Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Teacher Education

Silje Langsrud

Genome-wide levels of variation in space and time in Scandinavian subpopulations of the arctic fox (*Vulpes lagopus*)

Master's thesis in Natural Science with Teacher Education
Supervisor: Henrik Jensen, Ingerid J.H. Arnesen, Øystein Flagstad, Nina E. Eide
January 2020

NTNU
Norwegian University of Science and Technology

Silje Langsrud

Genome-wide levels of variation in space and time in Scandinavian subpopulations of the arctic fox (*Vulpes lagopus*)

Master's thesis in Natural Science with Teacher Education
Supervisor: Henrik Jensen, Ingerid J.H. Arnesen, Øystein Flagstad, Nina E. Eide
January 2020

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Teacher Education

NTNU
Norwegian University of
Science and Technology

Abstract

Use of genome-scale genetic data opens up new possibilities for looking into important questions in conservation biology. For example, conservation genetics and genomics has proven to be especially important in preservation and monitoring of threatened species. The Scandinavian arctic fox (*Vulpes lagopus*) is no exception. The population experienced a rapid decline in population size in the beginning of the 20th century. Unfortunately, the population was not able to recover, following low food availability (i.e. unstable rodent cycles) and interspecific competition and predation from the red fox (*Vulpes Vulpes*). In addition, the small size and fragmented structure of the Scandinavian arctic fox population in itself contributed to making the subpopulations having high risk of experiencing genetic drift and inbreeding depression. Over the last decades, several conservation and management measures have been implemented in both Sweden and Norway. The purpose was primarily to improve conditions and increase the number of arctic foxes in Scandinavia, and second, to restore subpopulations where the arctic fox had gone extinct and support existing subpopulations.

In this study the aim is to compare high-density genotype data for individuals in 6 subpopulations of the arctic fox in Scandinavia, to investigate the genetic variation within and between subpopulations, all over and regionally in the arctic fox genome. A custom Affymetrix Axiom 702K SNP-array for arctic fox and red fox was used to genotype the individuals. The results of this study suggest that the genetic variation (i.e. heterozygosity) within each subpopulation, for the most part, has increased over the study period. Accordingly, there has been a change in the genetic composition of most subpopulations during the study period, as shown by the observed levels of genetic differentiation across the whole and/or parts of the genome between sampling periods for the subpopulations. The results also indicate that the genome-wide and/or regional (within the genome) genetic differentiation between subpopulations has generally declined through the study period. Further studies with longer time periods using approaches similar to the ones in this study may be able to reveal how genetic drift, migration (gene flow) and selection interact to shape the variation within regions and across the whole genome, and thus make it possible to draw conclusions about genetic consequences of on-going conservation management actions. This study shows that high-density genome-wide genotype data, in combination with an arctic fox reference genome, open up new possibilities within conservation genomics and related questions for the Scandinavian arctic fox.

Sammendrag

Bruk av genom-skalert data åpner opp for nye muligheter angående genomiske spørsmål om bevaring. For eksempel har konserverings genetikk og genomikk vist seg å være viktig innen bevaring og overvåkning av truede arter. Den Skandinaviske fjellreven (*Vulpes lagopus*) er intet unntak. Populasjonen opplevde en rask nedgang i populasjonsstørrelsen på begynnelsen av 1900-tallet. Dessverre klarte ikke populasjonen å komme seg, som følge av lav mattilgjengelighet (dvs. ustabile gnagersykluser) og interspesifikk konkurranse og predasjon fra rødreven (*Vulpes vulpes*). I tillegg vil en liten populasjonsstørrelse og en fragmentert struktur hos den Skandinaviske fjellreven i seg selv bidra til at subpopulasjoner finner seg i en høy risiko for å oppleve genetisk drift og innavlsdepresjon. Det har derfor i løpet av de siste tiårene blitt iverksatt flere bevaring og forvaltningstiltak i Sverige og Norge. Formålet med iverksettingen var først og fremst å forbedre forholdene og øke antallet fjellrev i Skandinavia, dernest å gjenopprette subpopulasjoner hvor fjellreven var utryddet og støtte eksisterende subpopulasjoner.

Formålet i dette studiet er å sammenligne genotype data med høy tetthet for individer i 6 fjellrev subpopulasjoner i Skandinavia, for å undersøke den genetiske variasjonen innen og mellom subpopulasjonen, over hele og regionalt i fjellrev genomet. En Affymetrix Axiom 702K SNP-array tilpasset fjellrev og rødrev ble brukt til å genotype individene.

Studiets resultater antyder at den genetiske variasjonen (dvs. heterozygositet) innen hver subpopulasjon har for det meste økt over studiets omfang. Tilsvarende har det vært en endring i den genetiske sammensetningen for de fleste subpopulasjonene gjennom studieperioden, som vist av observerte nivåer av genetisk differensiering på tvers av hele og/eller deler av genomet mellom sampling perioder for subpopulasjonene. Resultatene indikerer også at den genetiske differensieringen mellom subpopulasjoner (over eller innen genomet) har generelt avtatt gjennom studieperioden. Videre studier med lengre tidsperioder, kan ved å bruke liknende tilnærminger som dette studiet, avdekke hvordan genetisk drift, migrasjon (gen flyt) og seleksjon samhandler i å forme variasjon innen regioner og på tvers av hele genomet, og dermed kunne trekke konklusjoner om genetiske konsekvenser av pågående bevaringsaksjoner. Dette studiet viser at genotype data (genom-bredt) med stor tetthet, i kombinasjon med et fjellrev-referanse-genom, åpner opp for nye muligheter innen konserverings genomikk og relaterte spørsmål for den Skandinaviske fjellreven.

Preface

This master thesis was written at the Norwegian University of Science and Technology (NTNU) and is a part of the arctic fox project at the Norwegian Institute of Natural Research (NINA). The thesis is one of the last things at my teacher education, and I´m so grateful for being allowed to participate in the research on the arctic fox. The whole process of writing this thesis (30 credits) has truly been educational.

I would like to express gratitude to my supervisors Henrik Jensen, Ingerid J.H. Arnesen, Øystein Flagstad and Nina E. Eide. I especially would like to thank my main supervisor Henrik Jensen for great guidance and tremendous support throughout the process. I would also like to thank Ingerid J.H. Arnesen for valuable information and guidance, including teaching me how to use Plink. I´m also hugely grateful to Øystein Flagstand for providing me with an excellent datafile over the different subpopulations, in addition to great guidance and insights. A huge thanks to Nina E. Eide for providing me with valuable insights about management actions implemented for the arctic fox. An extra thanks to you all, for giving me much appreciated inputs and feedbacks. I´m also thankful to Dilan Saatoglu and Sarah Lundregran for their advice on statistical challenges in R.

Finally, I would like to take the opportunity to thank my friends and family (including the best dog in the whole world), for all the love and support.

Silje Langsrud,

26th January 2020

Table of Contents

List of Figures

List of Tables

1. Introduction

In recent decades there has been an increased attention to the development of tools in the conservation of threatened species. Human activities with hunting, climate change, habitat fragmentation, pollution and invasive species constitute major threats and increase the extinction risk of species (Lande 1998; Isaac, 2004). How species respond to these threats and their vulnerability can vary both within the same species (i.e. heterogeneity in species response) and between different species (Isaac, 2004). Regardless, the most common reason why a species fails to recover probably originates from the struggles of dealing with a small population size (Larsson et al., 2019). Random genetic drift may have large consequences for small populations, especially when they are isolated (Lande 1988; Reed and Frankham, 2003). Through e.g. reduced immigration and emigration, fewer available potential mates and/or possibly social breakdowns, the populations' natural dynamics may collapse. McMahon, Teeling and Höglund (2014) outlines different scenarios regarding informativeness of genetic data, explains which forces that are most important under different circumstances and discuss potential actions that can be used if monitoring and preservation is necessary. One of the scenarios addresses the genetic problems that encounters small (and fragmentated) populations (McMahon, Teeling and Höglund, 2014). Populations that already have low genetic variation, are likely to experience rapid loss of genetic diversity due to genetic drift and in some cases inbreeding (McMahon, Teeling and Höglund, 2014). Because of the low genetic variation, selection has a reduced efficiency in small (isolated) populations, especially if they are suffering from inbreeding (Supple and Shapiro, 2018). In addition, if the population lacks necessary genetic variation for natural selection to act on (Supple and Shapiro, 2018) and if the loss of genetic variation continues, it may result in a lowered adaptability to environmental changes (Lande, 1988), increasing the likelihood of population extinction (McMahon, Teeling and Höglund, 2014).

There are two particular reasons for why it is important to preserve genetic diversity within species (Reed and Frankham, 2003). First, genetic diversity is necessary for adaptive evolution, and second, heterozygosity and the population`s fitness are anticipated to be correlated (Reed and Frankham, 2003). To understand how the management and conservation should be undertaken, there is need for more genetic data and genomic techniques. By scaling up to genomic data, the number of genetic markers increase (e.g. SNPs) and thereby, the precision of the estimated parameters should improve in accuracy (Shafer et al., 2015; Supple and Shapiro, 2018). One method of conservation management actions regarding genetic rescue is

translocations of individuals. This has been done successfully for e.g. the Florida panthers (*Puma concolor coryi*), even though there still are some challenges to overcome (Johnson et al., 2010). Examples of challenges the Florida panthers continued meeting are persistent inbreeding, infectious agents, habitat loss etc. (Johnson et al., 2010).

One of the most endangered mammals in Scandinavia is the arctic fox (*Vulpes lagopus*), listed as critically endangered in both the Norwegian and Swedish Red List for species (Artsdatabanken, 2015; ArtDatabanken, 2015). The species experienced a rapid reduction in the number of individuals in the beginning of the 20th century in Scandinavia because of a heavy hunting pressure (Lande, 1988; Hersteinsson et al., 1989). In 1930, the arctic fox was protected by legislation in Norway in the hope that the population would recover, but there was no sign that the species was able to do this naturally (reviewed in these action plans: Eide et al., 2017; Direktoratet for naturforvaltning, 2003). The same trend was observed in Sweden after the arctic fox was protected here by legislation in 1928 (Direktoratet for naturforvaltning, 2003). While hunting led to be a rapid decline in the arctic fox population, a combination of other factors may have influenced the unsuccessful recovery and may still threaten the species. The red fox (*Vulpes vulpes*) has a negative impact on the arctic fox through interspecific competition and predation (Elmhagen et al., 2017). Another main threat is low food availability caused by collapse in small rodent populations (Ims, Henden and Killengreen, 2008). Furthermore, the small population size of the arctic fox may in itself increase their vulnerability and lead to further decline in populations (Loison, Strand and Linnell, 2001).

Actions have been initiated since 1998 to support the arctic fox in Scandinavia (Eide et al., 2017). Life-Nature (SEFALO), the first project involving supplementary feeding and red fox control, were implemented in Finland and Sweden (Eide et al., 2017). These conservation measures were not implemented in Norway before 2004 (Eide et al., 2017). In 2010, a joint project to help the arctic fox in Scandinavia, by improving conditions and increase number of foxes within the regions of the project, was established between Norway and Sweden (Ericson, 2014a). The "Felles Fjellrev" project was in progress until 2014 and showed good results (Ericson, 2014b). Thereby in 2015, a declaration of intent was signed on the management of the Scandinavian arctic fox population (Eide et al., 2017). Following, a new project ("Felles Fjellrev II) and an action plan ("Handlingsplan for fjellrev (*Vulpes lagopus*) 2017 – 2021 Norge – Sverige) was compiled (Felles Fjellrev II, no date; Eide et al., 2017).

In Norway, the current "Arctic Fox Captive Breeding Program", was initiated from 2005 and onwards, to supplement arctic fox feeding and red fox culling, and with the goal to reestablish

extinct subpopulations and strengthen extant subpopulations (Landa et al., 2017; Eide et al., 2017.). The long-term vision behind the initiation of these actions is to reach a viable Scandinavian arctic fox population without need for further conservation and management measures (Eide et al., 2017). Through the breeding programme, arctic foxes have been released in several mountain areas (Ulvund et al., 2018). These include among others Hardangervidda, Snøhetta, and Saltfjellet (Eide et al., 2017; Ulvund et al., 2018).

The breeding program for arctic foxes was started with animals brought in from the nature, but in recent years, breeding animals have also been recruited from animals born in captivity (Landa et al., 2018). Because of the potential negative genetic effects of captivity there exists protocols for replacement of breeding animals, as limits to the number of generations that can be recruited from the breeding program's own breeding individuals (Landa et al., 2018). Additionally, the objective of the breeding program is that the genetic variation should represent the one that is still existing in Scandinavia (Landa et al., 2018). The breeding animals are therefore mixed in pairs with individuals that are not related and brought in from as many remaining Arctic fox subpopulations as possible (Landa et al., 2018). The breeding program may work as a buffer against loss of genetic variation, which is important for long-term persistence of the arctic fox. After the release of foxes from the breeding program, several immigration events have been observed (Dalén et al., 2006; Eide et al., 2017), and successful immigration may be an important contributor to the persistent of the arctic fox populations (Loison, Strand and Linnell, 2001).

Since actions started some subpopulations have recovered and many of the subpopulations are constantly increasing in size (Angerbjörn et al., 2013; Ulvund and Wallén, 2018). The genetic consequences of the initiatives (i.e. actions) are to a lesser extent mapped, but analysis of neutral genetic markers shows a general increase in genetic variation and lower differentiation (Hemphill et al. in review). The development of novel genomic resources such as an Arctic fox reference genome (Von Seth et al., in prep) and custom high-density single nucleotide (SNP) genotyping-arrays, representing functional genes will open up new possibilities for looking into conservation genomic assessments in the arctic fox (Hagen et al., in prep.). These new tools allow us to address questions not only related to patterns of neutral genetic variation (Dalén et al., 2006; Hemphill et al. in review and Hasselgren et al., 2018)., but also to identify important functional genes and their distribution across the Scandinavian arctic fox population.

In this project I will investigate the genetic variation within and between populations of the arctic fox in Scandinavia. I will focus on the genetic variation at the genome-wide level, and at different parts of the genome (scaffolds). Different subpopulations have experienced different

degrees of impact from the captive breeding program, and a comparison of genomic variation in space and time could point out the most important genomic impacts of the breeding programme. In line with the results in Hemphill et al. (in review), I expect to see a decrease in genetic differentiation between the subpopulations.

2. Materials and Methods

2.1 Study species and management

The arctic fox has a circumpolar distribution, where it inhabits arctic and alpine tundra (Angerbörn and Tannerfeldt, 2014). Globally, the species shows a stable population trend and is considered least concern in the IUCN Red List of Threatened Species (Angerbörn and Tannerfeldt, 2014). However, as mentioned the situation for the Scandinavian population is quite different, with small and fragmentated populations (Artsdatabanken, 2015). Since 2005, several conservation and management measures have been initiated, such as "the Arctic Fox Captive Breeding Program", supplementary feeding and red fox control (Landa et al., 2017; Eide et al., 2017). The breeding program was originally based on the genetic variation in wild-caught arctic fox juveniles that was left in the Scandinavian population (Landa et al., 2018; Landa et al., 2017). The first release of arctic foxes from the breeding program was on Saltfjellet (n = 2) in 2006 (Landa et al., 2017; Ulvund et al., 2018). Since then, there has been registered a total of 261 captive-bred fox releases in the subpopulations studied herein. They are distributed at Hardangervidda (n = 123), Snøhetta (n = 75) and Saltfjellet (n = 63) (Landa et al., 2017; Ulvund et al., 2018). In addition, the breeding program is to some extent an important source in the recovery of two of the Swedish subpopulations (Helags and Vindelfjällen) in this study. Though not through release of arctic foxes from the breeding program, but from immigration of arctic foxes that have been released from the program in other subpopulations in geographic proximity to these subpopulations (Eide et al., 2017; Landa et al., 2017; Ulvund et al., 2018). The genetic variation in the breeding station was therefore an important reference in this project. Since most of the foxes that was born in the station has been genotyped, we expect all the founder lines to be well represented.

2.2 Data collection and sampling

Through the surveillance work on the arctic fox, excrement and hair samples are collected from dens both in Norway and Sweden (Angelbjörn et al, 2008; Landa et al., 2017; Ulvund et al., 2018). After being genetically analyzed, the samples may be useful by providing a unique DNA-profile for each individual in the given area/subpopulation (Ulvund et al., 2018). Wild-born pups caught in for breeding and pups born at the breeding station are both DNA-sampled (tissue), in addition to ear tagging and a microchip inserted into their neck skin (Landa et al., 2019). The use of a microchip is useful when identifying which individuals that where present

in the population in a given year. Furthermore, DNA-samples can be genotyped at different kinds of markers and be used to track individuals in space and time, to e.g. obtain information on individual survival (Landa et al., 2018).

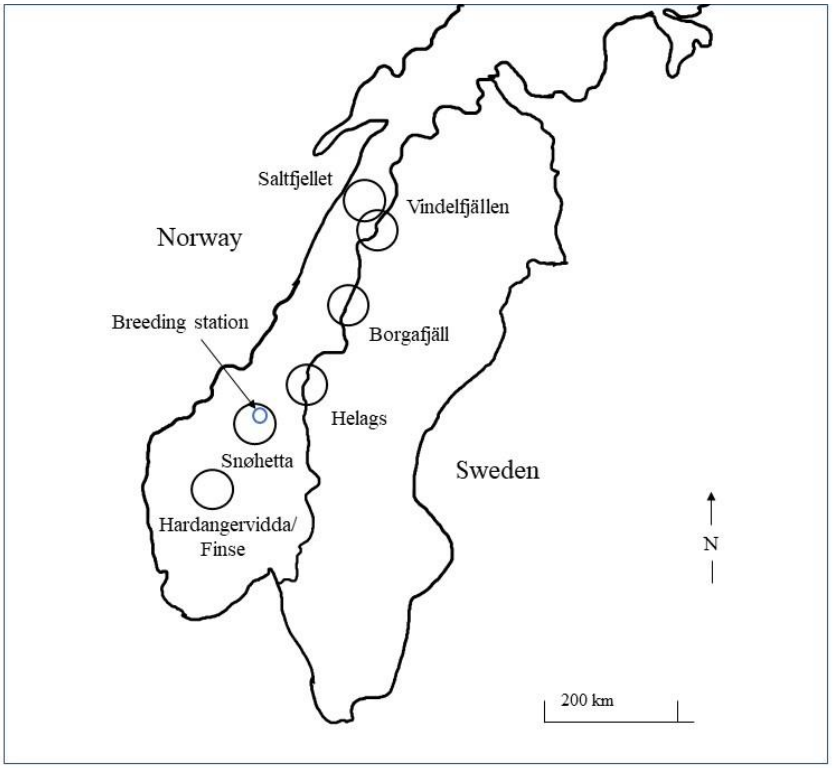


Figure 1. Map showing the 6 subpopulations investigated in this study: Hardangervidda, Snøhetta and Saltfjellet in Norway, and Helags, Borgafjäll and Vindelfjällen in Sweeden. The breeding station is indicated by a small, blue circle.

Data used in this project were obtained from tissue samples collected from pups either in the breeding station or during annual surveys of active dens in Norway and Sweden. Altogether, 703 arctic fox individuals were genotyped, but the final data set, from which I conducted the statistical analysis, was restricted to include only foxes identified as adults from the surveys described above (Table 1). This restriction was important in order to give a true representation of alive individuals in the study populations through time. Altogether, 253 individuals met these criteria and were included in the final analyses (Table 1).

Table 1. Overview of subpopulations and sample sizes. Early time period represents the early phase of the subpopulation, and late time period represents the late phase, last registered and analyzed data after the measures were implemented. Number of generations is the "no of generations" which has passed between early and late sampling period. It was calculated by taking (mid-year for last time period – mid-year for early time period)/4) for each subpopulation. The generation time for the arctic fox is approximately 4 years according to Loison et al., 2001.

Subpopulation	Early time period	Late time period	Number of generations	Sample size	
				Early	Late
Hardangervidda	2010-2011	2016-2018	1.625	19	52
Snøhetta	2008-2009	2016-2017	1.625	29	39
Saltfjellet	2009/2011	2014-2016	1.250	7	8
Borgafjäll	2010	2014	1.000	7	24
Helags	2008-2010(5)*	2013-2015	1.250	18	20
Vindelfjällen	2010-2011/2013	2015	0.917	11	19

2.3 Genotyping, quality control and data selection

To genotype the individuals, a custom Affymetrix Axiom 702K SNP-array for arctic fox and red fox was used. The design of the custom 702K Affymetrix Axiom array for the arctic fox is described in Hagen et al. (in prep.) Under development of the array, 500 000 high-quality SNPs were chosen so that they were positioned evenly along 4048 scaffolds in the draft reference genome for the arctic fox (von Seth et al. in prep.). Genotyping of 731 arctic fox individuals was carried out at the Center of Integrative Genetics (CIGENE). Only poly high resolution SNPs (366 149 SNPs) were used in further analysis. Due to sample quality, 28 individuals failed. Thereby, the genomic dataset consisted of 703 individuals for a total of 6 subpopulations before further quality control was conducted at the Norwegian Institute for Nature Research (NINA). PLINK 1.90 beta (Purcell et al., 2007), a whole genome association analysis toolset, was used to quality check and filter SNPs. SNPs with a low minor allele frequency (MAF < 0.01) were discarded. No individuals were removed due to low genotype rate (MIND < 0.05). Pedigree information for 1632 SNP indicated that (> 10) Mendelian errors were removed. This resulted in genotype data for a total number of 359 218 SNPs.

After receiving the genomic dataset, further quality control filtering, with other thresholds, was conducted by use of Plink version 1.90 beta (https://www.cog-genomics.org/plink2/; Purcell et al., 2007). Individuals with low genotype calls were discarded (MIND < 0.1), which means that

it excluded individuals who had more than 10% genotype missingness (Marees et al., 2018). In addition, only SNPs minor allele frequency (MAF) above 5% was included. This resulted in a genotype data file where 343 307 SNPs were typed in 703 individuals. Since data consisted of individuals sampled from multiple genetically differentiated subpopulations (as evident from analyses using low-density microsatellite data; Hemphill et al. in review), I expected the SNP genotype frequencies to deviate from Hardy-Weinberg frequencies, and there was no need to conduct a Hardy-Weinberg equilibrium test.

To ensure that SNPs included in the analyses were approximately independent (i.e. in linkage equilibrium, I carried out linkage disequilibrium (LD) pruning in PLINK with parameters 50 5 0.5. The specified parameters would (1) consider a window of 50 SNPs, (2) calculate LD between each pair of SNPs in the window and remove one of a pair of SNPs if the LD is greater than 0.5 (threshold), and (3) shift the window 5 SNPs forward and repeat the procedure (Purcell et al., 2007; http://zzz.bwh.harvard.edu/plink/summary.shtml#prune). The presence of LD can be a potential explanation for heterozygosity-fitness correlations in a population (Hansson et al., 2004). Genetic association among pairs of loci can be caused by physical linkage between loci, but similar patterns can also be shown by mutation, genetic drift and small populations sizes (Hansson et al., 2004). This resulted in genotype data for a total number of 450 individuals and 70 830 markers, of which 253 individuals were included in the statistical analysis (see above; Table 1).

2.4 Statistical analyses

Various packages developed for the Software R version 3.6.1 (R Core Team, 2012) was used for the genomic analyses. PGDSpider version 2.1.1.0 (Lischer and Excoffier, 2012) was used to convert the corresponding PED and MAP file to STRUCTURE format, since this was a convertible file format needed for at least two of the analyses performed. In following method sections, 2.4.2 estimating genetic differentiation and 2.4.3 sliding window analyses, two different F_{ST}-estimators was used, since the packages used different estimators. However, a correlation test was performed, which showed a strong correlation between the two estimators (Appendix, Figure A1, Table A1).

2.4.1 Estimating within-population genome-wide levels of genetic variation

Heterozygosity is a good measure of genetic diversity within populations, and it can provide valuable information about the history of the population (Samuels et al., 2016). The R package *adegenet* (Jombart, 2008) was used to estimate average observed heterozygosity for the different subpopulations. The average observed heterozygosity was estimated for both early and late period, separately.

2.4.2 Estimating genetic differentiation

The R package *adegenet* (Jombart, 2008) was used to read the structure files and convert them into a genind object, to further transform the data into hierfstat-format. The R package *hierfstat* (Goudet, 2005) was used to estimate the fixation index (F_{ST}), which can be used to investigate processes that influence the distribution of genetic variation between subpopulations, and within a subpopulation over time (Wright, 1949; 1965). The estimation of F_{ST} was conducted using the Weir and Cockerham estimator for F_{ST}. To look at the genetic differentiation between the arctic fox subpopulations at the genome-wide level, pairwise F_{ST} was estimated for each pair of subpopulations. To see how the genetic differentiation varies with time between the subpopulations, pairwise F_{ST} was estimated for pairs subpopulations for both early and late time periods. Furthermore, to examine within-subpopulation levels of genetic variation, pairwise F_{ST} was estimated within each subpopulation by using an early and a late time-period as representatives of two subpopulations.

Pairwise F_{ST} significance was assessed using 1000 bootstraps for the calculation of 95 % confidence intervals. This was conducted using the R package *diveRsity* (Keenan et al., 2013). In advance, the R package *zvau* (Lustrik and Skrbinsek, 2019) was used to convert the file into a gen file, since this was the format compatible with the function used for bootstrapping.

2.4.3 Sliding window analyses

The R package *PopGenome* (Pfeifer et al., 2014) was used in setting up sliding window analyses. Calculating estimates of statistics such as F_{ST} could be relatively noisy when it is calculated on a large number of SNPs. In R, by choosing a particular window size and a jump for that window, it calculates the mean for the statistics within that window (Pfeifer et al., 2014). By doing so, it captures the average variation across larger genomic regions (e.g. chromosomes) and the visualized data would be easier to interpret. Estimation of F_{ST} done along chromosomes,

or scaffolds in this case, using a "sliding window" approach, will provide more detailed information about any genomic regions where the level of genetic differentiation and/or variation is higher or lower than average.

The sliding window approach was performed on the two biggest scaffolds (i.e. largest number of SNPs) and the scaffold on which the gene(s) associated with fur color in the arctic fox are located (Tietgen et al. in prep). The latter scaffold was included in the analysis, since the blue foxes appear to have been introduced to the natural/wild subpopulations through the release of foxes from the breeding program. Although the selection on fur color does not appear to be strong (H. Jensen, *pers. comm*.), it would be interesting to investigate whether the differentiation within and between subpopulations are different in the genome region where the fur color gene is located, compared to genome regions on other scaffolds (i.e. genome-wide).

The R package *PopGenome* contains a function which makes it possible to split a VCF-file into multiple VCFs including only data for exactly one scaffold each (Pfeifer et al., 2014). Since *PopGenome* was not able to read the VCF-files transformed by Plink, the R package *vcfR* (Knaus and Grünwald, 2016) was used to manipulate the files into the right format. Subsequently, *PopGenome* was used to estimate the fixation index (F_{ST}, mode: nucleotide) in sliding windows, with window size 1 000 000 and window jump 500 000, along the scaffolds and genome-wide (Pfeifer et al., 2014). Thereby, the R package *ggplot2* (Wickham et al., 2019) was used to plot the results to see how F_{ST} varies between two of the subpopulations compared along the given scaffold.

2.4.4 Principal component analysis (PCA)

A principal component analysis is a technique of multivariate data analysis, which aims to reduce the dimensionality of large data sets (Maćkiewicz & Ratajczak, 1993; Mishra et al., 2017). This is done by transforming the variables into a set of smaller number of variables called principal components (PCs), without losing much information (Maćkiewicz & Ratajczak, 1993; Mishra et al., 2017). Eigenvalues represent the amount of genetic diversity represented by each PC (Jensen, 2019). In that case, each eigenvalue is the variance of the corresponding PC, and the eigenvalues of a PC divided by the sum of eigenvalues for all PC`s gives the proportion of variance in the total data set explained by the PC (Jensen, 2019).

The genind object was submitted to a Principal Component Analysis, using dudi.pca (Jombart, 2008), to get a summary of the genetic diversity among the individuals/subpopulations.

Preparation of the data was done using the R package *Adegenet* (Jombart, 2008), while the actual analysis was done using the R package *ade4* (Dray et al., 2018). The PCA was run on data containing all individuals in each subpopulation for the early and late time period. The purpose of the analysis was to (i) get a visualization of how genetically different the subpopulations are from each other (at each time period) and (ii) how the genetic clustering changes through time within each subpopulation. Therefore, individuals registered in the early period was retrieved from the PCA analysis and their PCs, containing the scores of each individual, was plotted together. The same procedure was done for the late period.

3. Results

3.1 Genetic variation within subpopulations

Genome wide

In general, no major changes in average observed heterozygosity (Ho) for each subpopulation was seen (Table 2). However, there are at least four interesting points to highlight. (i) The lowest level of average Ho was seen in the early period of Helags, and it has increased over time. (ii) Average Ho has also increased (a little) in both Hardangervidda, Saltfjellet and Vindelfjällen. (iii) Average Ho seems to be approximately constant in Snøhetta. Finally (iv), average Ho seems to have decreased slightly in Borgafjäll.

Table 2. Average observed heterozygosity for each subpopulation. Estimated for both early and late period, separately. See Table 1 for further information on which time periods that are included.

Subpopulation	**Hardangervidda**		**Saltfjellet**		**Snøhetta**		**Borgafjäll**		**Helags**		**Vindelfjällen**	
Time period	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late	Early	Late
Heterozygosity	0.3721	0.3758	0.3816	0.3852	0.3626	0.3619	0.3643	0.3474	0.3268	0.3368	0.3676	0.3694

3.2 Genetic differentiation within subpopulations

Genome wide

The results showed that there is some genetic differentiation between the two time periods within each subpopulation at the genome-wide level (Figure 2). However, only two of the estimated pairwise F_{ST}-values are significantly different from 0, as indicated by their 95% confidence intervals (CI`s) not overlapping with 0 (Table 3). The significant F_{ST}-values are for the subpopulations Snøhetta and Helags (Table 3). These two within-subpopulation comparisons provide also the two highest estimated F_{ST}-values, with a pairwise F_{ST}-value for Snøhetta of approximately 0.0294 and 0.0267 for Helags. Hardangervidda and Vindelfjällen show the lowest level of genetic differentiation between early and late time periods, with F_{ST}-values approximately 0.0116 for Hardangervidda and 0.0108 for Vindelfjällen (but not significantly different from zero; Table 3). The F_{ST}-value for Borgafjäll and Saltfjellet appears intermediate, but is also not significantly different from zero (Table 3).

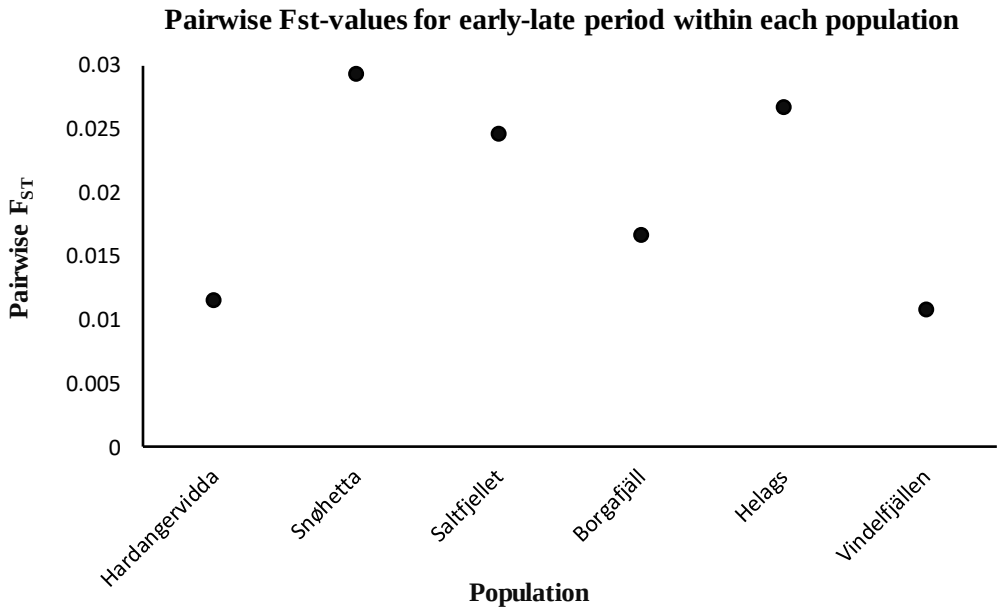


Figure 2. F_{ST} (Weir and Cockerham 1984) for each subpopulation. Calculated by comparing early and late time period within each subpopulation.

Table 3. Estimated pairwise F_{ST}-values (Weir and Cockerham 1984) between two time periods (early-late) within each subpopulation, and 95% confidence interval for each F_{ST}-value. F_{ST}-values significantly different from zero are indicated with an asterisk.

Subpopulation	F_{ST}	CI 95%
Hardangervidda	0.0116	-0.0044–0.0331
Snøhetta	0.0294*	0.0139–0.0500
Saltfjellet	0.0246	-0.0341–0.1168
Borgafjäll	0.0167	-0.0151–0.0713
Helags	0.0267*	0.0072–0.0523
Vindelfjällen	0.0108	-0.0114–0.0416

Sliding window analysis

The levels of genetic differentiation (F_{ST}) within subpopulations (early vs. late) were calculated for genomic regions using a 1000-kb sliding window for the two biggest scaffolds (0 and 3) as well as scaffold 11 (containing gene(s) coding for fur color) in the arctic fox genome (Figure 3; Appendix, Figure A3). In general, all within-population comparisons showed at least one large peak, but at somewhat different positions (Figure 3; Appendix, Figure A3). All three within-

comparisons for Helags showed however large peaks for the SNPs positioned towards the end of the scaffolds (Figure 3). The opposite trend was observed for two of the within-comparisons for Borgafjäll (Appendix, Figure A3(a,c,e)).

SNP AX-177360772 is the one closest to the MC1R gene on the dog-genome and is located at position 21 101 714 on scaffold 11 in the arctic fox genome (Tietgen et al. in prep.). Unfortunately, there is not observed any high peaks at position 21 Mb for neither of the within-subpopulation comparisons on scaffold 11 (Figure 3a; Appendix, Figure A3(e)).

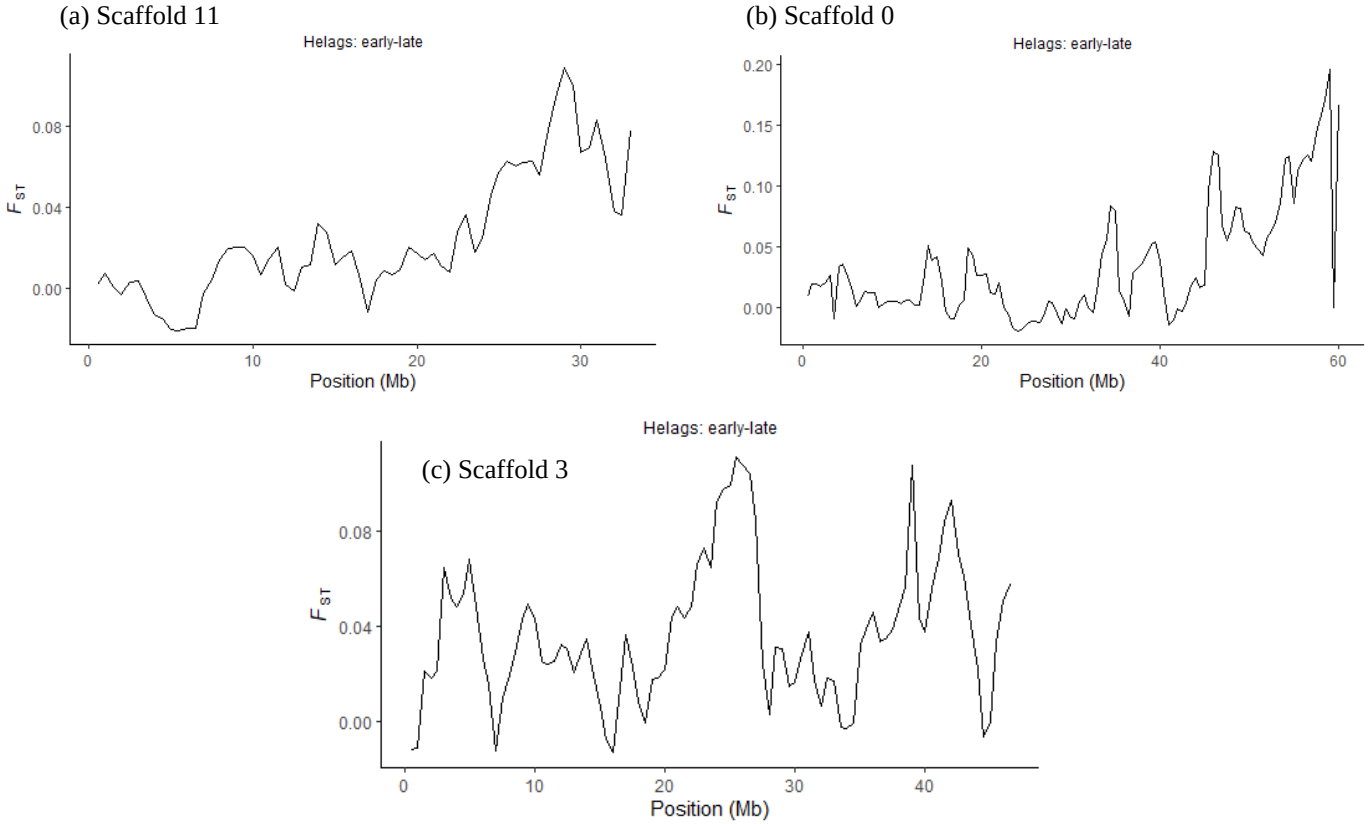


Figure 3. Plot of average F_{ST} calculated using a 1000-kb sliding window for the two biggest scaffolds (b and c) and for the scaffold containing gene(s) coding for fur color (a) in the arctic fox genome. The analysis is performed by comparing early and late period in Helags. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values. See appendix for further plots on the same scaffold analyzed for Borgafjäll and for the two biggest scaffolds

3.3 Genetic differentiation between subpopulations

Genome wide

The majority of the pairwise F_{ST}-values estimated between subpopulations within each time period were significantly different from zero, with the exception of three pariwise F_{ST}-values (Saltfjellet vs. Hardangervidda, Snøhetta vs. Saltfjellet and Vindelfjällen vs. Borgafjäll) for the

early period and only one pairwise F_{ST}-value (Saltfjellet vs. Hardangervidda) for the late period (Table 4-5).

The early time period for Helags compared to the 5 remaining subpopulations, all provide high pairwise F_{ST}-values (Table 4). Helags vs Saltfjellet and Helags vs Snøhetta shows the highest levels of genetic differentiation, with a pairwise F_{ST}-value for the former of approximately 0.1571 and 0.1563 for the latter (Table 4). The two pairs of subpopulation comparisons with the lowest pairwise F_{ST}-values are Saltfjellet vs Hardangervidda and Vindelfjällen vs Borgafjäll, with F_{ST}-values approximately 0.0177 for Saltfjellet vs Hardangervidda and 0.0167 for Vindelfjällen vs Borgafjäll (but not significantly different from zero; Table 4).

As seen for the early time period, Helags still shows high pairwise F_{ST}-values for the late period (Table 4-5). However, only three of these estimated pairwise F_{ST}-values (Snøhetta vs Helags, Borgafjäll vs Helags and Vindelfjällen vs Helags) are higher than shown for other late period pairs of subpopulation comparisons (Table 5). The two pair of subpopulation comparisons with the lowest pairwise F_{ST}-values are Saltfjellet vs Hardangervidda and Snøhetta vs Hardangervidda, with F_{ST}-values approximately 0.0106 for Saltfjellet vs Hardangervidda (but not significantly different from zero) and 0.0259 for Snøhetta vs Hardangervidda (Table 5).

The overall direction of temporal change in the levels of genetic differentiation between subpopulations at the genome-wide level (as measured by pairwise F_{ST}-values) were not found to be consistent (Figure 4). In total, 9 of the comparisons resulted in a decrease in pairwise F_{ST} over time (blue lines in Figure 4), whereas the remaining 6 comparisons resulted in an increase in pairwise F_{ST} over time (red lines in Figure 3). 5 of the 9 comparisons resulting in a decrease in pairwise F_{ST} belongs to the Helags comparisons. How much the F_{ST}-values increased or decreased differed between the subpopulation pairs, with the biggest change seen in Saltfjellet vs. Helags (Figure 4). The smallest change was observed in Hardangervidda vs. Snøhetta with a decrease in Pairwise F_{ST} of approximately 0.0012 (Figure 4; Table 4-5). When comparing the 95% confidence intervals (CI) for early and late time period, to see if some of the changes are "significant" (i.e. 95% CI do not overlap), 5 of the decreases in pairwise F_{ST} seems "significant" and none of the increases in pairwise F_{ST} seems "significant" (Figure 4, Table 4-5).

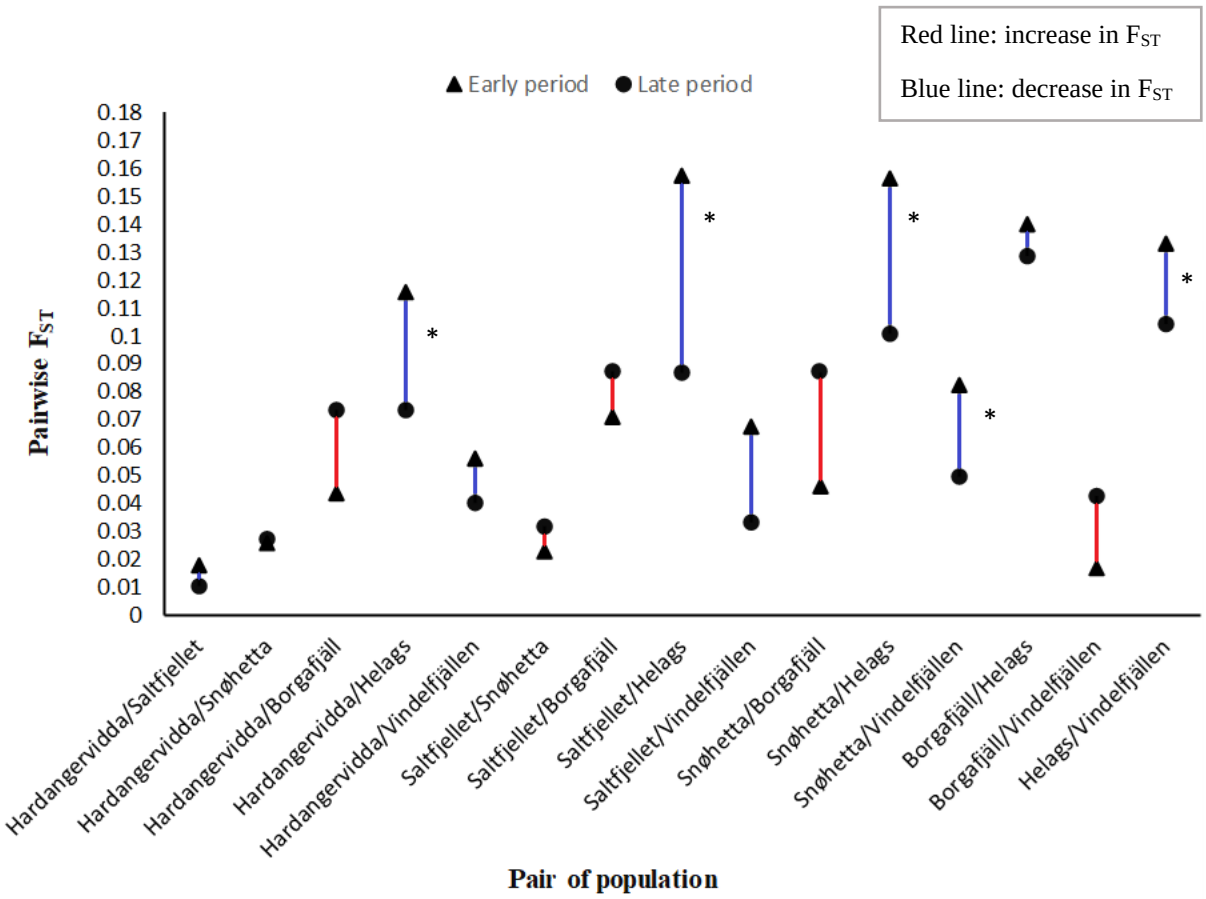


Figure 4. Pairwise F_{ST} (Weir and Cockerham 1984) for each pair of subpopulations in the two time-periods (▲early and ● late). Red lines represent an increase in pairwise F_{ST} from early to late period. Blue lines represent a decrease in pairwise F_{ST} from early to late period. "Significant" decreases in F_{ST} are marked with an asterisk.

Table 4. Matrix of pairwise F_{ST}-values (Weir and Cockerham 1984) between the subpopulations for early time periods (above the diagonal), and the 95% confidence interval for the corresponding F_{ST}-value (below the diagonal). F_{ST}-values significantly different from zero are indicated with an asterisk. See Table 1 for further information on which time periods that are included.

Subpopulation	Hardangervidda	Saltfjellet	Snøhetta	Borgafjäll	Helags	Vindelfjällen
Hardangervidda	_	0.0177	0.0259*	0.0435*	0.1156*	0.0562*
Saltfjellet	-0.0137–0.0715	_	0.0227	0.0708*	0.1571*	0.0674*
Snøhetta	0.0064–0.0559	-0.0111–0.0931	_	0.0459*	0.1563*	0.0826*
Borgafjäll	0.0138–0.0870	0.0246–0.1430	0.0085–0.1010	_	0.1399*	0.0167
Helags	0.0901–0.1467	0.1330–0.2048	0.1274–0.1821	0.1154– 0.1765	_	0.1332*
Vindelfjällen	0.0336–0.0889	0.0309–0.1342	0.0586–0.1134	-0.0270–0.0803	0.1135–0.1607	_

Table 5. Matrix of pairwise F_{ST}-values (Weir and Cockerham 1984) between the subpopulations for late time periods (above the diagonal), and the 95% confidence interval for the corresponding F_{ST}-value (below the diagonal). F_{ST}-values significantly different from zero are indicated with an asterisk. See Table 1 for further information on which time periods that are included.

Subpopulation	Hardangervidda	Saltfjellet	Snøhetta	Borgafjäll	Helags	Vindelfjällen
Hardangervidda	_	0.0106	0.0271*	0.0736*	0.0734*	0.0400*
Saltfjellet	-0.0138–0.0534	_	0.0319*	0.0872*	0.0870*	0.0334*
Snøhetta	0.0172–0.0403	0.0050–0.0720	_	0.0870*	0.1009*	0.0497*
Borgafjäll	0.0625–0.0884	0.0628–0.1249	0.0758–0.1032	_	0.1285*	0.0427*
Helags	0.0610–0.0879	0.0511–0.1415	0.0840–0.1217	0.1154– 0.1464	_	0.1041*
Vindelfjällen	0.0285–0.0563	0.0076–0.0735	0.0384–0.0655	0.0262–0.0642	0.0903–0.1203	_

Sliding window analysis

The levels of genetic differentiation (F_{ST}) between subpopulation pairs early and late, were also calculated for genomic regions using a 1000-kb sliding window for the two biggest scaffolds (0 and 3) as well as scaffold 11 (containing gene(s) coding for fur color) in the arctic fox genome, respectively (Figure 5). Additional plots for scaffold 0 and 3, and initial plots for the scaffold associated with fur color in the arctic fox genome can be found in the Appendix, Figure A4-A6. In general, there seems to be some fluctuations in all comparisons on scaffold 0, 3 and 11 (Figure 5; Appendix, Figure A4-A6). However, F_{ST}-values vary from very small, almost flat (e.g. Vindelfjällen/Borgafjäll in Figure 5b) to imply higher peaks (e.g. Helags/Borgafjäll in Appendix, Figure A5(b,c)).

For the comparison Helags/Borgafjäll (early period), there was a large peak at around 40 Mb on scaffold 0 (Figure 4a). When Borgafjäll compared to the other 4 subpopulations (Vindelfjällen, Hardangervidda, Saltfjellet and Snøhetta) in the early period, no such peak was observed (Figure 4a). On the other hand, a tendency towards small peaks around 40 Mb for all comparisons with Borgafjäll was observed for the late period (Appendix, Figure A4(c)). When comparing Helags to all subpopulations (early period), a peak around 40 Mb was observed for

each comparison on scaffold 0 (Appendix, Figure A4(a)). Similar observations were seen for all subpopulation comparisons with Helags for the late period (Appendix, Figure A4(b)).

Overall, small variations were observed in F_{ST} along scaffold 3 (early and late period) for both Helags and Borgafjäll compared with each of the other subpopulations (Figure 5b; Appendix, Figure A5(a-c)). There was, however, a larger peak between 5 and 10 Mb for Borgafjäll/Helags and Vindelfjällen/Helags subpopulation pairs than for the other subpopulations (Figure 5b). Otherwise, the comparisons tend to fluctuate together, or at least in the same way, except for the comparison Helags/Borgafjäll (late period) which have higher peaks between 20 and 30 Mb than the other subpopulations compared to either Helags or Borgafjäll (Figure 5b; Appendix, Figure A5(c)).

(a) Scaffold 0 – early period

(b) Scaffold 3 – late period

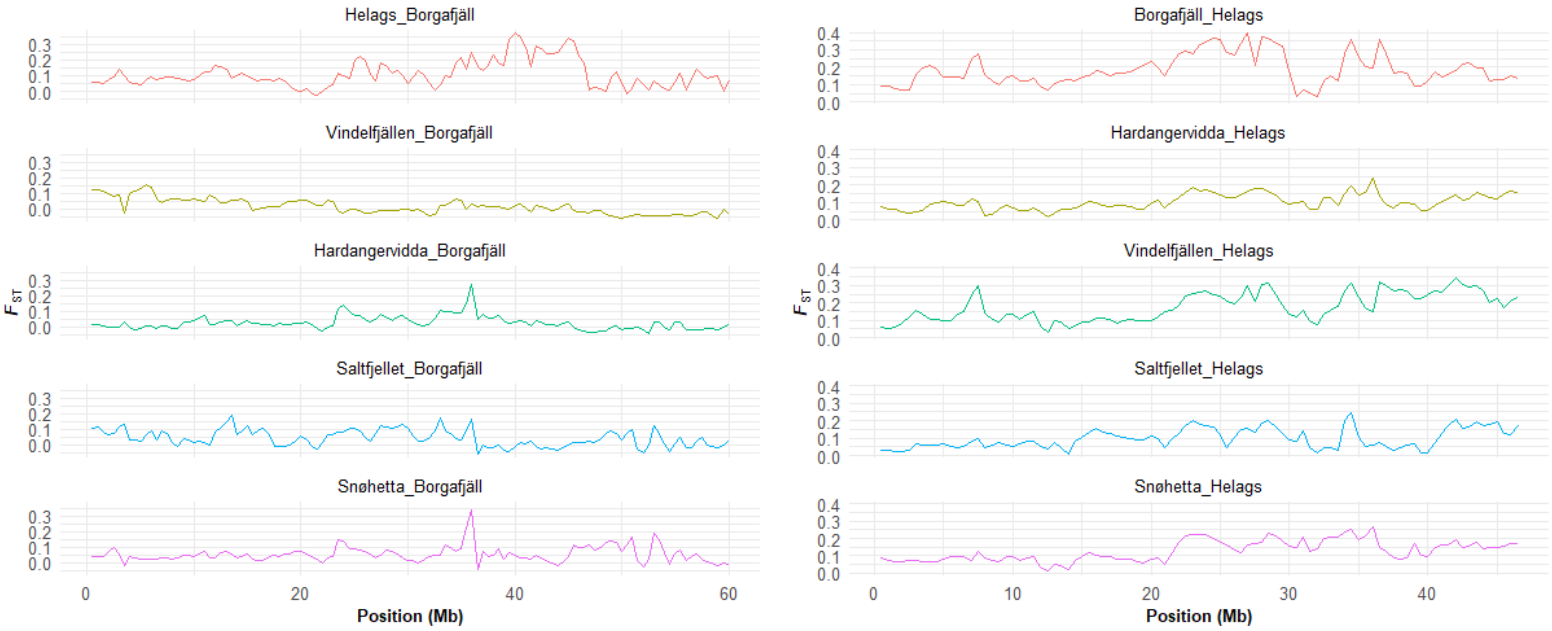


Figure 5. Plots of F_{ST} calculated within 1000-kb sliding window for the two biggest scaffolds (0 and 3) in the arctic fox genome. In (a) early time period in Borgafjäll is compared with each of the other populations. In (b) late period in Helags is compared with each of the other populations. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values within the sliding windows.

3.4 Principal component analysis (PCA)

Seeking a summary of the genetic diversity among the sampled individuals/subpopulations

As expected, the first eigenvalues were large for the first principal components (PCs), and smaller for the consecutive PCs (Appendix, Figure A6). The two first eigenvalues explained 7.9 and 5.2 % of the variance, respectively, which means that approximately 13 % of the variation was explained by the two first principal components combined (Appendix, Table A3, Figure A7).

The genetic difference between individuals (i.e. points) in a PCA plot is represented by the distance of the points (Figure 6). That is, the further away the points are, the more genetically different they are. Overall, in both early and late period of the program, Hardangervidda and Saltfjellet had points (i.e. individuals) quite spread (Figure 6a, b). Especially Helags seemed to be quite divergent from the other subpopulations (Figure 6a, b). In addition to Helags, Vindelfjällen and Borgafjäll formed more tightly grouped clusters in the early period, with the exception of two individuals from Borgafjäll located in the lower left-hand corner of the PCA plot (Figure 6a).

As expected, a denser pattern was observed in the PCA plot for the late period (Figure 6b). It appeared that the points (i.e. individuals) was located more around the center of the plot (Figure 6b), and not as far apart as observed in the beginning (Figure 6a). Both Helags and Vindelfjällen, seemed not to form less tightly clustered groups (Figure 6b). Snøhetta seemed to be somewhat more spread and a large proportion of the points was now closer to Borgafjäll (Figure 6b).

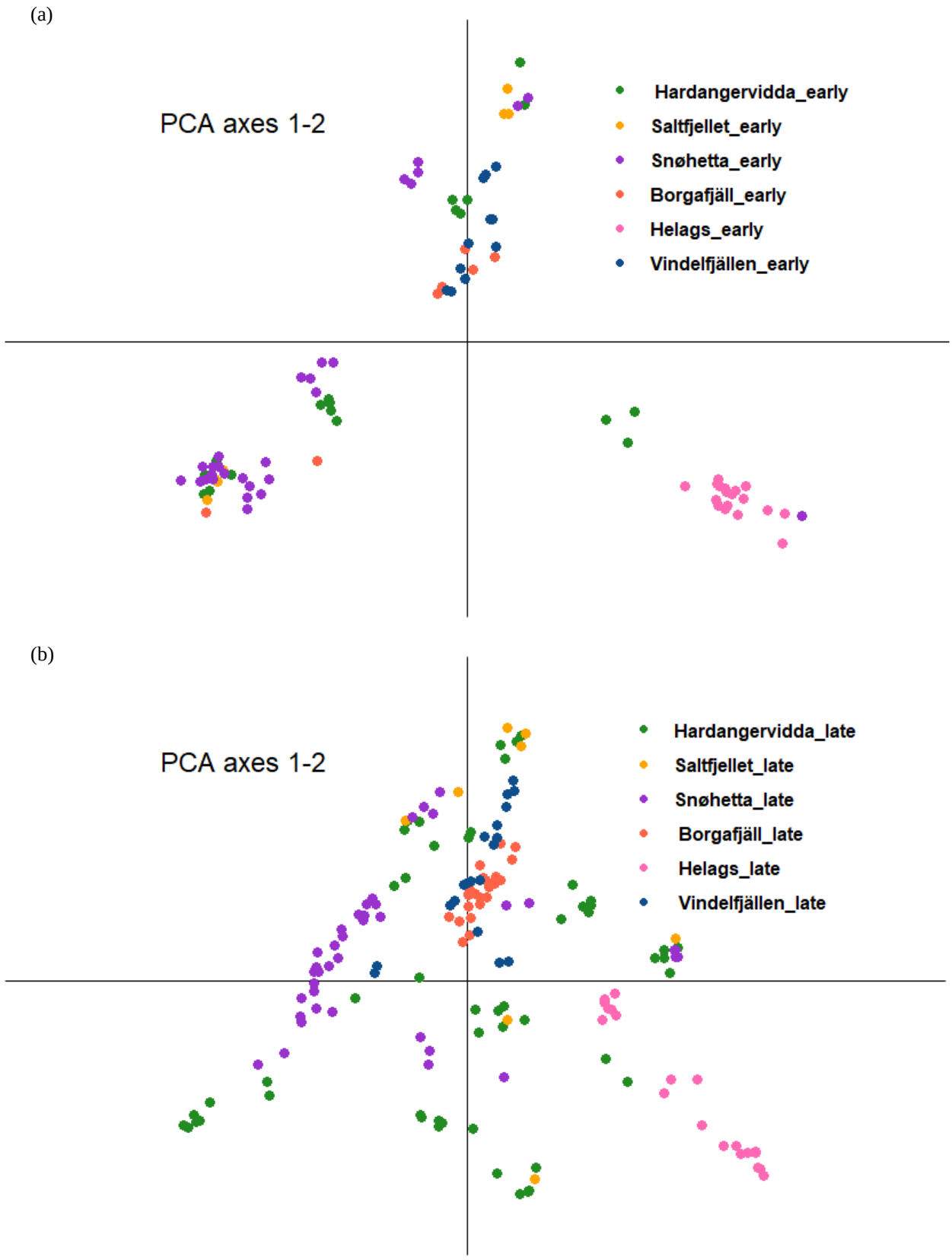


Figure 6. Principal component analysis (PCA) showing the genetic clustering of 6 arctic fox subpopulations in Scandinavia. Figure (a) shows early time-period and figure (b) shows late time-period. See Table 1 for further information on which time periods that are included. The x-axis represents the first principal component and the y-axis represent the second principal component. The axes have the exact same limits. The two first principal components explain 7.9 and 5.2 % of the variance, respectively (see Appendix, Table A3).

4. Discussion

In recent decade, the Scandinavian arctic fox population has increased significantly in number of individuals (Ulvund and Wallén, 2018). In addition, several subpopulations have been reestablished. The implementation of several conservation and management measures has probably been a major influence regarding this positive trend. Several studies indicate that translocations of individuals may be a good approach in supplementing threatened species and/or to reintroduce species to areas where they no longer exist (Smith and Clark, 1994; Servheen et al., 1995; Johnson et al., 2010). Also, captive breeding and release has shown to be successful for many species (Phillips and Parker, 1988; Kleiman, 1989; Johnson et al., 2010). Previous studies show that the success of the reintroduction is dependent on certain conditions, e.g. a viable captive-bred population, suitable available habitats (including habitat preservation and management), release and long-term monitoring and so on (Kleiman, 1989; Landa et al., 2017). Successful captive-breeding programs may therefore act as models for what to base future conservation and management actions (reintroductions) on. As seen in previous studies on the arctic fox (Landa et al., 2017; Hemphill et al. in review), the current captive-breeding program has successfully strengthened and re-established subpopulations in several Scandinavian mountain areas.

4.1 Genetic variation and differentiation within subpopulations

Genome-wide

In general, all within-population estimates indicated a level of genetic differentiation between early and late time period (Figure 2). Genetic drift and immigration (both naturally and as a result of release from the breeding station), has probably had an impact on the genetic variation within the subpopulation, and thereby indirectly affecting the temporal genetic differentiation. A case study on the Florida Panther (Phillip, 1995) points out in particular the importance of genetic drift in small populations. In an endangered species, a small population size could have an effect on the outcome even though selection is at work and the population is experiencing gene flow (Phillip, 1995). In my study though, the time period is probably too small for selection to work.

Average observed heterozygosity (Ho) has increased (a little) in Hardangervidda, Saltfjellet and Vindelfjällen (Table 2). This may be due to immigration and mixing of lineages from the captive breeding programme, which has affected all three populations considerably. In addition,

population sizes have increased (Ulvund et al., 2018), counteracting loss of genetic variation by genetic drift. A comparable increase in Ho was not seen in Snøhetta even though the genetic differentiation between early-late period within Snøhetta was one of the highest one of all subpopulations. Increased differentiation may be explained by overrepresentation of one or two founder lineages from the breeding programme during the early release phase, when some pairs in the breeding station contributed with a disproportionally high proportion of released pups (Ø. Flagstad, *pers. comm.*). With a more balanced lineage representation in the population over time, allele frequencies may have shifted, without increasing the overall genetic variation.

The early period of Helags represents the original subpopulation, which was quite inbreed and originated from only five individuals (Noren et al., 2016). From 2011 and onwards, at least three immigrants (AF0089, AF0091, AF0120) had established and breed in Helags ((Ø. Flagstad, *pers. comm.*; Hasselgren et al., 2018). Two of them were found to be siblings, the so-called "Blues brothers", and they experienced good reproductive success (Hasselgren et al., 2018). The immigration of these three foxes could therefore be one possible explanation for why Helags is one of the subpopulations that showed the highest pairwise F_{ST} within-comparison, in addition to the low average observed heterozygosity (Ho) seen in the early period, and the increase of average Ho over time. Hasselgren et al. (2018) documented positive effects (i.e. higher fitness and increased population growth) as a result of the three male's immigration. Even though the immigration brought about an "genetic rescue effect" in Helags, Hasselgren et al. (2018) points out that it is difficult to predict the long-term effects of such sporadic immigration events in an inbred population and that it may also result in negative consequences if the population is not supplemented by further immigration.

Before the conservation and management actions were implemented, Borgafjäll was the biggest population and had the most stable rodent cycles (Ø. Flagstad, *pers. comm.*). Relatively few conservation measures have been conducted here, nor is it believed that this subpopulation has been significantly affected by the Breeding program (Ø. Flagstad, *pers. comm.*; Ulvund et al., 2018). The low level of genetic differentiation within Borgafjäll (comparing early-late period), in addition to the slight reduction of average Ho, may be a result of little natural immigration and no release of captive bred foxes. That is, drift has probably reduced average Ho in Borgafjäll.

There was no significant relationship between the number of generations between early and late period and the change in F_{ST} within each subpopulation (Appendix, Figure A2). This indicate that there are other important factors, rather than time between sampling, for the F_{ST} within

subpopulations, such as Ne (drift) and natural (im)migration and immigration of foxes from the breeding programme. However, few data points probably provide little power to find any firm connections. If the data covered more populations and a larger variation in the number of generations, it is conceivable that a significant relationship would appear, which would then mean that drift and immigration have had longer time to bring about changes.

Sliding window

The time periods in this study are probably too short for recombination to happen. Meaning that there is a very low chance for observing any patterns resulting from selection on a regional scale. It is more likely to observe patterns resulting from drift and/or immigration. So, the observed trend in large peaks for SNPs positioned at towards the end of the scaffold, may probably only be a result of drift and/or immigration. However, Tietgen et al. (in prep.) has found SNP-markers which has a significant correlation with coat color for the arctic fox (and which is also in LD with the candidate gene for coat color, MC1R), up to 15 million Bp from the MC1R-gene. This means that the peak observed for the within-comparison for Helags on scaffold 11 (Figure 3a) may potentially represent elevated differentiation on gene (s) that either affect coat color or that are in LD with coat color (s).

4.2 Genetic differentiation between subpopulations

Genome-wide

In general, the results from the pairwise F_{ST}-analysis, conducted comparing two subpopulations for both early and late period separately, showed mostly signs of moderate genetic divergence. However, in the early period, there was several comparisons which can be assigned to the category "great divergence". High values of pairwise F_{ST} (i.e. strong differentiation) could indicate that there is low gene flow between the subpopulations. Dalén et al. (2006) report that immigration (dispersal) among populations at that time was probably very low. Pongratz, Gerace and Michiels (2002) also found that populations further apart (i.e. from different lakes rather than within the same lake) were more genetic differentiated (i.e. higher F_{ST}). Nevertheless, an increase in immigration between some of the subpopulations after the release from the breeding program and implementation of additional actions (i.e. supplementary feeding and culling of red fox) have been reported (Eide et al., 2017; Ulvund et al., 2018). However, since the genetic diversity within the breeding station was supposed to represent the genetic diversity left in the Scandinavian arctic fox population, a decrease in genetic

differentiation between the subpopulations from early to late period is seen in several of the comparisons.

The patterns of genetic differentiation could also provide information about the evolutionary past of a population. E.g. if a population recently went through a bottleneck, the genetic diversity could be very low (Dalén et al., 2006). However, no evidence for any recent Bottlenecks was found in Dalén et al. (2006) study of the arctic fox. Rather, the results support a hypothetical fragmentation, which may origin from the altitudinal expansion of the red fox (Dalén et al., 2006).

Several of the reductions in pairwise F_{ST} from early to late period seem to be significant. After the release of captive bred foxes, several of the pairwise F_{ST} comparisons between subpopulations decreased in value (Figure 4). A high proportion of the decreases seemed to be significant, suggesting increased connectivity in the metapopulation (Slatkin, 1987; Wade and McCauley, 1988; Hale et al., 2001). The ones with the largest decrease in pairwise F_{ST} (early-late period) was the comparisons with Helags. Since the early time period for Helags was quite inbred and originally based on only five individuals, it is not surprisingly that the genetic differentiation between Helags and other subpopulations would decline after influence from the breeding program or immigration from foxes connected to the breeding program.

Sliding window

Several studies indicate that variation of diversity within the genome can provide valuable information about different evolutionary effects (Akhunov et al., 2010; Bentley et al., 2017). E.g. it could say something about the strength of selection in different parts of the genome. Usually, the estimates of population differentiation between some of the loci will not be independent when using a lot of genome-wide loci (Lotterhos, 2019). However, the data set was pruned, so that the loci used in the analyses should not be in high linkage disequilibrium. An individual "brings" along its whole genome when released and/or migrates to a new subpopulation. Independent sorting of homologous chromosomes, in addition to recombination, "breaks up" the link between genes/markers in the following generations (Hunter, 2015). The temporal dimension of this study is relatively short, so big "chunks" of the genome will not yet be broken and the "immigrant effect" (natural and released) will thereby at this stage only be seen at the genome level.

4.3 Principle components analysis

A principle components analysis (PCA) reduces the dimensionality and can help with visualization of the data and to find clusters (Lever, Krzywinski and Altman, 2017). The genetic clustering of the subpopulations in the PCA plot seemed to be in alignment with the results from the pairwise F_{ST} comparisons. A higher level of migration between the subpopulations has been reported after the releases from the breeding station (Hemphill et al. in review; Hasselgren et al., 2018). According to Pontgratz, Gerace and Michiels (2002) this leads to less genetic differentiation. Meaning, that the observed denser pattern with less space between the individuals/subpopulations are partly a result of natural immigration between the subpopulations and/or release of arctic foxes from the breeding program. However, the population structure estimates may be disproportionately affected by non-independence among SNPs due to linkage disequilibrium (Lotterhos, 2019). Even though a linkage disequilibrium pruning was performed, there are studies indicating that the SNPs will never be completely independent (i.e. they are quasi-independent) (Lotterhos, 2019). In addition, several studies show and discuss the different factors that may influence a PCA (Novembre and Stephens; Przeworski, 2009; Lotterhos, 2019). However, there are situations where the PCA will provide a good data summary, e.g. when the populations are in proximity of the source population (Przeworski, 2009). This seems to hold for my data set, reflecting less genetic differentiation between subpopulations, which indeed is compatible with increased migration across the metapopulation.

The fact that Helags (early period) represents the original subpopulation, and not individuals released from the breeding program, is probably the reason for why Helags seemed to be quite divergent from the other subpopulations in the PCA analysis (Figure 7a). In addition, after immigration from three male arctic foxes which is linked to the breeding program, the distance decreased (i.e. less divergent, Figure 7b). This result is in alignment with the increase in average Ho for Helags, and the observed reductions in pairwise F_{ST} for Helags compared to the other subpopulations. Even though immigration often is thought to be beneficial to small, fragmented populations (e.g. by reducing inbreeding levels), there are studies showing that these benefits could become short-lived (Hagenblad et al., 2009; Hedrick and Fredrickson, 2010; Adams et al., 2011). After a while, if the immigration becomes too overwhelming and the immigrant ancestry replaces other individuals in the population (i.e. losing to much local variation), then it could have an opposite effect, by making the population more inbred (Hedrick and Fredrickson, 2010; Hasselgren et al., 2018).

4.4 Limitations and future perspectives

The Software R comes with some restrictions according to available functions and statistics in each package. Unfortunately, the Weir and Cockerham (1984) estimator for F_{ST}, which was used in calculation of pairwise F_{ST}, was not available for the R package used in the sliding window approach. A strong positive correlation was however found when comparing the F_{ST} estimator used in the sliding window approach with the Weir and Cockerham (1984) estimator (Appendix, Figure A1). In addition, the Weir and Cockerham method is perhaps the most common used estimator for estimating F_{ST}.

Weir and Cockerham`s (1984) F_{ST} work fine with moderate sample sizes (n=15, 20, 25), but tends to slightly overestimate genetic differentiation when the sample size is quite low. The statistical power of this study is probably low due to the low number of individuals in certain subpopulations (Table 1). However, it has been shown that using a high number of markers (number of SNPs) and an appropriate estimator, then the sample size can be reduced without losing much statistical power (Weir and Cockerham, 1984; Menashe, 2008).

The initial thought was to estimate pairwise F_{ST} for each year to compare and look at changes in F_{ST}. Then it would have been possible to see if the F_{ST} fluctuates or if there is overall directional change. But due to missing year-data for the Swedish populations and limited time, it was necessary to make groups of early and late years to have sample size that were at least 7 individuals. Even though the results provide the opportunity to look at changes in F_{ST} over time, the results are somewhat limited considering that they do not provide any information about changes in F_{ST} on an annual basis. In addition, it may be important to take into account the differences in sample size and consider that there may be a chance that the individuals do not represent the genetic composition of the subpopulation. E.g. for Borgafjäll, only a small number of the total population is represented in the data set. Also for Saltfjellet, an important part of the population is missing. The foxes represented in my study only represents individuals released from the breeding programme, whereas the native population is missing.

This study, in alignment with Hemphill et al. (in review), shows some of what is possible to do with the arctic fox data in Scandinavia now. This study shows that high-density genome-wide genotype data, in combination with an arctic fox reference genome, open up new possibilities within conservation genomics and related questions for the Scandinavian arctic fox. With available genomic data on a threatened species which undergoes various conservation management actions, it is possible to investigate effects on both a genome-wide (i.e. at many loci) and regional/locus-specific levels. The latter is the foundation of genome-scans/F_{ST}-outlier

analysis, such as conducted in sliding window analyses. Further studies with longer time periods using approaches similar to the ones in this study may be able to reveal how genetic drift, migration (gene flow) and selection interact to shape the variation within regions and across the whole genome, and thus make it possible to draw conclusions about genetic consequences of on-going conservation management actions.

The results from this study and Hemphill et al. (in review), in alignment with previous research on the Scandinavian arctic fox populations, could all help in the conservation of the arctic fox forwards. Knowing and understanding the genetic variation and differentiation within and between threatened species, can provide valuable information about where or what kind of conservation actions that is further needed and/or which subpopulations that seems to have recovered after the implemented actions.

5. Conclusion

The results from this study suggest that the reintroductions from the Norwegian Arctic Fox Captive Breeding Program has so far had success in a genetic conservation perspective. By releasing captive-bred foxes from the program, the genetic differentiation seemed to decrease between the subpopulations in Scandinavia, suggesting increased migration and subsequent reproductive contribution from the migrants. These results are not unexpected, but the question remains whether the identified changes occurred surprisingly rapidly. Both the generation time and temporal scale of this study was relatively short. Therefore, it will be important to further investigate how the genetic variation and diversity both within and between the subpopulations will change over a much larger temporal scale. Then it will be possible to find and reveal several factors affecting the genetic diversity of the subpopulations (e.g. selection), and thereby draw stronger conclusions.

References

Adams, J.R. *et al.* (2011) Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. *Proceedings of The Royal Society.* 278, pp. 3336-3344. doi: 10.1098/rspb.2011.0261.

Akhunov, E.D. *et al.* (2010) Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics.* 11(702). doi: 10.1186/1471-2164-11-702.

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

Angerbjörn, A. *et al.* (2008) Saving the endangered Fennoscandian Alopex lagopus – SEFALO+. Final report with comments on the financial report. Stockholm, Sweden. Available at: http://internt.zoologi.su.se/en/research/alopex/publications/Angerbjoernetal2008.pdf (accessed 26 January 2020).

Angerbjörn, A. *et al.* (2013) Carnivore conservation in practice: replicated management actions on a large spatial scale. *Journal of Applied Ecology.* 50, pp. 59-67. doi: 10.1111/1365-2664.12033.

Angerbjörn, A. and Tannerfeldt, M. (2014) *Vulpes lagopus*. *The IUCN Red List of Threatened Species*. Available at: http://dx.doi.org/10.2305/IUCN.UK.2014-2.RLTS.T899A57549321.en (accessed 30 August 2019).

Artsdatabanken (2015) *Vulpes lagopus (Linnaeus, 1758).* Available at: https://artsdatabanken.no/Rodliste2015/rodliste2015/Norge/48032 (accessed 25 August 2019).

ArtDatabanken (2015) Fjällräv (*Vulpes lagopus*). Available at: https://artfakta.se/naturvard/taxon/vulpes-lagopus-100005 (accessed 26 January 2020).

Bentley, A.R., Callier, S. and Rotimi, C.N. (2017) Diversity and inclusion in genomic research: why the uneven progress? *Journal of Community Genetics.* 8(4), pp. 255-266. doi: 10.1007/s12687-017-0316-6.

Dalén, L. *et al.* (2006) Population structure in a critically endangered arctic fox population: Does genetics matter? *Molecular Ecology.* 15(10), pp. 2809-2819. doi: 10.1111/j.1365-294X.2006.02983.x.

Direktoratet for naturforvaltning (2003) *Handlingsplan for fjellrev.* Rapport 2003-2. Available at: http://tema.miljodirektoratet.no/old/dirnat/attachment/37/Rapport%202003-2.pdf (accessed 25 August 2019).

Dray, S. et al. (2018) ade4: Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Science. Available at: https://CRAN.R-project.org/package=ade4 (accessed 19 January 2019).

Eide, N. E. *et al.* (2017) *Handlingsplan for fjellrev (Vulpes lagopus), Norge-Sverige 2017-2021*. Miljødirektoratet M-794 | 2017. Available at: https://www.miljodirektoratet.no/globalassets/publikasjoner/M794/M794.pdf (accessed 25 August 2019).

Elmhagen, B. *et al.* (2017) Homage to Hersteinsson & Macdonald: Climate warming and resource subsidies cause red fox range expansion and Arctic fox decline. *Polar Research*. 36(3). doi: 10.1080/17518369.2017.1319109.

Ericson, M. (2014a) Felles Fjellrev 2010 - 2013: Slutrapport. Stockholm, Sweden. Available at: http://www.fellesfjellrev.se/wp-content/uploads/FellesFjellrev_Slutrapport_131231.pdf (accessed 26 January 2020).

Ericson, M. (2014b) Projekt Felles Fjellrev förlängning 2014: Slutrapport. Stockholm, Sweden. Available at: http://www.fellesfjellrev.se/wp-content/uploads/FellesFjellrev_Slutrapport_140831.pdf (accessed 26 January 2020).

Felles Fjellrev II (no date) Available at: http://www.fellesfjellrev.se/om-felles-fjellrev/detaljerad-projektbeskrivning/ (accessed 26 January 2020).

Goudet J. (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*. 5, pp. 184–186. doi: 10.1111/j.1471-8278 .2004.00828.x.

Hagen, I. J. *et al.* (in prep.) Genomics in arctic fox conservation and ecology: development and utility of the custom fox Affymetrix 702k array.

Hagenblad, J. *et al.* (2009) Population genomics of the inbred Scandinavian wolf. *Molecular Ecology*. 18(7), pp. 1341-1351. doi: 10.1111/j.1365-294X.2009.04120.x.

Hale, M.L et al. (2001) Impact of Landscape Management on the Genetic Structure of Red Squirrel Populations. *Science*. 293(5538), pp. 2246-2248. Available at: https://www.jstor.org/stable/3084725 (accessed 27 January 2020).

Hansson, B. *et al.* (2004) Does Linkage Disequilibrium Generate Heterozygosity-Fitness Correlations In Great Reed Warblers? *Evolution.* 58(4), pp. 870-879. doi: 10.1554/03-368.

Hasselgren, M. *et al.* (2018) Genetic rescue in an inbred Arctic fox (*Vulpes lagopus*) population. *Proceedings. Biological Science.* 285(1875). doi: 10.1098/rspb.2017.2814.

Hedrick, P.W. and Fredrickson, R. (2010) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservative Genetics.* 11, pp. 615-626. doi: 10.1007/s10592-009-9999-5.

Hemphill, E.J.K. *et al.* (in review) Genetic consequences of conservation action: restoring the arctic fox (Vulpes lagopus) population in Scandinavia. Biological Conservation.

Hersteinsson, P. *et al.* (1989) The Arctic Fox in Fennoscandia and Iceland: Management Problems. *Biological Conservation.* 49(1), pp. 67-81. doi: 10.1016/0006-3207(89)90113-4.

Hunter, N. (2015) Meiotic Recombination: The Essence of Heredity. *Cold Spring Harbor Perspectives in Biology*. 7(12). doi: 10.1101/cshperspect.a016618.

Ims, R.A., Henden, J.A. and Killengreen, S.T. (2008) Collapsing population cycles. *Trends in Ecology & Evolution.* 23(2), pp. 79-86. doi: 10.1016/j.tree.2007.10.010.

Isaac, N. J. B. and Cowlishaw G. (2004) How species respond to multiple extinction threats. *Proceedings. Biological Science.* 271(1544), pp. 1135-1141. doi: 10.1098/rspb.2004.2724.

Jensen, H. (2019) Trinidadian guppy. *BI3083 Evolutionary and Ecological Genetics*. Available at: https://ntnu.blackboard.com (accessed 14 October 2019).

Johnson, W.E. *et al.* (2010) Genetic Restoration of the Florida Panthern. *Science*. 329(5999), pp. 1641-1645. doi: 10.1126/science.1192891.

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

Keenan, K. *et al.* (2013) diversity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecology Evolution*. 4(8), pp. 782-788. doi: 10.1111/2041-210X.12067.

Kleiman, D.G. (1989) Reintroduction of Captive Mammals for Conservation: Guidelines for reintroducing endangered species into the wild. *BioScience*. 39(3), pp. 152-161. Available at: https://www.jstor.org/stable/1311025 (accessed 21 January 2020).

Knaus, B.J. and Grünwals, N.J. (2016) VCFR: a package to manipulate variant call format data in R. *Molecular Ecology Resources*. 17(1), pp. 44-53. doi: 10.1111/1755-0998.12549.

Landa, A. *et al.* (2017) The endangered Arctic fox in Norway— the failure and success of captive breeding and reintroduction. *Polar Research,* 36(sup1), 9. doi: 10.1080/17518369.2017.1325139.

Landa, A. *et al.* (2018) *Avlsprogrammet for fjellrev – Årsrapport 2017*. (NINA Rapport;1527). In: Norsk Institutt for Naturforskning (NINA). Available at: http://hdl.handle.net/11250/2500715 (accessed 25 August 2019).

Landa, A. *et al.* (2019) *Avlsprogrammet for fjellrev – Årsrapport 2018.* (NINA Rapport; 1685) In: Norsk Institutt for Naturforskning (NINA). Available at: http://hdl.handle.net/11250/2601583 (accessed 6 January 2020).

Lande, R. (1988) Genetics and Demography in Biological Conservation. Science. 241(4872), pp. 1455-1460. Available at: www.jstor.org/stable/1702672 (accessed 9 January 2020).

Lande, R. (1998) Anthropogenic, Ecological and Genetic Factors in Extinction and Conservation. *Research on Population Ecology.* 40(3), pp. 259-269. doi: 10.1007/BF02763457.

Larsson, P. *et al.* (2019) Consequences of past climate change and recent human persecution on mitogenomic diversity in the arctic fox. *Philosophical Transactions of the Royal Society B.* 374(1788). doi: 10.1098/rstb.2019.0212.

Lever, J., Krzywinski, M. and Altman, N. (2017) Principal component analysis. *Nature Methods*. 14, pp. 641-642. doi: 10.1038/nmeth.4346.

Lischer, H.E.L and Excoffier, L, (2012) PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics.* 28: pp. 298-299.

Loison, A., Strand, O. and Linnell, J. D. C. (2001) Effect of temporal variation in reproduction on models of population viability: a case study for remnant arctic fox (Alopex lagopus) populations in Scandinavia. *Biological Conservation*. 97(3), pp. 347–359. doi: 10.1016/S0006-3207(00)00134-8.

Lotterhos, K.E. (2019) The Effect of Neutral Recombination Variation on Genome Scans for Selection. *G3 (Bethesda)*. 9(6), pp. 1851-1867. doi: 10.1534/g3.119.400088.

Lustrik, R. and Skrbinsek, T. (2019) romunov/zvau: Misc functions used by our ecology lab. Available at: https://rdrr.io/github/romunov/zvau/ (accessed 13 January 2020).

Maćkiewicz, A. and Ratajczak, W. (1993) Principal Components Analysis (PCA)*. *Computers & Geosciences*. 19(3), pp. 303-342. doi: 10.1016/0098-3004(93)90090-R.

Marees, A.T. *et al.* (2018) A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research.* 27(2). doi: 10.1002/mpr.1608.

McMahon, B.J., Teeling, E.C. and Höglund, J. (2014) How and why should we implement genomics into conservation? *Evolutionary Applications.* doi: 10.1111/eva.12193.

Menashe, I. et al. (2008) PGA: power calculator for case-control genetic association analyses. *BMC Genetics*. 9(36). doi: 10.1186/1471-2156-9-36.

Mishra, S. *et al.* (2017) Multivariate Statistical Data Analysis- Principal Component Analysis (PCA). *International Journal of Livestock Research*. 7(5), pp. 60-78. doi: 10.5455/ijlr.20170415115235.

Noren, K. et al. (2016) Inbreeding depression in a critically endangered carnivore. *Molecular Ecology*. 25(14), pp. 3309-3318. doi: 10.1111/mec.13674.

Novembre, J. and Stephens, M. (2008) Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*. 40(5), pp. 646-649. doi: 10.1038/ng.139.

Pfeifer, B. *et al.* (2014) PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Molecular Biology Evoutionl*. 31(7), pp. 1929-1936. doi: 10.1093/molbev/msu136.

Phillips, M.K. and Parker, W.T. (1988) Red Wolf Recovery: A Progress Report. *Conservation Biology*. 2(2), pp. 139-141. Available at: https://www.jstor.org/stable/2386098 (accessed 21 January 2020).

Pongratz, N., Gerace, L. and Michiels, N.K. (2002) Genetic differentiation within and between populations of a hermaphroditic freshwater planarian. *Nature*. 89, pp. 64-69. doi: 10.1038/sj.hdy.6800102.

Przeworski, M. (2009) A Genealogical Interpretation of Principal Components Analysis. *PLoS Genetics*. 5(10). doi: 10.1371/journal.pgen.1000686.

Purcell, S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 81(3), pp. 559-575. doi: 10.1086/519795.

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Available at: http://www.R-project.org/ (accessed XX).

Reed, D.H. and Frankham, R. (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*. 17(1), pp. 230-237. Available at: https://www.jstor.org/stable/3095289 (accessed: 9 January 2020).

Samuels, D.C. et al. (2016) Heterozygosity Ratio, a Robust Global Genomic Measure of Autozygosity and Its Association with Height and Disease Risk. Genetics. 204(3), pp. 893-904. doi: 10.1534/genetics.116.189936.

Servheen, C., Kasworm, W.F. and Thier, T.J. (1995) Transplanting grizzly bears Ursus arctos horribilis as a management tool – results from the cabinet mountains, Montana, USA. Biological Conservation. 71(3), pp. 261-268. doi: 10.1016/0006-3207(94)00035-O.

Shafer *et al.* (2015) Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution.* 30(2). doi: 10.1016/j.tree.2014.11.009.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science*. 236(4803), pp. 787-792. doi: 10.1126/science.3576198.

Smith, K.G. and Clark, J.D. (1994) Black bears in Arkansas: characteristics of a successful translocation. *Journal of Mammalogy*. 75(2), pp. 309-320. doi: 10.2307/1382549.

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

Ulvund, K. *et al.* (2018) *Fjellrev i Norge 2018. Resultater fra det nasjonale overvåkingsprogrammet for fjellrev.* (NINA Rapport; 1603). Available at: http://hdl.handle.net/11250/2577860 (accessed 25 August 2019).

Ulvund, K. and Wallén, J. (2018) Overvåkning av fjellrev 2018 – revidert utgave/Inventering av fjällräv 2018 – reviderad utgåva. Bestandsstatus for fjellrev i Skandinavia/Beståndsstatus för fjällräv i Skandinavien 1-2018. Norsk institutt for naturforskning (NINA) og/och Naturhistoriska riksmuseet (NRM).

Wade, M.J. and McCauley, D.E. (1988) Extinction and Recolonization: Their Effects on the Genetic Differentiation of Local Populaitons. 42(5), pp. 995-1005. Available at: https://www.jstor.org/stable/2408915 (accessed 27 January 2020).

Weir, B.S. and Cockerham, C.C. (1984) Estimating F-Statistics for the analysis of population structure. *Evolution*. 38(6), pp. 1358-1370. doi: 10.1111/j.1558-5646.1984.tb05657.x.

Wickham, H. *et al.* (2019) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Available at: https://CRAN.R-project.org/package=ggplot2 (accessed 19 January 2019).

Wright, S. (1949) The genetical structure of populations. Ann. Eugen. 15(1), pp. 323-354. doi: 10.1111/j.1469-1809.1949.tb02451.x.

Wright, S. (1965) The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution*. 19(3), pp. 395-420. doi: 10.2307/2406450.

Appendix

I: Supplementary Tables

Table A1. Table of Pairwise F_{ST} estimated by using the nucleotide method implemented in PopGenome and the Weir and Cockerham (1984) method for both early and late period. See Appendix, Figure A1 for graphic representation of the relationship between the two estimators.

Pair of subpopulation	Early period		Late period	
	F_{ST} (nucleotide)	**F_{ST} (Weir and Cockerham)**	**F_{ST} (nucleotide)**	**F_{ST} (Weir and Cockerham)**
Hardangervidda/Saltfjellet	0.0146	0.0177	0.0080	0.0106
Hardangervidda/Snøhetta	0.0235	0.0259	0.0266	0.0271
Hardangervidda/Borgafjäll	0.0416	0.0435	0.0747	0.0736
Hardangervidda/Helags	0.1143	0.1156	0.0750	0.0734
Hardangervidda/Vindelfjällen	0.0542	0.0562	0.0388	0.0399
Saltfjellet/Snøhetta	0.0168	0.0227	0.0292	0.0319
Saltfjellet/Borgafjäll	0.0625	0.0708	0.0837	0.0872
Saltfjellet/Helags	0.1494	0.1571	0.0793	0.0870
Saltfjellet/Vindelfjällen	0.0630	0.0674	0.0315	0.0334
Snøhetta/Borgafjäll	0.0406	0.0459	0.0880	0.0870
Snøhetta/Helags	0.1558	0.1563	0.1020	0.1009
Snøhetta/Vindelfjällen	0.0781	0.0826	0.0485	0.0497
Borgafjäll/Helags	0.1331	0.1399	0.1287	0.1285
Borgafjäll/Vindelfjällen	0.0130	0.0167	0.0422	0.0427
Helags/Vindelfjällen	0.1286	0.1332	0.1026	0.1041

Table A2. Summary information for arctic fox scaffolds (n = 3) used for sliding window analyses. Given is the scaffold name, the length of the scaffold in base pairs and the number of SNPs on each scaffold. The final column represents the figures which shows the F_{ST} analysis for that scaffold.

Scaffold name	Length (bp)	# SNPs	Figure
0	60512252	2088	4(a), A3(a, b), A4
3	47483000	2178	4(b), A3(c, d), A5
11	33748078	1342	5, A3(e), A6

Table A3. The six first principal components (PC) from the principal component analysis. Given for each PC is its eigenvalues, how many percent of the variance it explains (second column) and how many percent of the variation is explained by PC and lower PS combined (third column).

	Eigenvalues	Variance.percent	Cumulative.variance.percent
Dim.1	1000.170	7.897	7.897
Dim.2	666.327	5.261	13.158
Dim.3	531.849	4.199	17.357
Dim.4	397.038	3.135	20.492
Dim.5	325.740	2.572	23.492
Dim.6	291.513	2.302	25.365

II: Supplementary Figures

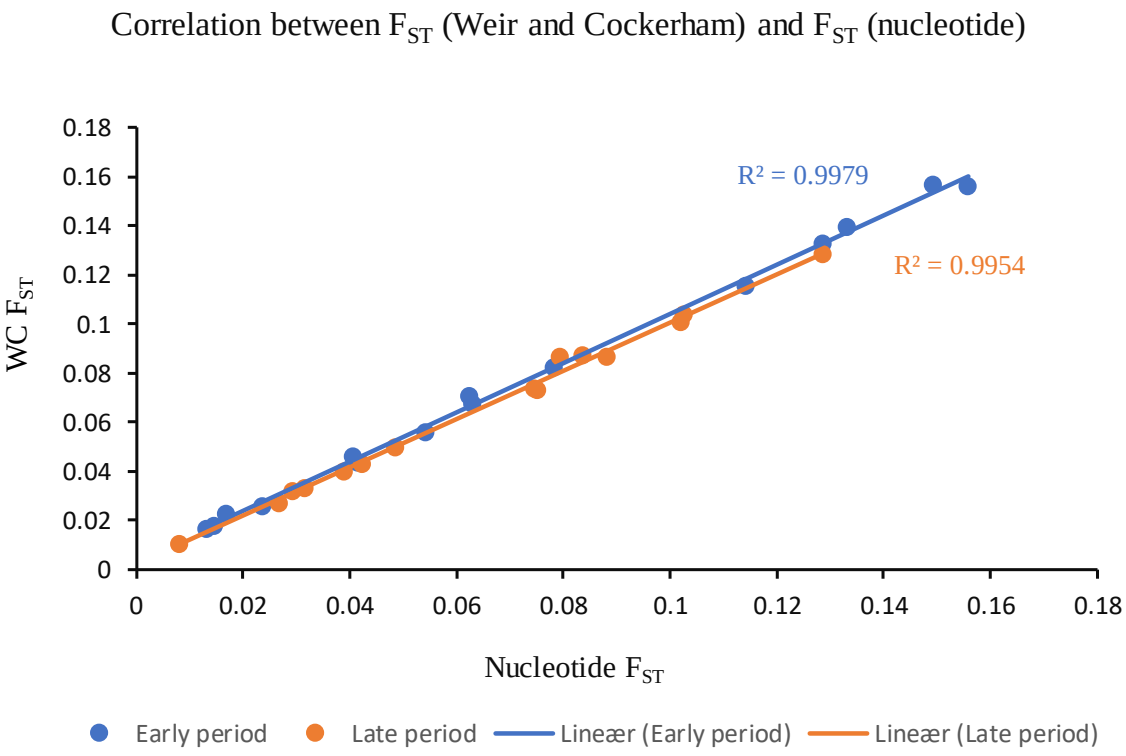


Figure A1. Relationship between two estimators of F_{ST}. Nucleotide estimator implemented in the R package PopGenome on the x-axis and Weir and Cockerham (1984) on the y-axis. For early period $R^2 = 0.9979$ and for late period $R^2 = 0.9954$. See Appendix, Table A1 for estimated pairwise F_{ST}-values.

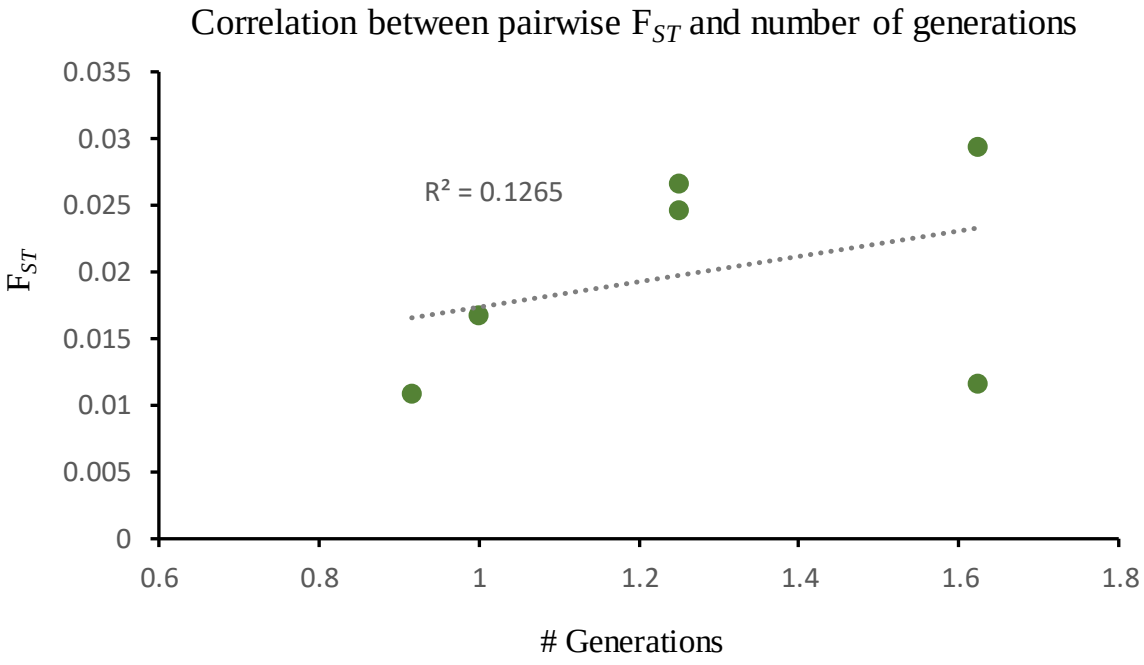


Figure A2. Relationship between F_{ST} and number of generations. The x-axis represents number of generations that has passed between early and late sampling period and the y-axis represents the pairwise F_{ST} within each subpopulation. R^2 = 0.1265. Spearman correlation: 0.3557 = ROT(0.1265).

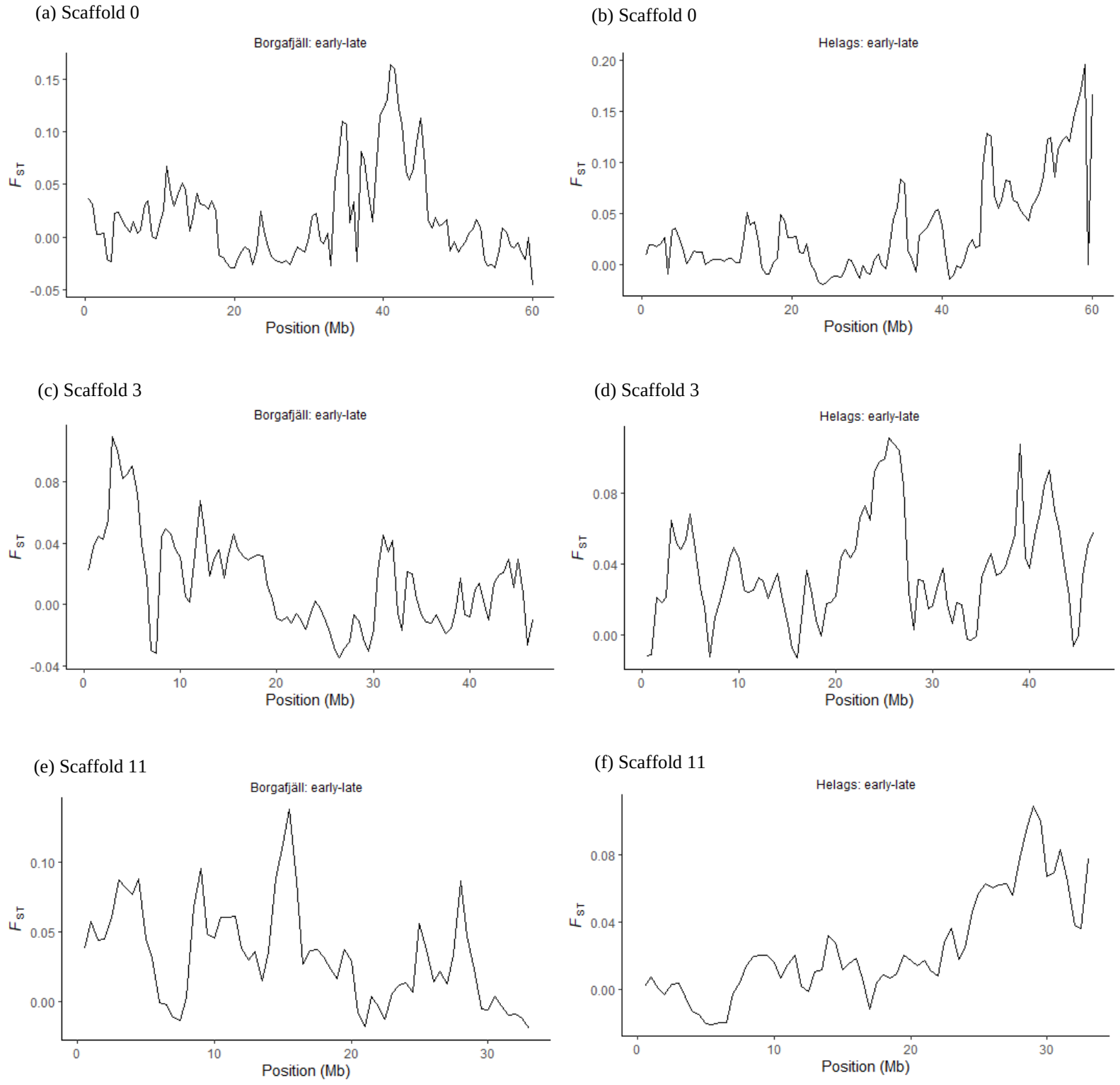
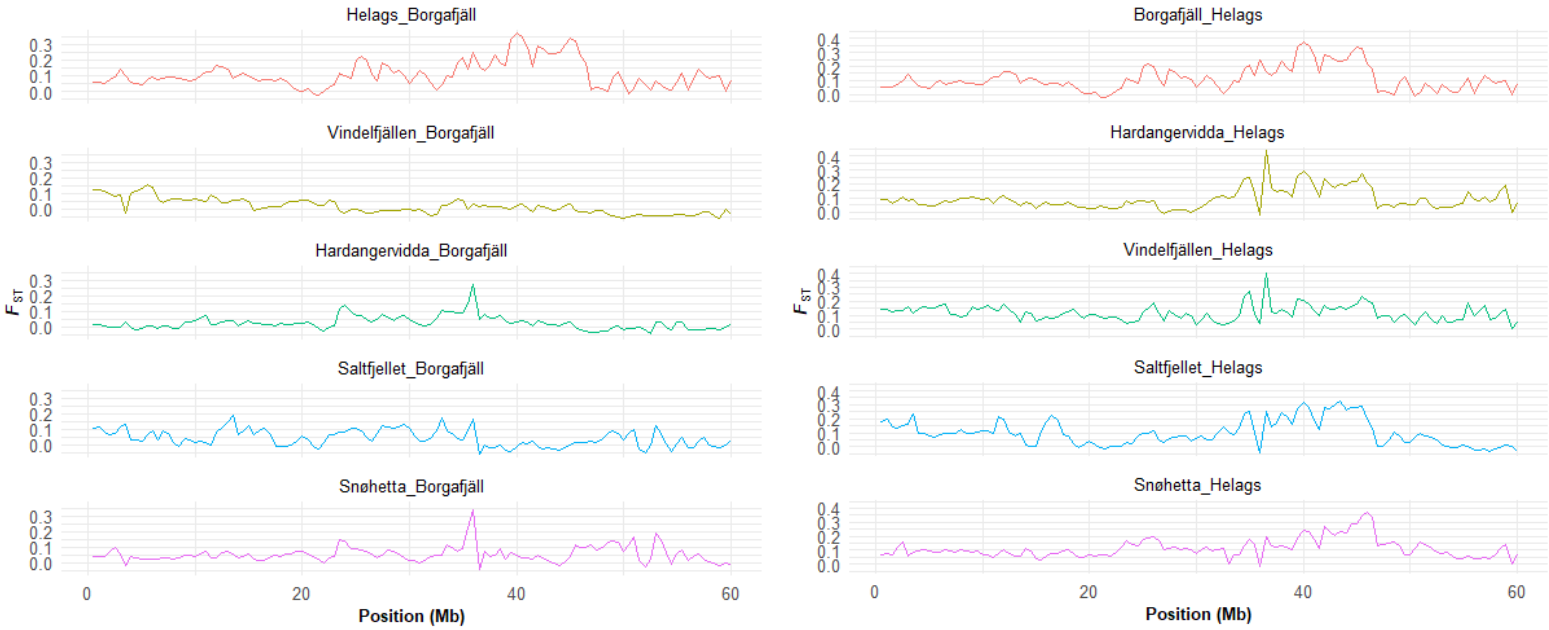


Figure A3. Plots of F_{ST} calculated within 1000-kb sliding window for the two biggest scaffolds (0 and 3) and the scaffold containing gene(s) coding for fur color (11) in the arctic fox genome. The analysis is performed by comparing early and late period for Borgafjäll in a, c and e and for Helags in b, d and f. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values.

(a) Scaffold 0 – early period

(b) Scaffold 0 – early period



(c) Scaffold 0 – late period

(d) Scaffold 0 – late period

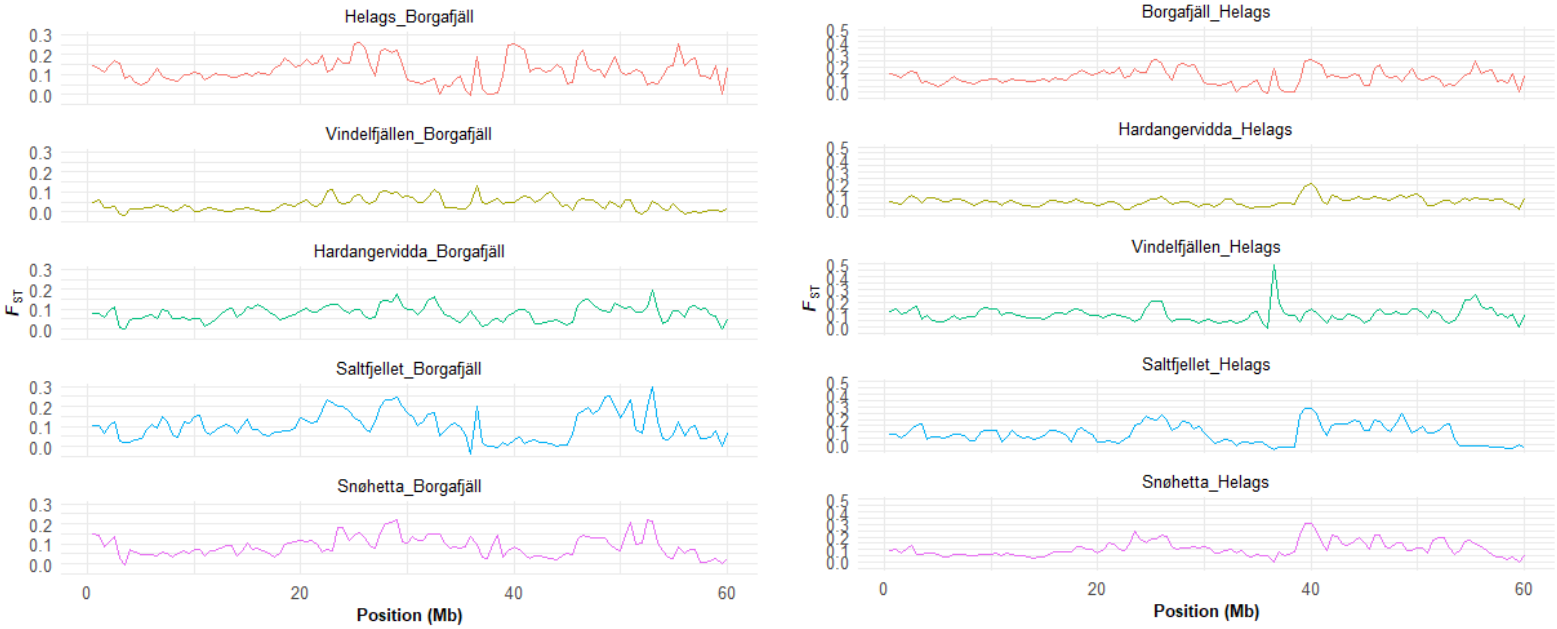
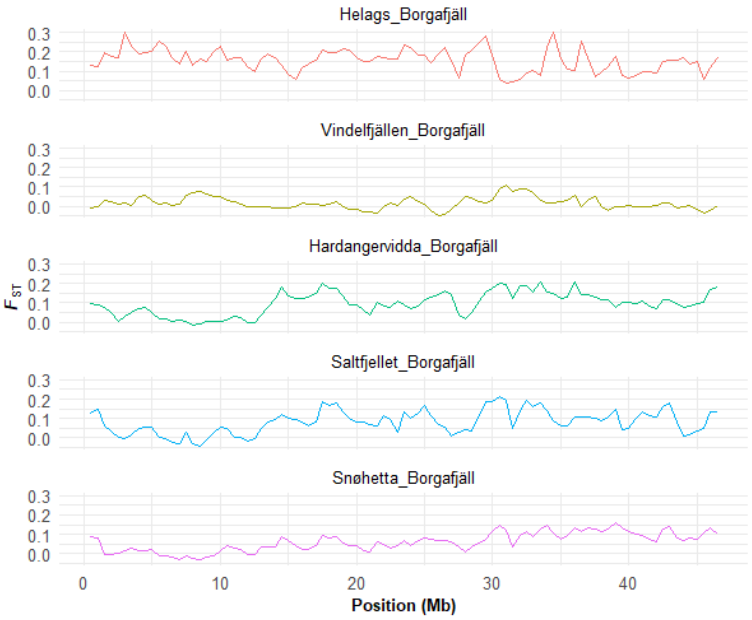
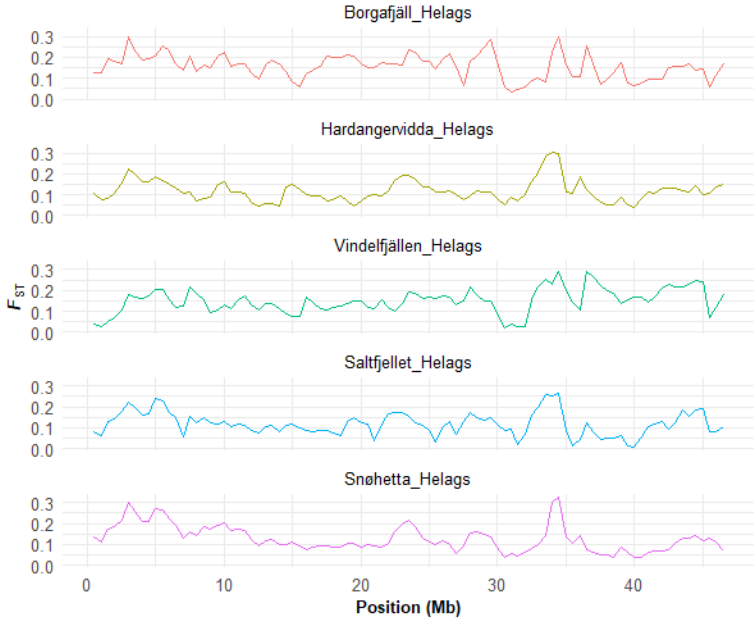


Figure A4. Plots of F_{ST} calculated within 1000-kb sliding window for one of the biggest scaffolds (0) in the arctic fox genome. (a) and (b) represents the early time period in the population and (c) and (d) represents the late time period in the populations. In (a) and (c) Borgafjäll is compared with each of the other populations. In (b) and (d) Helags is compared with each of the other populations. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values within the sliding windows.

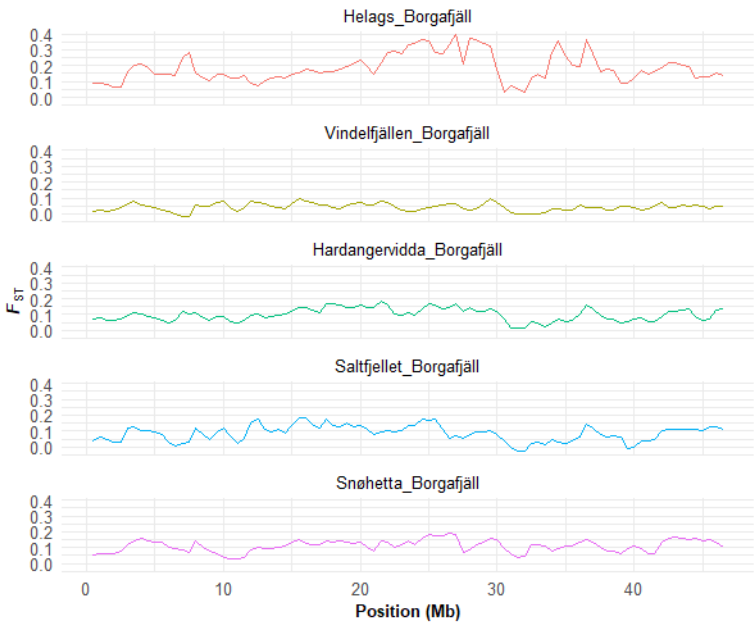
(a) Scaffold 3 – early period



(b) Scaffold 3 – early period



(c) Scaffold 3 – late period



(d) Scaffold 3 – late period

Figure A5. Plots of F_{ST} calculated within 1000-kb sliding window for one of the biggest scaffolds (3) in the arctic fox genome. (a) and (b) represents the early time period in the populations and (c) and (d) represents the late time period in the population. In (a) and (c) Borgafjäll is compared with each of the other populations. In (b) anc (c) Helags is compared with each of the other populations. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values within the sliding windows.

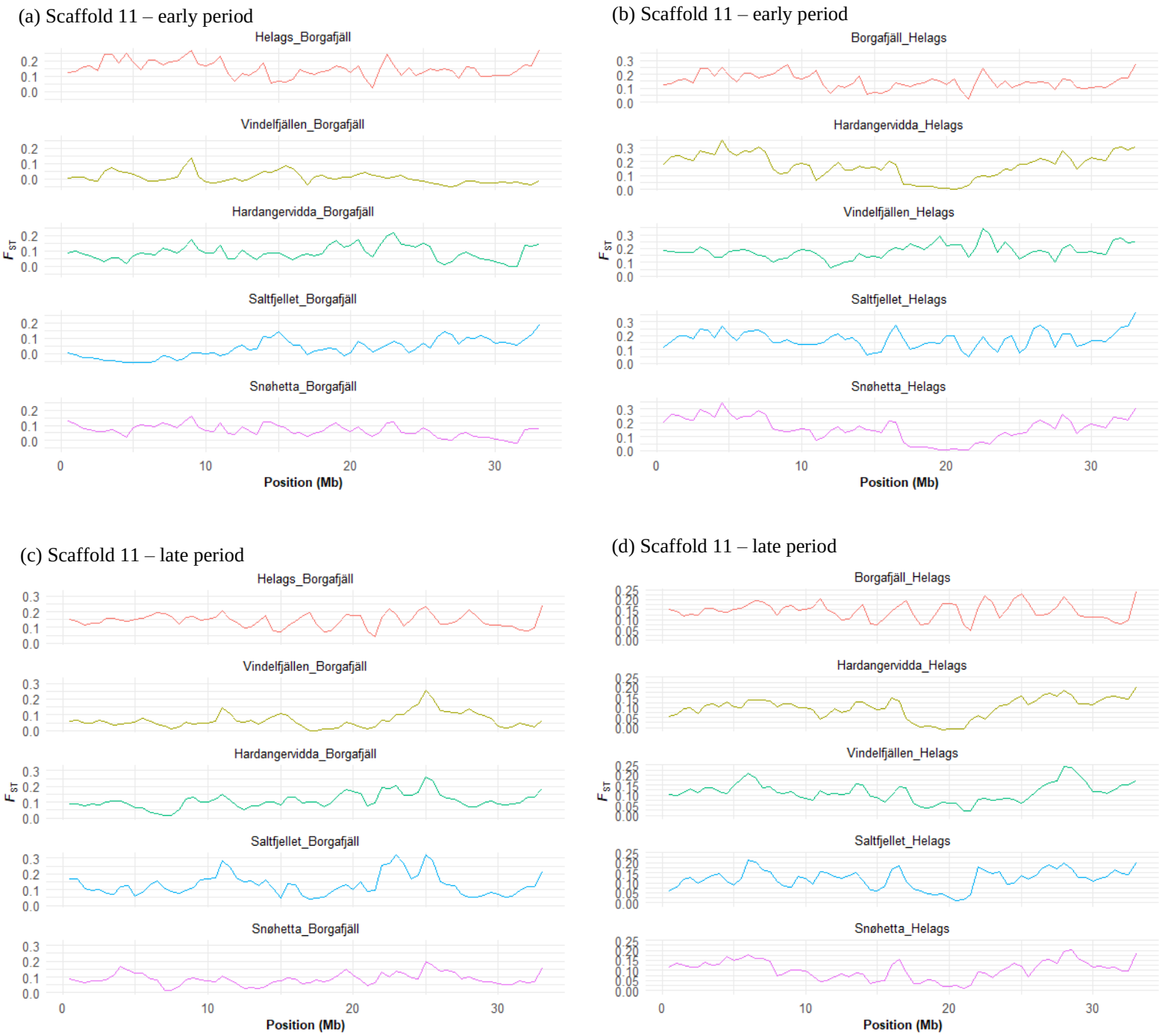


Figure A6. Plots of F_{ST} calculated within 1000-kb sliding window for the scaffold containing gene(s) coding for fur color (11) in the arctic fox genome. (a) and (b) represents the early time period in the populations and (c) and (d) represents the late time period in the populations. In (a) and (c) Borgafjäll is compared with each of the other populations. In (b) and (d) Helags is compared with each of the other populations. The x-axis represents the SNPs position in Mb and the y-axis represents the average F_{ST}-values within the sliding windows.

(a) (b)

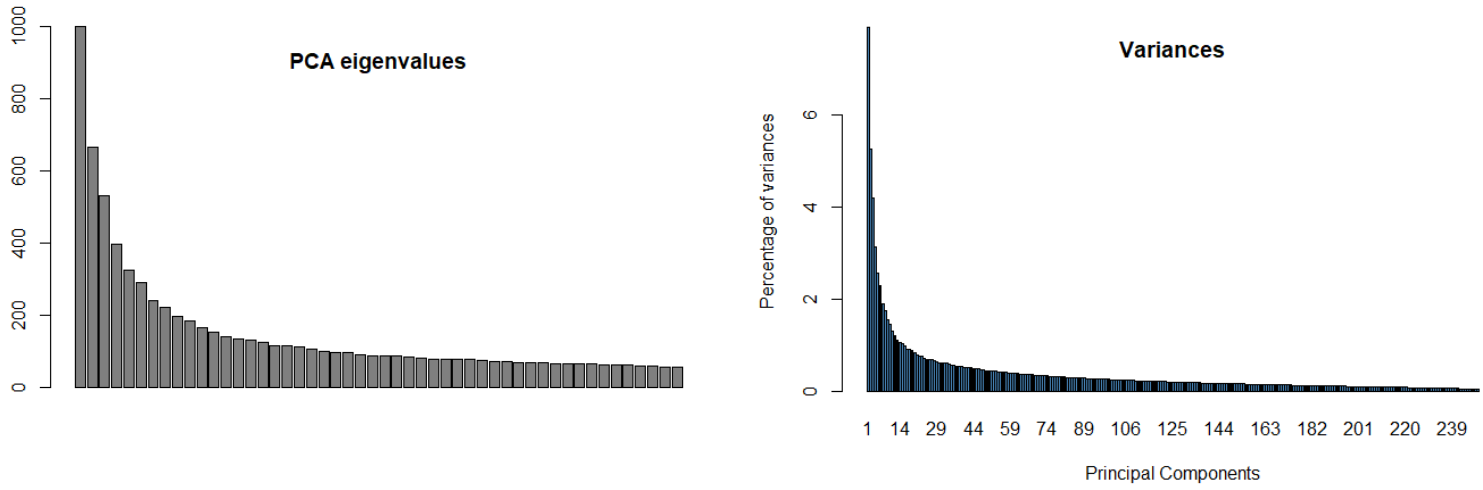
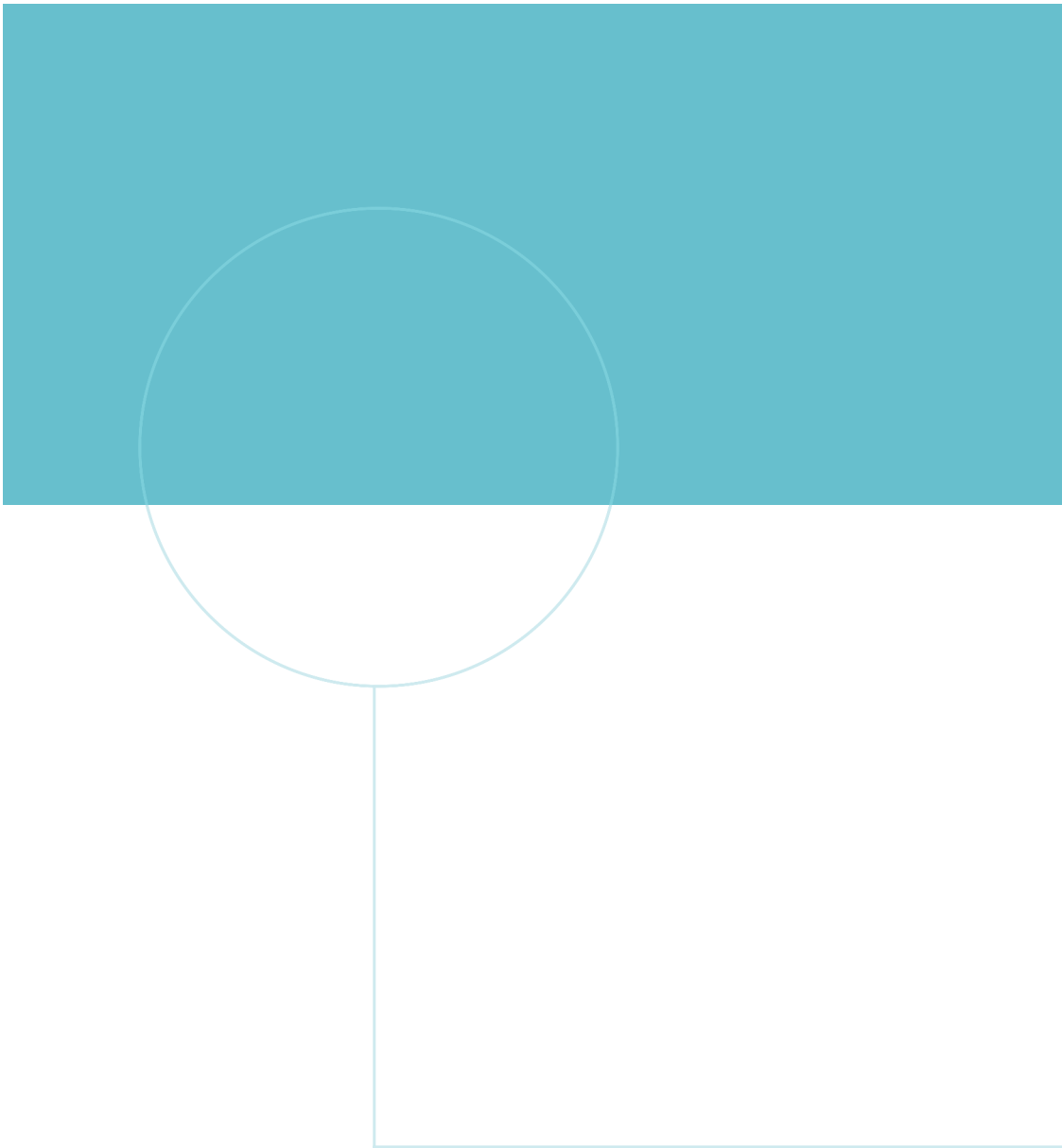


Figure A7. (a) Bar plot showing the eigenvalues of the first 50 principal components from the principal component analysis (PCA) and (b) showing how many percent of the variation is explained by each principal component.



NTNU

Norwegian University of
Science and Technology