viability of vertebrate populations? Biol Conserv 143:537–544. https://doi.org/10.1016/j.biocon.2009.07.026

Radwan J, Kawałko A, Wójcik JM, Babik W (2007) MHC-DRB3 variation in a free-living population of the European bison, Bison bonasus. Mol Ecol 16:531–540. https://doi.org/10.1111/j.1365-294X.2006.03179.x

Reimers E (1983) Mortality in Svalbard reindeer. Ecography (Cop) 6:141–149. https://doi.org/10.1111/j.1600-0587.1983.tb01075.x

Sasmal I, Jenks JA, Waits LP, et al (2013) Genetic diversity in a reintroduced swift fox population. Conserv Genet 14:93–102. https://doi.org/10.1007/s10592-012-0429-8

Schubert M, Ermini L, Sarkissian C Der, et al (2014) Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat Protoc 9:1056–1082. https://doi.org/10.1038/nprot.2014.063

Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. BMC Res Notes 9:1–7. https://doi.org/10.1186/s13104-016-1900-2

Siddle H V., Kreiss A, Eldridge MDB, et al (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci U S A 104:16221–16226. https://doi.org/10.1073/pnas.0704580104

Sinding MHS, Gopalakrishan S, Vieira FG, et al (2018) Population genomics of grey wolves and wolf-like canids in North America. PLoS Genet 14:e1007745. https://doi.org/10.1371/journal.pgen.1007745

Stien A, Loel LE, Mysterud A, et al (2010) Icing events trigger range displacement in a high-arctic ungulate. Ecology 91:915–920. https://doi.org/10.1890/09-0056.1

Taylor RS, Horn RL, Zhang X, et al (2019) The Caribou (Rangifer tarandus) genome. Genes (Basel) 10:4–9. https://doi.org/10.3390/genes10070540

Taylor SS, Jamieson IG (2008) No evidence for loss of genetic variation following sequential translocations in extant populations of a genetically depauperate species. Mol Ecol 17:545–556. https://doi.org/10.1111/j.1365-294X.2007.03591.x

Uboni A, Horstkotte T, Kaarlejärvi E, et al (2016) Long-term trends and role of climate in the population dynamics of eurasian reindeer. PLoS One 11:1–20. https://doi.org/10.1371/journal.pone.0158359

Vors LS, Boyce MS (2009) Global declines of caribou and reindeer. Glob Chang Biol 15:2626–2633. https://doi.org/10.1111/j.1365-2486.2009.01974.x

Vucetich JA, Waite TA (2000) Is one migrant per generation sufficient for the genetic management of fluctuating populations? Anim Conserv 3:261–266. https://doi.org/10.1017/S1367943000000986

Weeks AR, Sgro CM, Young AG, et al (2011) Assessing the benefits and risks of translocations in changing environments: A genetic perspective. Evol Appl 4:709–725. https://doi.org/10.1111/j.1752-4571.2011.00192.x

Weiser EL, Grueber CE, Jamieson IG (2013) Simulating Retention of Rare Alleles in Small Populations to Assess Management Options for Species with Different Life Histories. Conserv Biol 27:335–344. https://doi.org/10.1111/cobi.12011

White DJ, Watts C, Allwood J, et al (2017) Population history and genetic bottlenecks in translocated Cook Strait giant weta, Deinacrida rugosa: recommendations for future conservation management. Conserv Genet 18:411–422. https://doi.org/10.1007/s10592-016-0916-4

Whitlock MC, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. Heredity (Edinb) 84:452–457. https://doi.org/10.1046/j.1365-2540.2000.00693.x

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Willi Y, Griffin P, Van Buskirk J (2013) Drift load in populations of small size and low density. Heredity (Edinb) 110:296–302. https://doi.org/10.1038/hdy.2012.86

Williams BW, Scribner KT (2010) Effects of multiple founder populations on spatial genetic structure of reintroduced American martens. Mol Ecol 19:227–240. https://doi.org/10.1111/j.1365-294X.2009.04455.x

Williams CE, Serfass TL, Cogan R, Rhodes Jr OE (2002) Microsatellite variation in the reintroduced Pennsylvania elk herd. Mol Ecol 11:1299–1310

Williams RN, Rhodes OE, Serfass TL (2000) Assessment of Genetic Variance Among Source and Reintroduced Fisher Populations. J Mammal 81:895–907. https://doi.org/10.1644/1545-1542(2000)081<0895:aogvas>2.3.co;2

Willis K, Willis RE (2010) How many founders, how large a population? Zoo Biol 29:638–646. https://doi.org/10.1002/zoo.20310

Wisely SM, Santymire RM, Livieri TM, et al (2008) Genotypic and phenotypic consequences of reintroduction history in the black-footed ferret (Mustela nigripes). Conserv Genet 9:389–399. https://doi.org/10.1007/s10592-007-9351-x

Wright DJ, Spurgin LG, Collar NJ, et al (2014) The impact of translocations on neutral and functional

genetic diversity within and among populations of the Seychelles warbler. Mol Ecol 23:2165–2177. https://doi.org/10.1111/mec.12740

Wright S (1931) Evolution in mendelian populations. Bull Math Biol 52:241–295. https://doi.org/10.1007/BF02459575

Supporting Information

SI 1. *DNA extraction appendix*

Peeters et al. (2020). A Qiagen (Hilden, Germany) DNeasy Blood & Tissue extraction kit was used to extract DNA from ~20 mg of ear tissue from each of the eight Daudmannsøyra samples, along with an extraction blank. The fur and outer skin layer were removed, and each tissue sample was diced with a scalpel and then placed into a 1.5-ml Eppendorf tube with 180 µL buffer ATL. 20 µL of proteinase K was added to each sample tube and vortexed. Samples were then incubated at 56°C in a Thermoshaker set at 700 RPM for 5.5 hours. Each sample was vortexed after incubation, then 200 µL buffer AL and 200 µL 96% ethanol were added to each sample and immediately vortexed. The mixture was then pipetted into a DNeasy mini spin column placed inside a 2 ml collection tube and centrifuged at 8000 RPM for one minute. Mini spin columns were then placed in a new 2 ml collection tube, 500 µL buffer AW1 added, and centrifuged at 8000 RPM for one minute. Mini spin columns were again placed in a new 2 ml collection tube, 500 µL buffer AW2 added, and centrifuged for three minutes at 14000 RPM. Finally, the mini spin columns were then placed in 1.5 ml centrifuge tubes and 200 µL buffer AE was pipetted onto the DNeasy membrane, and incubated for one minute. The samples were then centrifuged at 8000 RPM for one minute to elute. The DNA concentration was quantified for all using a Thermo Fisher Scientific Qubit 2.0 fluorometer (Indiana, USA).

SI 2. *Library build protocol*

Genomic library building was based on the BEST v 1.1 (Blunt End Single Tube) 96 well-plate format protocol (Vanessa Bieker, personal communication) based on the method presented in Carøe et al., (2018). Samples were pipetted into a 96 well plate and diluted with EB buffer if necessary, so all wells contain 60 µL volume containing 500 – 5000 ng DNA. A Covaris ME220 focused ultrasonicator was used for DNA fragmentation with a target fragment length of 400 bp.

Blunt end repair

8 µL of end-repair master mix (0.4 µL T4 DNA polymerase at 0.03 U/µL concentration, 1 µL T4 polynucleotide kinase at 0.25 U/µL, 0.4 µL dNTPs at 0.25 mM each, 4 µL T4 NDA ligase buffer (NEB) at 1x concentration, and 2.2 µL reaction enhancer) was added to 32 µL of extract from each sample in a 96 well plate and incubated at 20°C for 30 minutes followed by 65°C for 30 minutes and cooled to 4°C.

Adapter ligation

2 µL adapter solution (20 µM) was added and samples were vortexed. 8 µL ligase master mix (1 µL T4 DNA ligase buffer (1x concentration), 6 µL PEG-4000 (6.25%), 1 µL T4 DNA ligase (8 U/µL)). Samples were incubated at 20 °C for 30 minutes followed by 65 °C for 10 minutes and cooled to 4 °C.

Adapter fill-in

10 µL of fill-in master mix (2 µL Isothermal amp. Buffer (0.33x concentration), 0.8 µL dNTPs (0.33mM), 1.6 µL Bst 2.0 Warmstart pol. (0.21 U/µL), 5.6 µL H_2O) was then added to each sample which were incubated at 65 °C for 15 minutes followed by 80 °C for 15 minutes and then cooled to 4 °C.

Library purification

100 µL SPRI beads were added to each sample, incubated for five minutes at room temperature, then the plate was placed on a magnetic rack and supernatant was removed once bead pellets had formed. Each sample was washed twice with 200 µL 80% ethanol. The plate was removed from the magnetic rack and 33 µL EBT buffer was used to elute the DNA. The plate was sealed with aluminum foil and incubated at 37 °C for 10 minutes, and then returned to the magnetic rack where the DNA containing supernatant was transferred to a new well plate after bead pellets had formed.

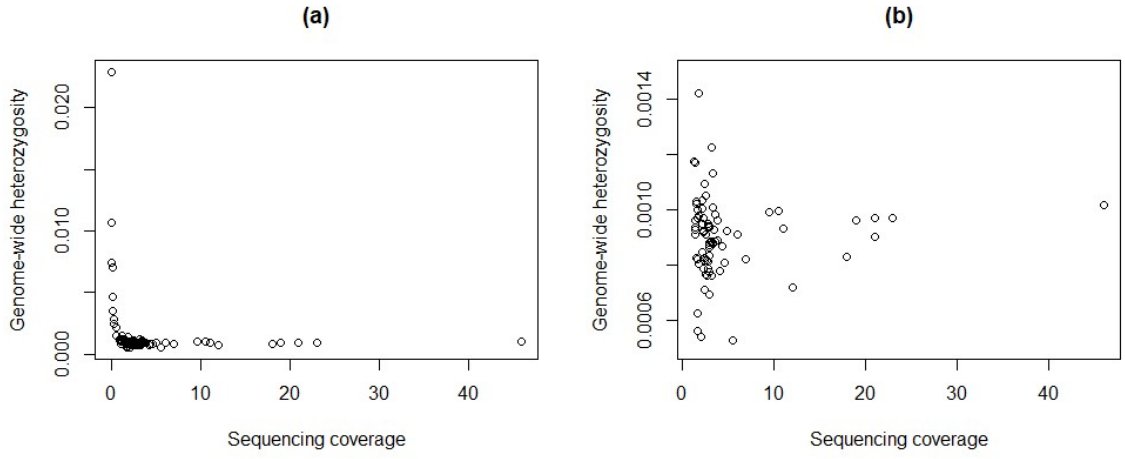
Indexing PCR and sequencing

Each sample was indexed with a unique combination of F and R index primers. For each sample, 10 µL of library template and a unique combination of F and R index primers (2 µL of each at 0.2 µM) was added to 96 µL index PCR master mix (0.8 µL dNTPs (0.2 nM), 1 µL AmpliTaq Gold polymerase (0.5 U/µL), 10 µL AmpliTaq Gold buffer (1x conc), 10 µL $MgCl_2$ (2.5 mM), 2 µL BSA (0.4 mg/ml), 62.2 µL H_2O), resulting in a total volume of 100µL for each library. All samples were given nine cycles of PCR during indexing. PCR consisted of holding libraries at 95 °C for 10 minutes, followed by nine cycles of 95 °C for 30 seconds, 60 °C for one minute, and 72 °C for 45 seconds. After the final cycle, libraries were held at 72 °C for

five minutes and then cooled to 4 °C. 100 µL of SPRI beads (Rohland & Reich, 2012) were mixed with the 100 µL PCR product for each library, which were put onto magnetic racks after incubating at room temperature for five minutes. Libraries were then washed twice with 200 µL 80% ethanol, then removed from the magnetic rack. 33 µL EB buffer was added, libraries were vortexed and incubated at 37 °C for 10 minutes and placed back on the magnetic rack. Once beads had reformed, the DNA-containing supernatant was collected and transferred to a new plate. Samples were then run on an Agilent TapeStation 4200 to estimate molarity of samples for pooling. A pooling design was devised based on these morality values, resulting in samples being combined into eight equimolar pools for sequencing. Samples were sequenced on an Illumina HiSeq 4000 (one 150-bp paired-end flowcell) at the NTNU Genomics Core Facility.

SI 3. *Effect of coverage on genome-wide heterozygosity estimates*

A strong negative non-linear relationship was found between sequencing coverage and heterozygosity estimates, with very low coverage samples having strongly upward biased heterozygosity scores (Supp. Fig. A. below). After filtering samples to 1.3x minimum coverage, no significant relationship was observed (linear model heterozygosity~coverage, slope =1.542e-06, P = 0.53).



Supp. Fig. A. Heterozygosity – sequencing coverage relationship. a) including all samples; b) only including samples with sequencing coverage > 1.3x.

SI 4. *ROH simulation description*

The simulation was designed as follows:

24 unique chromosomes were simulated by creating a vector with 10,000 repeats of a single unique letter for each (A – X) to represent markers spaced 100 Kbp apart. Chromosomes were paired together to form a 2 x 10,000 data frame representing individual animals. These individual data frames are placed in a list of founder individuals. The simulation was set to loop for 40 cycles, with each cycle representing one year. Founders were assumed to be three years old in the first year. Individuals that are of breeding age, set to 3 – 10 years (cycles) are extracted from the list and into a "breeding pool" list each year. When the number of breeding-age individuals exceeded 140, a random sample of 140 of the available individuals of breeding age was taken to maintain a stable population size. Within the breeding pool each year, individuals were randomly paired together by randomising their order in the list and sequentially moving through the randomised list pairing together individuals.

To simulate recombination and Mendelian inheritance, each year within each breeding individual, every site along the chromosome was given a 1/1000 probability of recombination (representing a recombination rate of 1 cM/Mbp) generating designated recombination sites. Recombination to produce "gametes" involved replacing the portion of a chromosome vector above each recombination site with the characters from the opposite homolog within the individual, in sequential order from the first recombination site to the last. After recombination, one of these "gamete" vectors was extracted from both individuals in a breeding pair and combined into a new data frame to form a new individual. The new individuals produced each year were then put into a sub-list (corresponding to the year/cycle it was created in) nested within a combined list of all individuals. The breeding pool selection moves along a sliding window of these sub-lists each cycle so that individuals were only selected for breeding when they were three years old, and so that individuals over 10 years old were no longer selected (effectively removed from breeding population). Age-based survival probability was implemented by randomly removing a proportion of individuals in each age class each year (not implemented for founders).

All homozygous sites (sites containing the same character in both chromosomes of an individual) were therefore identical by descent (IBD). To identify RoH, tracts of matching characters were extracted from all individuals produced in year 40 (70 individuals) to simulate

offspring born in 2018 at the first reintroduction site in Brøggerhalvøya. Inbreeding coefficients were estimated as the proportion of the genome that was homozygous.

This script therefore allowed the frequency and length runs of homozygosity and overall inbreeding coefficients/loss of heterozygosity attributable to the reintroduction bottleneck and subsequent inbreeding in the simulated population to therefore be quantified.

Supplementary Figures

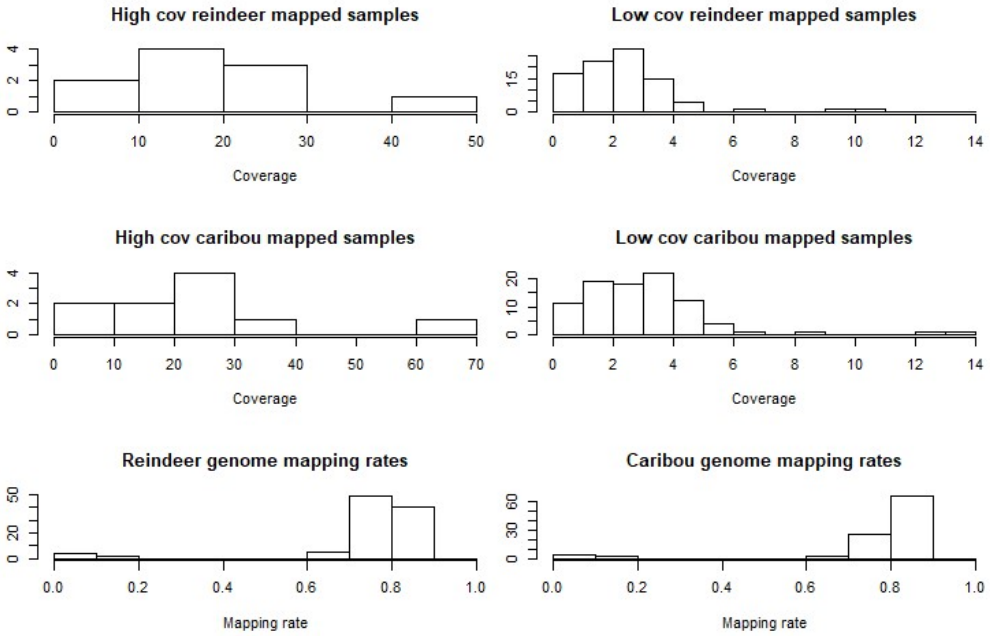


Figure S1. Coverage frequency distribution of sequenced reindeer samples mapped to reindeer and caribou genomes from two sequencing runs (high coverage and low coverage).

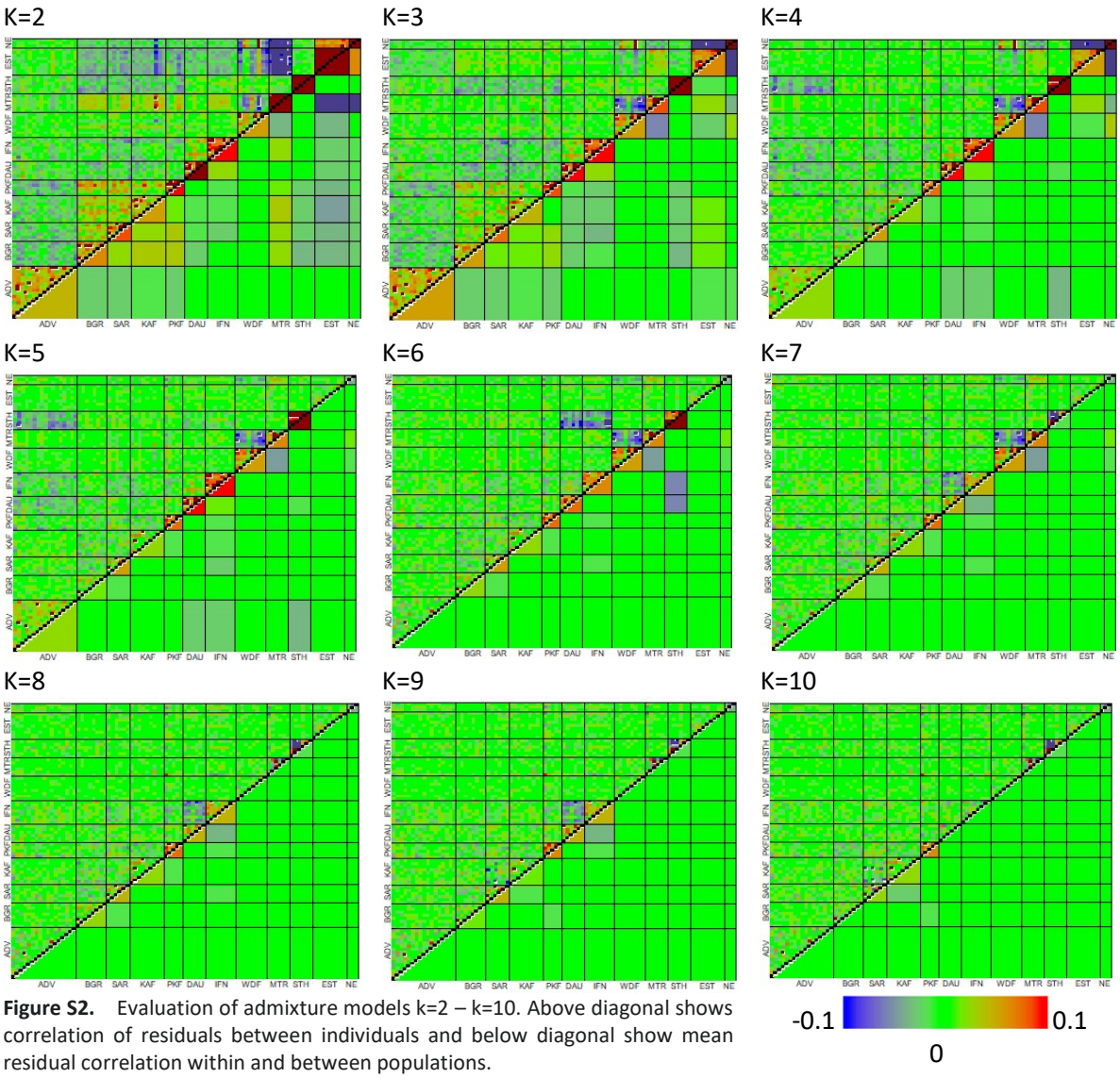


Figure S2. Evaluation of admixture models k=2 – k=10. Above diagonal shows correlation of residuals between individuals and below diagonal show mean residual correlation within and between populations.

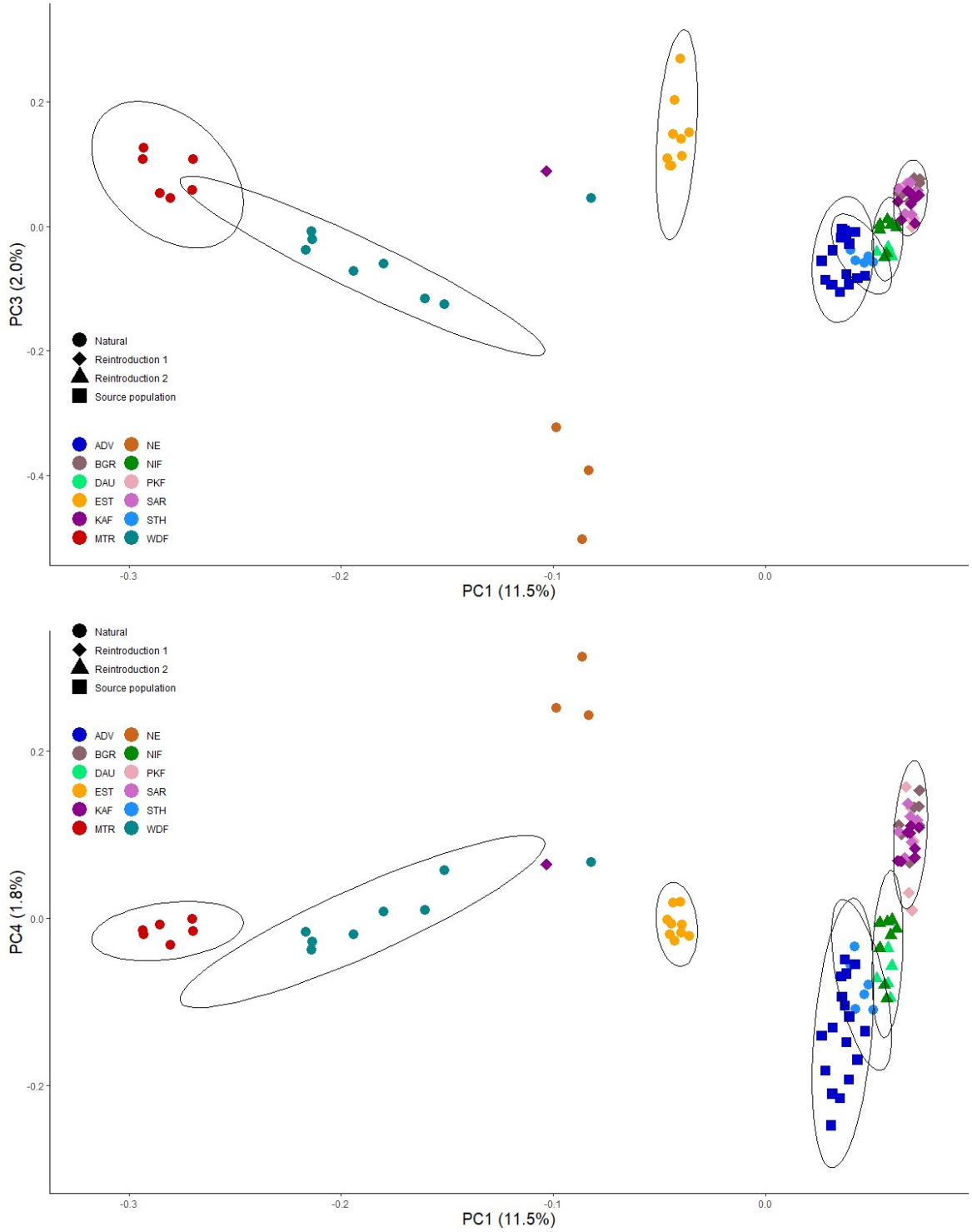


Figure S3. PC1 and (a) PC3, (b) PC4 from principal component analysis of the Svalbard reindeer meta-population. Variation explained by each PC axis indicated on axis labels. on Ellipses represent normal 95% confidence intervals of population means (Reintro 1 (BGR, SAR, KAF, PKF) and Reintro 2 (DAU, NIF) combined, and all other sub-populations individually). Two admixed individuals were not included in population ellipse estimations, and the North East Land population had too few samples for an ellipse estimate. ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karls Forland, DAU = Daudmannsøyra, NIF = North Isfjorden, WDF = Wijdefjorden, MTR = Mitrahalvøya, STH = Southern Spitsbergen, EST = Eastern Svalbard, NE = North East Land.

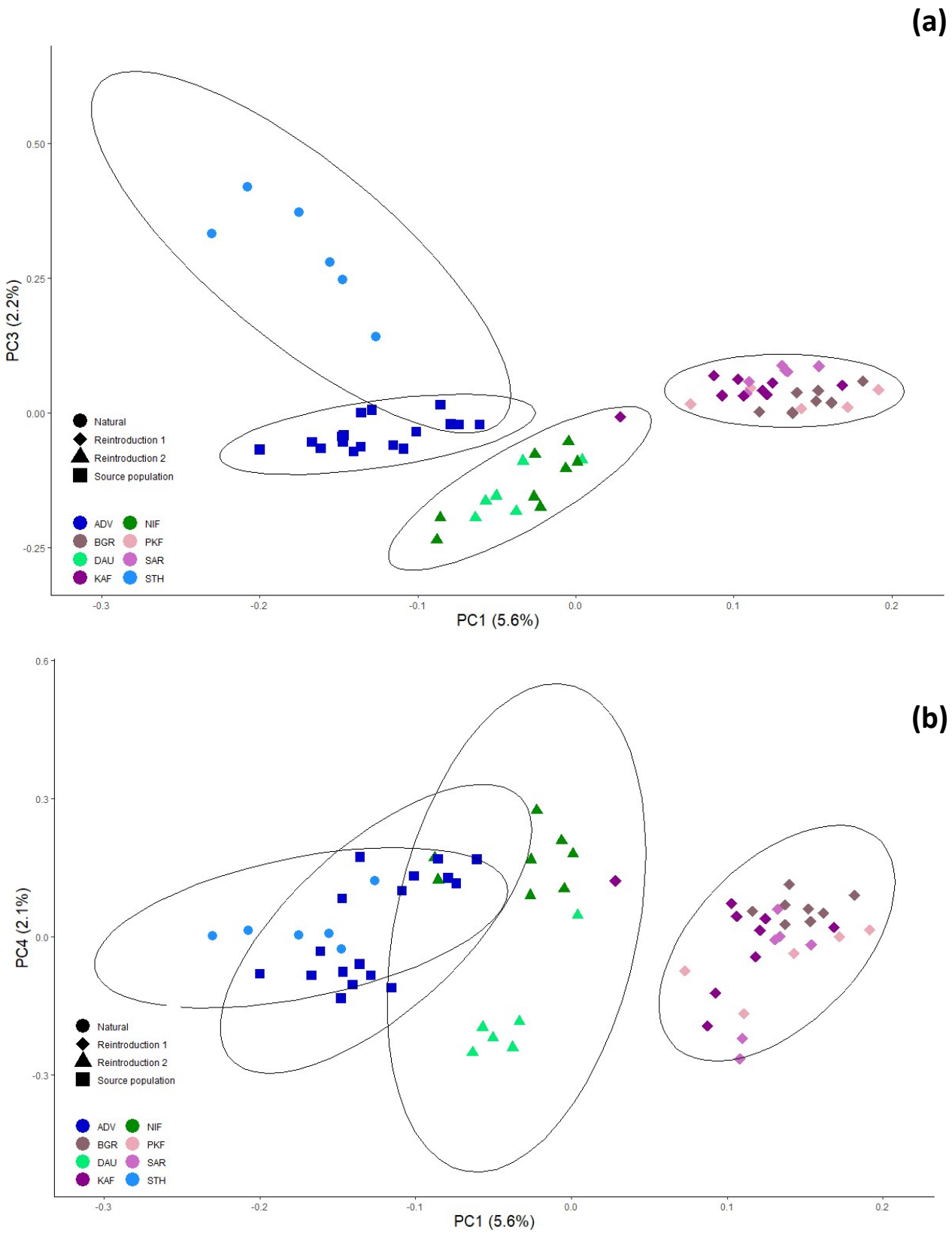


Figure S4. PC1 and (a) PC3, (b) PC4 from principal component analysis including only the ADV population, reintroduced populations 1 and 2, and STH population. Variation explained by each PC axis indicated on axis labels. Ellipses represent normal 95% confidence intervals of population means (Reintro 1 (BGR, SAR, KAF, PKF) and Reintro 2 (DAU, NIF) combined, and ADV and STH individually). ADV = Adventdalen, BGR = Brøggerhalvøya, SAR = Sarsøyra, KAF = Kaffiøyra, PKF = Prins Karls Forland, DAU = Daudmannsøyra, STH = Southern Spitsbergen.

NTNU

Norwegian University of
Science and Technology