Master's thesis

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

Lukas Tietgen

Fur colour in the arctic fox – Genetic architecture and consequences for fitness

Master's thesis in Biology
Supervisor: Henrik Jensen (IBI)
Co-supervisors (NINA): Ingerid J. H. Arnesen, Nina E. Eide, Øystein Flagstad and Arild Landa

June 2020

NTNU
Kunnskap for en bedre verden

Lukas Tietgen

Fur colour in the arctic fox – Genetic architecture and consequences for fitness

Master's thesis in Biology
Supervisor: Henrik Jensen (IBI)
Co-supervisors (NINA): Ingerid J. H. Arnesen, Nina E. Eide, Øystein Flagstad and Arild Landa,
June 2020

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

NTNU
Kunnskap for en bedre verden

Abstract

Genome-wide association studies (GWAS) are powerful, yet somewhat underused in wild populations. Previous studies have often failed to identify major effect genes underlying adaptive traits. In this study, I used a large SNP dataset from a wild population of a non-model species to investigate the genetic architecture underlying arctic fox *Vulpes lagopus* fur colour. Using this whole-genome approach, I was able to identify the MC1R gene as the sole causal gene for arctic fox fur colour. Further, through measures of fitness that link ecological and evolutionary processes, I documented selection on arctic fox fur colour. This analysis was made possible by the large encounter dataset of the Arctic Fox Monitoring Programme in Norway that allowed the reconstruction of life history for arctic fox individuals for over a decade.

I showed that female foxes that are heterozygous at the fur colour locus have higher mean fitness than homozygous individuals which is explained through both increased survival and reproduction. The effect was found to be mostly independent of examined environmental variables that may affect the arctic fox. MC1R is located in a genomic region that is densely covered with genes and many of these are very likely to covary with MC1R, showing the potential for indirect selection. The pleiotropic effects of MC1R in the melanocortin system are also a possible foundation for indirect selection. Some of the traits in linkage with MC1R and identified as potential candidates behind indirect selection are lipid metabolism and developmental processes such as the Wnt signalling pathway. Foxes heterozygous at the fur colour locus showed also a higher degree of genome-wide heterozygosity which might contribute to their overall fitness advantage through heterozygote advantage. Support for direct selection on arctic fox fur colour based on camouflage or thermoregulation is weak. However, more knowledge about differences other than fur colour between the arctic fox fur colour morphs is needed to disentangle the potential effects of indirect selection.

My study provides evidence that whole-genome analyses can be successfully applied to wild non-model species and help to identify major effect genes underlying adaptive traits. Furthermore, I have shown how this approach can be used to identify knowledge gaps that future research can investigate.

Keywords

Adaptive trait, Captive breeding, Endangered species, Fitness estimation, GWAS, Indirect selection, Linkage disequilibrium, Pleiotropy, Wild population

Sammendrag

‘Genome-wide association studies’ (GWAS) er nyttige verktøy, men er likevel ikke veldig mye brukt til å studere ville populasjoner. Tidligere studier har ofte ikke klart å identifisere gener som ligger til grunn for adaptive trekk. I dette studiet har jeg brukt et stort SNP-datasett fra en vill populasjon av en ikke-modell-organisme for å undersøke den genetiske arkitekturen som styrer fjellrevens *Vulpes lagopus* pelsfarge. Ved å bruke denne tilnærmingen som ser på hele genomet, klarte jeg å identifisere MC1R-genet som det eneste genet som bestemmer pelsfarge hos fjellreven. Gjennom målinger av fitness som kobler sammen økologiske og evolusjonære prosesser, har jeg dokumentert seleksjon på pelsfarge i fjellrev. Med den store mengden aktivitetsdata for fjellrev som har blitt samlet inn av det norske overvåkningsprogrammet, kunne jeg rekonstruere livshistorien til fjellrevindivider fra over et tiår tilbake.

Jeg har vist at tisper som er heterozygote på pelsfarge-lokuset har høyere gjennomsnittlig fitness enn homozygoter, og dette forklares av både økt overlevelse og reproduksjon. Denne effekten var for det meste upåvirket av de undersøkte miljøvariablene som kan påvirke fjellreven. MC1R ligger plassert på et sted i genomet som inneholder mange gener, og mange av disse kovarierer med stor sannsynlighet med MC1R. Dette viser et potensiale for indirekte seleksjon. De pleiotropiske effektene MC1R har i melanokortin-systemet kan også være et grunnlag for indirekte seleksjon. Lipidmetabolisme og utviklingsprosesser som Wnt-signalisering er noen av egenskapene til genene som er koblet til MC1R og som er identifisert som potensielle kandidater for indirekte seleksjon. Heterozygote rever hadde høyere heterozygositet i genomet. Dette bidrar potensielt til en generell fitness-fordel i form av ‘heterozygot-fordelen’.

Jeg fant ikke bevis for at det skjer direkte seleksjon på pelsfarge basert på kamuflasje eller termoregulering. Det kan likevel være andre forskjeller enn pelsfarge mellom de ulike pelsfargevariantene, men for å linke dette til indirekte seleksjon, trengs det mer kunnskap om forskjeller i for eksempel atferd eller fysiologi mellom pelsfargevariantene.

Min studie viser at GWAS-analyser kan brukes i ville populasjoner med ikke-modell-organismer for å identifisere gener bak adaptive trekk. I tillegg har jeg vist hvordan denne tilnærmingen kan brukes for å oppdage kunnskapshull som kan utforskes i fremtidige studier.

Acknowledgements

I would like to thank all my supervisors, Henrik Jensen, Ingerid J.H. Arnesen, Nina E. Eide, Øystein Flagstad and Arild Landa for allowing me to participate in the exciting field of arctic fox research and all the support and guidance they have given me during the last two years.

I want to thank everyone involved in collecting field data and genetic samples, this includes many volunteers of whom I was lucky enough to meet during the fieldwork and a lot of people working for SNO. Further, I would like to thank Stefan Blumentrath for extracting snow data, Oddmund Kleven for facilitating the Fluidigm genotyping, Alina Niskanen for help with genetic parentage analyses, and Sarah Lundregan and Martin Kuiper for fruitful discussions about GWAS, BLAST and GO analyses, as well as Cecilia Di Bernardi for sharing insights into her ongoing arctic fox research. I also want to thank Thomas Kvalnes for being patient enough to answer many of my questions regarding modelling fitness and statistical issues.

I want to thank everyone that I have shared my everyday life with during my studies. Finishing this thesis ends six years of Trondheim life and it would not have been the same without the many people I have met. I want to thank my family for trying to support me as best as they could despite being over 1000 km away. I want to thank the Thorseth Grevskott family for opening their home for me when the Covid-19 pandemic hit. Lastly, I need to thank Ragnhild for being an incredible support and keeping me sane through stressing times!

The Norwegian Captive Breeding Programme (19087015) and The Arctic Fox Monitoring Programme in Norway (contract 18087019) contributed a significant amount of data to this study all funded by the Norwegian Environmental Agency. This study was part of ECOFUNC funded by the Research Council of Norway (grant 244557/E50). The work was also partly supported by the Research Council of Norway through its Centres of Excellence funding scheme (grant 223257).

Lukas Tietgen,

Lommedalen, juni 2020

TABLE OF CONTENTS

INTRODUCTION

Ever since Charles Darwin enlightened the world with the concept of natural selection, and the term *fitness* was coined by Spencer (1896) and Fisher (1930), biologists have been interested in understanding how individuals differ in terms of survival or reproductive success. For a long time, researches were limited to observing how differences in phenotypic traits made one individual more successful, more fit, than its conspecifics. Recent advances in molecular analytical methodologies and the increased availability of genomic data have, however, allowed us to connect variation in these phenotypic traits directly to the third condition for evolution: their causal genes (Stinchcombe & Hoekstra, 2008; Andrew et al., 2013; Ellegren, 2014).

A commonly used method for mapping genes for phenotypic traits is to conduct a genome-wide association study (GWAS) (Bush & Moore, 2012). While being widely used for mapping the genes of human diseases (Visscher et al., 2017), the use of GWA studies in natural animal populations is still somewhat limited (Jensen et al., 2014; Santure & Garant, 2018). Aside from methodological issues (e.g. reproducibility of associations, density of genetic markers, relatedness, and sample size (Korte & Farlow, 2013; Jensen et al., 2014; Santure & Garant, 2018)), most studies that map genes underlying fitness-related traits, find that these traits are polygenic and thus struggle to detect significant associations between single genetic markers and the trait in question (Hecht et al., 2013; Robinson et al., 2013; Santure et al., 2013; Comeault et al., 2014; Santure et al., 2015; Wenzel et al., 2015; Kardos et al., 2016; Silva et al., 2017; Lundregan et al., 2018). However, several studies have shown in the past that GWAS is capable of identifying single genes or genomic regions underlying fitness-related traits in wild populations. Johnston et al. (2011) found the gene underlying polymorphism for horn morphology, an important fitness-related trait, in wild Soay sheep *Ovis aries*. Likewise, Barson et al. (2015) discovered a large effect locus explaining variation in sea age at maturity, a highly variable and fitness-related trait in Atlantic salmon *Salmo salar*. Recently, the impact of that locus was supported, however, alongside loci on other chromosomes, thus showing a polygenic basis after all (Sinclair-Waters et al., 2020). The adaptive significance of beak morphology in the different Darwin's ground finches *Geospiza* is well known. Lawson and Petren (2017) used GWAS to examine the genetic architecture and found a region of interest on chromosome 1A.

Colouration may be one of the most conspicuous phenotypic traits in animals and has been the subject of research for many decades, if not centuries (Poulton, 1890; Beddard, 1892). Animal

colouration can have many different purposes (Caro, 2005) and implications of colouration on fitness have been shown in a wide range of animal species (Roulin et al., 2001; Nachman et al., 2003; Brommer et al., 2005; Gratten et al., 2008; van den Brink et al., 2011; Roff & Fairbairn, 2013; Zimova et al., 2014; Atmeh et al., 2018; Zimova et al., 2018). Since colouration is such a conspicuous trait, it is appealing to solely account for differences in fitness to the colour phenotype. However, it is important to keep in mind that there might be more to a trait than the phenotype itself. Hadfield et al. (2007) showed that colour phenotypes do not always coincide with genetic patterns. Additionally, an association between colouration and other phenotypic traits, such as sexual behaviour, aggressiveness, stress response and energy homeostasis has been shown in different species (Ducrest et al., 2008; San-Jose & Roulin, 2017). Such covariation clearly raises the question of how theory can predict evolutionary consequences of selection on a phenotypic trait when the gene underlying the trait is linked to or affects other phenotypic traits that themselves could affect fitness. This issue demonstrates the importance of more studies aimed at gaining insight into the genetic architecture behind adaptive traits.

The arctic fox *Vulpes lagopus* is a species with interesting colouration features. The species occurs in multiple distinct fur colour morphs and undergoes seasonal moult (Chesemore, 1970; Audet et al., 2002). The two common colour morphs are described as the white and the blue morph. White arctic foxes have a completely white winter fur, whereas their summer fur is mostly brown with lighter ventral sides. The blue morph is uniformly dark brown or charcoal year-round, with a lighter colouration during winter. Fur colour in arctic foxes appears to be inherited by a simple Mendelian trait at one autosomal locus, where the blue morph is the result of the dominant allele (Slagsvold, 1949; Våge et al., 2005). Despite this simple genetic basis suggested for fur colour morphs, the white colour morph makes up over 90% of the global arctic fox population (Norwegian Polar Institute). However, the relative frequency of the two morphs varies across the species distribution (Barabash-Nikiforov, 1938; Braestrup, 1941; Fetherston, 1947; Chesemore, 1968; Hersteinsson, 1989) and even across different environments, as shown in Iceland (Hersteinsson, 1989). In Iceland, the observed differences in frequencies are thought to reflect distinct selection advantages of the two colour morphs in different habitats (Hersteinsson, 1989). The exact mechanisms underlying the global distribution of arctic fox fur colour morphs are despite this not well studied and understood.

Molecular analysis by Våge et al. (2005) showed that two cysteine amino acid substitutions within the region of the Melanocortin-1-receptor gene (MC1R) co-segregated with the arctic

fox fur colour morphs. MC1R is known to regulate melanin-based colouration in a wide range of animal species (Robbins et al., 1993; Våge et al., 1997; Nachman et al., 2003; Schmutz & Berryere, 2007; Roulin & Ducrest, 2013; San-Jose et al., 2015). It is thus not surprising that MC1R is involved in arctic fox fur colouration. However, the study design of Våge et al. (2005) was focused on MC1R and would not have been able to detect other genes contributing to the colour morphs. MC1R is part of a gene complex where five melanocortin receptors (MC1R-MC5R) share the same melanocortin ligands (Ducrest et al., 2008). Pleiotropic co-variation between melanin-based colouration and traits governed by MC2R-MC5R can thus be expected and is, in fact, found in different species (Ducrest et al., 2008).

In this study, I first examined the genetic basis and architecture of fur colour in a wild population of arctic fox. Second, I quantified selection on fur colour genotypes using measures of fitness that link ecological and evolutionary processes. Third, I investigated how frequencies of fur colour alleles and genotypes varied spatially and temporally in the wild population. Finally, I investigated the potential for indirect phenotypic effects of fur colour genes through pleiotropy or physical linkage with other genes, and how these effects could affect the observed patterns of fitness and genotype frequencies.

METHODS

Unless otherwise stated, all analyses were performed in statistical software R 3.6.1 (R Core Team, 2019).

THE ARCTIC FOX IN FENNOSCANDIA

In the early 20th century, the Fennoscandian arctic fox populations were close to extinction, likely due to large hunting pressure (Hersteinsson et al., 1989; Linnell et al., 1999b). Over 12 years, the arctic fox was protected in Sweden (1928), Norway (1930) and Finland (1940) (Linnell et al., 1999a). Despite protection, Fennoscandian populations did not recover during the following years (Hersteinsson et al., 1989; Linnell et al., 1999a).

In Norway, a National Monitoring Programme has been in place since 2003 and will be continued until at least 2021 (Andersen et al., 2003; Direktoratet for naturforvaltning, 2003; Eide et al., 2017). Additionally, The Norwegian Arctic Fox Captive Breeding Program was initiated in 2000 to re-establish extinct natural populations and support extant ones. The first pups were released in 2006 (Landa et al., 2017). Also, in Sweden, extensive monitoring is undertaken. Census sizes have increased in both countries. The minimum population estimate

of adult foxes in Norway is 221 for the period 2017-2019 (Ulvund et al., 2019). In Sweden, the minimum population was estimated to be 144 adult individuals in 2019 (Wallén et al., 2020). Despite this increase, the arctic fox is listed as critically endangered on the Norwegian red list of threatened species (Henriksen & Hilmo, 2015) and endangered on the Swedish red list of threatened species (ArtDatabanken, 2015). In Finland, no breeding arctic foxes have been observed since 1996, and arctic foxes only occur sporadically in the border areas (Angerbjörn et al., 2013).

DATA COLLECTION

Arctic fox data used in this study originate from the described monitoring and conservation programmes, collected in the period 2007-2019. Foxes born in the breeding station at Sæterfjellet, Oppdal, are released into the wild early the following year. These foxes are ear-tagged with unique colour codes and marked with HPT12 Biomark© PIT-tags for later identification. Before release, ear tissue is sampled for DNA analysis. Wild born arctic foxes are monitored through surveys at known den localities. Main periods for annual surveys are during the late winter (1st of March – 15th of May) and summer (20th of June – 15th of August) (Ulvund et al., 2019). During summer surveys, wild-born pups are trapped and marked with PIT-tags, in three of the subpopulations (Hardangervidda, Snøhetta and Varangerhalvøya, Figure 1). The fur colour phenotype is registered and ear tissue for DNA analysis is sampled. During den surveys in both summer and winter, non-invasive

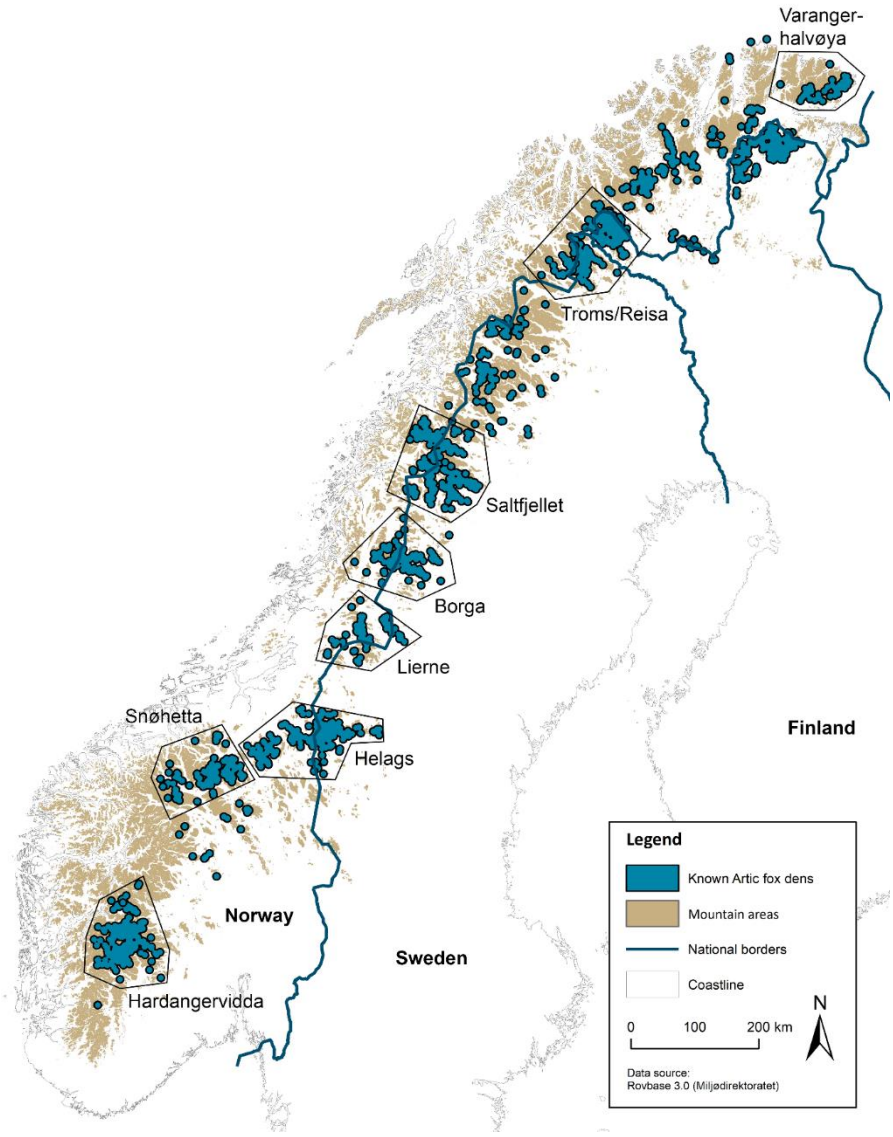


Figure 1: Overview over known historic arctic fox den sites in Fennoscandia. Polygons mark the subpopulations used in this study.

sampling of genetic material (scats and hair) is conducted to identify individuals and estimate population size (Ulvund et al., 2019). Additionally, wildlife cameras and Biomark chip readers at feeding stations allow for year-round monitoring and identification of marked individuals. See Landa et al. (2017) for more details regarding methods and data collection of the captive breeding and release programme.

Affymetrix genotyping

701 arctic fox individuals were successfully genotyped using a custom Affymetrix Axiom 702k SNP-array with 507 000 arctic fox specific single nucleotide polymorphisms (SNPs). Samples constituted ear tissue and DNA from these were extracted using the Qiagen DNeasy 96 Blood & Tissue Kit (Qiagen GmbH, Hilden, Germany). More details on the array's design are given in Hagen et al. (in prep.). Only somatic SNPs classified as poly high resolution (Bassil et al., 2015) were kept for this study's analyses (361 289 SNPs). The SNP positions were obtained from an arctic fox reference genome assembly which consists of 4 048 scaffolds with SNP positions given within every scaffold (von Seth et al., in prep.).

Data quality control

The software PLINK 1.90 (Chang et al., 2015; Purcell & Chang, 2019) was used to check the Affymetrix SNP-genotyping data set for Mendelian errors based on a microsatellite pedigree. 1 632 SNPs with more than 10% error rate and 12 fox individuals with more than 5% error rate were excluded in the process. 448 SNPs were discarded due to low minor allele frequency (MAF < 0.01). Additional data quality control was performed using the *check.marker* function in the GenABEL R package (Aulchenko et al., 2007). Here, no SNPs were excluded due to low call rate (<95%). Neither were individuals excluded due to low call rate (<95%) or extreme high heterozygosity (FDR<1%). Eight genotyped individuals that lacked fur colour phenotype were excluded. After quality control, the data set consisted of 359 218 autosomal SNPs and 681 arctic fox individuals (562 white, 119 blue).

Genome-wide Association Study

A genome-wide association (GWA) analysis was used to investigate associations between autosomal genetic markers (SNPs) and the arctic fox fur colour morphs. Preparation of the data and the GWA analysis itself were done using the GenABEL R package (Aulchenko et al., 2007). A genomic relatedness matrix (GRM) based on identity of state (IBS) was included in the model to account for relatedness. Unlinked markers (SNPs in linkage equilibrium) should be used for most accurate relatedness estimates (Santure et al., 2010). Thus, the *indep* function

of PLINK 1.90 (Chang et al., 2015; Purcell & Chang, 2019) was used to create a subset of 40 539 unlinked SNPs prior to the GRM calculation. Recommended parameters 50 5 2 were used.

In the GWA analysis, a polygenic model including the full GRM was fitted and in conjunction, a mixed model approximation was performed to test for association between arctic fox fur colour and the genetic markers included in the study. Due to genomic inflation ($\lambda = 1.92$) p-values were corrected for lambda. As a GWA analysis runs a single test for each SNP, multiple testing is an issue (Bush & Moore, 2012). The significance threshold α was therefore adjusted using the conservative Bonferroni method where 0.05 was divided by the number of SNPs included in the analysis (Bonferroni, 1936; Johnson et al., 2010).

Fluidigm genotyping

For facilitation of non-invasive genetic monitoring of the arctic fox population in Norway, a microfluidic SNP-array was recently developed (Kleven et al., in prep.). The Fluidigm SNP-array included 87 autosomal markers, including the SNP chosen to describe the arctic fox fur colour genotype (AX-176934441; see Results section). DNA genotyped on this platform was extracted from hair, scat and tissue using the Maxwell Tissue Kit (Promega, Madison, WI, USA) and the Qiagen DNeasy 96 Blood & Tissue Kit (Qiagen GmbH, Hilden, Germany), following the protocols provided by the manufacturers. DNA samples were genotyped on a 96.96 Dynamic Array with integrated fluidic circuits (IFCs) using the Fluidigm EP1 instrument according to the manufacturer's protocol and scored using the Fluidigm SNP genotyping analysis software (https://www.fluidigm.com/software). Further details about the SNP-array is provided in Kleven et al. (in prep.).

912 arctic fox individuals were genotyped using the Fluidigm SNP-array. Of these, 109 were also genotyped using the Affymetrix SNP-array. The AX-176934441 genotype was identical across the two SNP-arrays in all these individuals. Of the remaining 803 individuals only genotyped using the Fluidigm platform, fur colour phenotype was known for 444 individuals (329 white, 115 blue). These individuals were used as a relatively independent dataset to verify the association between the top GWAS SNP and fur colour because they were not included in the data set used for the GWA analysis.

A detailed overview of the individuals SNP genotyped on both the Affymetrix and Fludidigm SNP array is given in Table S1.

GENE ANALYSIS

BLAST searches (Altschul et al., 1990) were performed to investigate genes located in the vicinity of SNPs that, based on the GWA analysis, were significantly associated with arctic fox fur colour. An annotated arctic fox genome is yet to be published, thus the annotated and well-studied dog genome *CanFam 3.1* (Hoeppner et al., 2014) was used as reference genome in BLAST analyses. The BLAST searches were performed using the BLAST+ 2.9.0 software (Camacho et al., 2009). For the searches, sequences of 71 base pairs (bp) were used (35 bp up- and downstream of the SNP in addition to the SNP itself). To identify the most correct BLAST hits, the e-value was required to be below 0.001 and query coverage needed to be higher than 70% (50 bp). When a SNP had multiple hits that met the requirements, the hit with the lowest e-value was chosen. SNPs that did not have a hit meeting the requirements were excluded from further analysis (n=6). To check whether using the dog genome as reference, positions of significant SNPs from the GWA analysis on the arctic fox scaffolds were compared to their respective positions in the dog genome based on the BLAST output. The SNPs occurred in the same order in both species (Figure S2) indicating a good fit between the two genomes.

SELECTION ANALYSES

Estimation of selection was made possible by the extensive encounter data on arctic foxes in Norway collected as part of the national monitoring programme. Life history data (annual survival and fecundity) was available for 1 181 individuals. All individuals were genotyped at the fur colour locus as described at either the Affymetrix or the Fluidigm SNP-array. The dataset used in the selection analyses included data from 12 years (2007-2018) and eight subpopulations (Borga, Snøhetta, Hardangervidda, Lierne, Helags, Saltfjellet, Troms/Reisa and Varangerhalvøya; Figure 1, Table S1).

Encounter data

Encounter data for the individuals was compiled from a range of sources: observation and trapping during den surveys, DNA from faeces and hair samples, Biomark© and Trovan© chip readers at feeding stations, and records from wildlife cameras. This allowed for a dataset with high-resolution encounter data for the study individuals. Arctic foxes suffer high mortality during winter (October-April) (Meijer et al., 2008). Therefore, I chose to operate with *pre-breeding census* describing the period from 1st of April to 31st of March the following year. The beginning of April also coincides roughly with the end of the mating season. Individuals alive at that time are recorded in the dataset. Individual annual survival was recorded as 1 in a

given census year for individuals that were observed after 1st of April the next year (otherwise 0).

Parentage was determined for 1 497 individuals based on individual genotype data on 85 autosomal SNPs by using the *Sequoia* R package (Huisman, 2017). This parentage approach uses birth year and SNP-genotype data to assign offspring to their genetic parents. In cases where the birth year of an individual was unknown (n=205), it was assumed that it was an adult born the previous year if the first observation of that fox was made before 1st of July. If the first observation was made after 1st of July, it was assumed to be a juvenile born the same year. This threshold was chosen to coincide roughly with the emergence of pups from the den. Default settings were used in the pedigree construction in *Sequoia*, except for genotyping rate which was set to 0.002. To obtain a pedigree as informative as possible, dummy parents (n = 158) were also assigned via sibship clustering. The resulting genetic pedigree included genetic mother for 1 400 and father for 1 392 of the 1 655 individuals (1 497 real and 158 dummy individuals) in the pedigree. Among all parent-offspring pairs in this pedigree, two SNPs had two Mendelian errors and nine SNPs had one Mendelian error. The genetic parentage was also compared to a previously constructed pedigree based on microsatellite genotype data and social information. The two pedigrees agreed for 84% of the parental links, and in the few cases with differences, the SNP-based parentage was carefully checked against social data and adjusted in 101 cases. The final pedigree was used to determine the number of pups that emerged from the den (and were genotyped) of all adults present in a subpopulation in a given year. Annual fecundity was then determined as the number of pups that survived to recruit into the next years' population (i.e. were alive after 1st April next year). In addition, to analyse the probability of breeding a dichotomous variable (1/0) was made to indicate if an individual had been found to breed or not in any given year. This variable took the value 1 if pups had been recorded (0 otherwise). Adults not recorded as parents of any pups in the pedigree in a given year were assumed to have produced no pups that year.

Calculating ecological fitness

To estimate individual ecological fitness and to be able to account for age structure in the population, individuals were assigned to one of five age classes x (1-5+, where 1-4 designates the age of individuals in years, and 5+ designates all individuals five years or older), hereafter referred to as *age*. Only 35 individuals in the dataset were older than five years. Based on data on annual individual survival and fecundity (number of recruits) across all subpopulations and years in the study, the projection matrices for each sex (l_m and l_f, Table S3) of the population

were calculated using R package *lmf* (Engen et al., 2012; Kvalnes, 2013). The elements of l_m and l_f are fecundities in the first row and survival rates in the sub diagonal. Each element is estimated as mean values of individual survival or fecundity records for each age class over all years. Left and right eigenvectors of l describe the reproductive values v and stable age distribution u at population equilibrium respectively. The eigenvectors are scaled so that $\sum uv=1$, and $\sum u=1$. For each age class, reproductive values v_x are interpreted as the expected contribution an individual in age class t makes towards the equilibrium population. Individual fitness for a given individual was defined as:

$$\Lambda = \frac{W}{v_x} = \frac{\left(\frac{no.offspring * v_1}{2}\right) + (survival * v_{x+1})}{v_x}$$

Here, W is the individual reproductive value, *no. offspring* is the number of offspring the individual has produced in year t and survival is a binary variable of survival in year t (0/1). Additionally, individual fitness without weighting with reproductive values W^* was calculated for each individual and year t:

$$W^* = \frac{no.offspring}{2} + survival$$

See Engen et al. (2014) for further information on the topic of calculating and estimating ecological fitness.

Estimating selection

To investigate the effect of the fur colour genotype on the variation in individual fitness, a generalised linear mixed effect model (GLMM) with Poisson distribution and log link function was fitted with the *lme4* package in R (Bates et al., 2015). The Poisson distribution requires the response variable to be integers, hence, $2W^*$ was used as response variable in place of Λ. Then, to fit the model in a way that provides the correct parameter estimates and standard errors as if Λ was the response variable, an offset value c was defined to establish the relationship between $2W^*$ and Λ:

$$c_t = \frac{2{W^*}_t}{\Lambda_t}$$

When $\Lambda_t=0$, c_t was set to 1. In the model, the offset value was set to be *log(c)* due to the log link function. Additionally, model weights ω were defined to ensure correct estimation of

standard errors of the model estimates where v_t designates the age-specific reproductive values for each individual in year *t*.

$$\omega_t = v_t * \frac{1}{c_t}$$

To account for variation between years and subpopulations, random intercepts *for subpopulation* and *year* were included in the model. Models were fitted for females and males separately. Likelihood ratio tests (LRT) between models containing only the intercept and models containing *genotype* as predictor variable were performed to assess the effect of the genotype on mean fitness. Model predictions were calculated and visualised using R packages *ggeffects* (Lüdecke, 2018) and *ggplot2* (Wickham, 2016).

Decomposing fitness components

To investigate mechanisms behind differences in mean fitness, the relationships between fur colour genotype and the two fitness components included in the ecological fitness value Λ (survival and fecundity) were also analysed in separate models. Fecundity was modelled for adult individuals using a Poisson distributed GLMM with log link function. *Number of recruits* was used as the response variable and *genotype* as the predictor variable. Breeding probability of adult individuals was modelled using GLMMs with a binomial distribution and logit link function. The dichotomous (1/0) variable *detected breeding* was used as the response and *genotype* was used as the predictor variable. For adult survival, a GLMM with a binomial distribution and logit link function was fitted. A categorical *survival* variable (1/0) was used as the response variable and *genotype* was used as the predictor variable. For all response variables, the potential effect of sex and age was tested using LRTs for comparing models with and without *sex* and *age* as covariates. Whether the effect of genotype differed with sex or age (i.e. whether there was an interaction between *genotype* and the covariates) was tested in the same manner. The projection matrices (Table S3) have shown a non-linear relationship between age and fecundity and survival, thus a second-degree polynomial of age (age^2) was included in the models that contained age. In all cases, LRTs between models with and without the polynomial term for age were performed and supported the use of age^2. Random intercepts for *subpopulation* and *year* were included in all models to account for variation between years and subpopulations. Due to zero-inflation in the response variable *No. recruits* (Figure S4), fecundity models were fitted with a zero-inflation parameter using the *glmmTMB* R package (Brooks et al., 2017). LRTs between models with and without the zero-inflation parameter were performed and supported the need to include the parameter. The other models were fitted using

the *lme4* R package (Bates et al., 2015). The analysis on recruitment probability (i.e. juvenile survival until 1st of April the year following birth) was done using a restricted dataset based on the dens where summer fieldwork has been conducted. This allowed for most precise estimates of recruitment probability since the summer surveys capture approximately all pups that emerge from a den. *Recruitment* was used as a dichotomous (1/0) response variable in a binomial distributed GLMM with logit link function. *Genotype* was used as the predictor. The potential effect of sex on recruitment probability was tested using LRTs for comparing models with and without *sex* covariate, as well as testing the interaction between *genotype* and *sex*. Random intercepts for *birth year, subpopulation* and *den* were included in the models to account for interannual variation and non-independence of individuals sampled from the same den and subpopulation. Model predictions were calculated and visualised using R packages *ggeffects* (Lüdecke, 2018) and *ggplot2* (Wickham, 2016).

Testing for effect of environmental variables

Arctic foxes are strongly affected by winter climate and food access. Therefore, variables describing these environmental factors are included in the analysis to look for effects on fitness. Additionally, the potential effect of origin (i.e. wild- or captive-born individuals) was tested.

The abundance of small rodents usually varies in cycles between three to five years. Abundance data were gathered as part of the Norwegian Terrestrial Ecosystems Monitoring (TOV), where annual estimates of small rodent abundance per subpopulation are calculated based on the number of rodents trapped/100 trap-nights (Framstad, 2017). Based on Angerbjörn et al. (2013), rodent abundances were categorised in four rodent phases: low phase (1), increase phase (2), peak phase (3) and decline phase (4). These phases may, but do not need to follow each other and are thus used as factors rather than numerical in this analysis.

To estimate winter conditions, first day of snowfall and last day of snowfall were used. Here first day of snowfall describes the day of year (DOY) where first snowfall after 1st of September with subsequent accumulation of snow on the ground happens in year *t*. Last day of snowfall describes the DOY with latest snowfall and following accumulation of ground snow before 1st of September in year *t.* Snow data was retrieved from the Norwegian Water Resources and Energy Directorate (Saloranta, 2014) and was extracted for a 2.54 km buffer area around used arctic fox den sites in the study subpopulations. This buffer area is an average of annual home ranges of resident arctic foxes presented in Landa et al. (1998). Within a subpopulation and year, values were averaged across all buffer areas. Since small scale movements of the

individuals are not known, the averaging approach was chosen to remove the risk of wrongly assigning small scale snow data (that might be influenced by microhabitat) to individuals. Both snow variables, first day of snowfall and last day of snowfall, were mean-centred before the analyses to create biologically meaningful intercepts (i.e. intercept represents the response variable at mean first or last day of snow fall).

In this study, it was of interest whether fur colour genotypes were affected differently by environmental factors. Therefore, it was only investigated whether there was an interaction effect between *genotype* and the different environmental variables rather than the main effect. For each response variable (mean fitness in females and males, and the fitness components adult survival, recruitment probability, fecundity and breeding probability), the models (including error distribution, random factors etc.) presented in the sections *Estimating selection* and *Decomposing fitness components* were expanded with one environmental variable at a time. For each environmental variable and response variable, one model with additive effects (genotype + *ecological variable*) and one with an interaction (genotype * *ecological variable*) was fitted. In cases where *age* or *sex* was found to be significant covariates to genotype (see *Decomposing fitness components*), these were included in both the additive and interaction model. LRTs were performed to test whether there was support for the interaction between fur colour genotype and the environmental variable tested.

The variables investigated can have effects that are offset one year back or forth (e.g. reproduction in year *t* may be affected by both rodent phase in year *t* and the year before *t*-1). Additionally, not all environmental variables are expected to affect the fitness variables (e.g. breeding probability in year *t* cannot ecologically be affected by rodent phase in year *t*+1). Thus, not all environmental variables were tested for all fitness measures. Table 1 shows which environmental factors (and which temporal offsets) were tested for the different fitness measures.

Table 1: Overview over which predictor variables (and their temporal offsets) were included when investigating environmental effects on different fitness components (response variables). A + designates predictor variables included in the analysis for a given fitness component. Variables were chosen after ecological significance.

Ecological predictor variable	Response variable				
	Fitness (females and males)	**Survival**	**Recruitment probability**	**Fecundity**	**Breeding probability**
Origin	+	+		+	+
Rodent phase t	+	+	+	+	+
Rodent phase t+1	+	+	+		
Rodent phase t-1	+			+	+
First snowfall t	+	+	+		
First snowfall t-1	+			+	+
Last snowfall t	+			+	+
Last snowfall t+1	+	+	+		

Temporal changes in genotype frequencies

After I quantified selection, it is was of interest to see whether potential effects of selection are visible in a spatio-temporal frame. For this, proportions of the three fur colour genotypes in the two subpopulations where most arctic foxes had been released during the captive breeding and release program (Hardangervidda and Snøhetta). For each year, proportions were calculated based on the total number of adult arctic foxes (i.e. those that survived until 1st of April). Furthermore, the number of captive-born foxes released into the wild subpopulations were counted to be able to assess potential effects of released foxes on spatio-temporal variations in genotype frequencies.

GENE ONTOLOGY

Genes within 10 kbp of significant SNPs were analysed for gene ontology (GO) term enrichment using the GOstat tool (Beißbarth & Speed, 2004). The distance of 10 kbp was chosen to ensure strong linkage between the SNP and the gene. In lack of a dog-specific GO-database, the goa_human database was used. P-values for overrepresentation significance were corrected based on false discovery rate (FDR). For genes within 10 kbp of a significant SNP that is in high LD ($r^2 >= 0.5$) with the top SNP, gene functions were investigated using the UniProt knowledgebase (The UniProt Consortium, 2018). These genes were also included in a GeneMANIA network analysis (Warde-Farley et al., 2010). GeneMANIA uses a large dataset of functional association data to analyse among others, known co-expression between the input genes.

GENOME-WIDE HETEROZYGOSITY

Heterozygote advantage (Sellis et al., 2011) might affect the fitness of the different fur colour genotypes. To investigate this, genome-wide heterozygosity was calculated for 689 individuals based on individual genotype data on a set of 359 218 autosomal SNPs distributed across the arctic fox genome using the *GenABEL* R package (Aulchenko et al., 2007). Differences in genome-wide heterozygosity were modelled with a linear mixed effect model with a Gaussian error distribution. The full model included *origin* (wild-born or captive-born) and *genotype* as predictor variables, as well as their interaction. *Birth year* and *birth subpopulation* were included as random factors to account for variation between subpopulations and years. The effect of the predictors and the interaction was assessed by comparing the full model with models without the interaction and only one of the predictor variables respectively using likelihood ratio tests.

RESULTS

GENE MAPPING

The genome-wide association analysis revealed a total of 495 SNPs significantly associated with arctic fox fur colour at a Bonferroni-adjusted significance level that accounted for the large number of SNPs tested ($p < 1.39E-07$) (Figure 2). These significant SNPs were located on four different scaffolds of the arctic fox genome (Figure 2, Table 2). Most significant SNPs were located on arctic fox scaffold number 11, and of these, AX-177333963 was the SNP with the lowest p-value (λ-corrected $p = 6.84E-63$). Detailed information about position, sample size, effect size, major and minor alleles, including χ^2 test statistics for the all significant SNPs, can be found in Supplementary material 5.

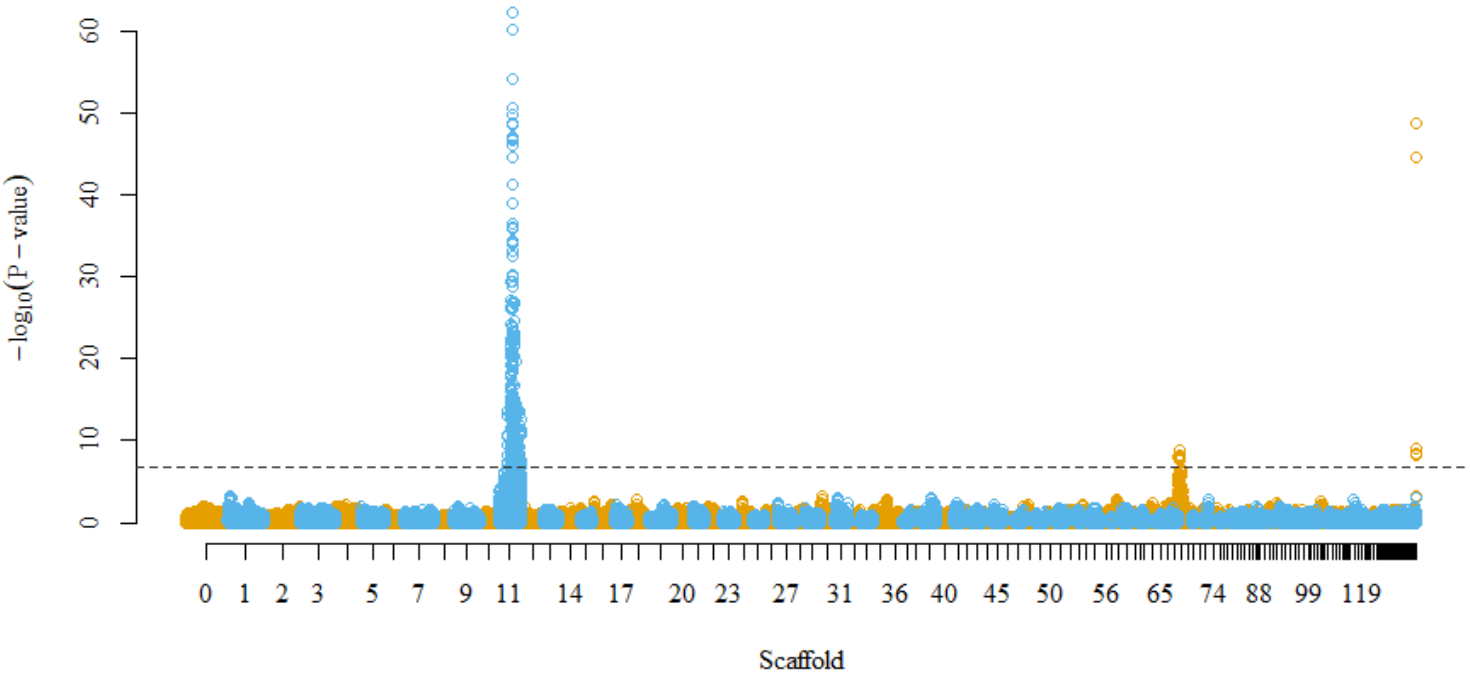


Figure 2: Manhattan plot showing results of genome-wide association study of arctic fox fur colour using information on 359 219 SNPs typed in 681 individuals. P-values are given on a negative log scale. The significance threshold is shown with a dashed horizontal line. Colours alternate between scaffolds. Note that scaffolds are ordered after scaffold length, and that scaffolds to the right are short, thus seemingly merged into each other in the graph.

Table 2: Distribution of SNPs significantly associated with arctic fox fur colour across arctic fox scaffolds. For each scaffold, the total number of SNPs included in the GWA-analysis and number of significant SNPs are given. Also, the SNP with the lowest p-value on each scaffold is given as "top SNP".

Scaffold	SNPs	Significant SNPs	Top SNP
11	6769	477	AX-177333963
68	1636	13	AX-177295135
1772	4	3	AX-176900926
2224	2	2	AX-177181351
Total	**8411**	**495**	

All 495 significant SNPs were BLASTed against the annotated dog genome *CanFam 3.1*. I obtained a match in the dog genome for 489 SNPs (486 on chromosome 5, two on chromosome 27 and one on chromosome 17). A total of 438 SNPs were intragenic in the dog genome, whereas the remainder 51 SNPs were located in intergenic regions. The intragenic SNPs were distributed within 98 different genes (Table S6). An additional 57 genes were found less than 20 kbp away from a significant SNP, with 34 of these being closer than 10 kbp from a significant SNP (Table S6). The position of significant SNPs on chromosome 5 stretched from 52 617 594 to 76 592 936 bp (Figure 3), a distance that appears to be outside of strong linkage

disequilibrium in the arctic fox genome (Figure S7). A total of 379 genes are located in this region of the dog genome.

The Affymetrix SNP array used in this study does not include any SNP located in the intragenic region of the candidate gene MC1R. In fact, the closest SNP to MC1R, AX-177360772, lies 5 535 bp downstream of MC1R and is non-significant (p = 0.15). SNP AX-176934441 is the closest significant SNP (5 961 bp upstream; p = 6.7E-61). Because the SNP with lowest p-value (AX-177333963; p = 6.8E-63) is located 19 830 bp upstream of MC1R, I chose AX-176934441 as the diagnostic SNP for the alternative genotypes at the MC1R gene in further analyses.

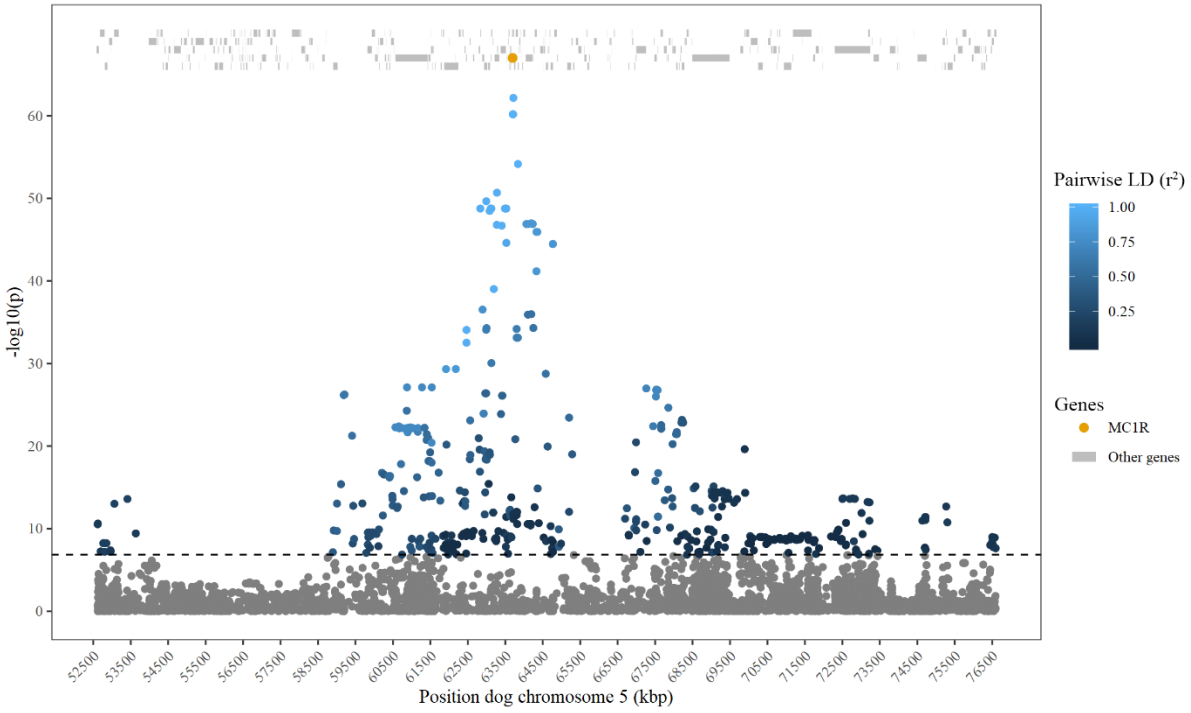


Figure 3: Plot showing BLAST determined dog chromosome 5 locations of 486 SNPs significant in GWAS of fur colour in arctic fox. On the y-axis, significance levels of the SNPs in the GWAS are shown on a negative log scale. Pairwise linkage disequilibrium (r^2) between top SNP AX-177333963 and the other significant SNPs is shown by the blue colour gradient. All dog genes in the region are shown with grey lines. The position of putative causal gene MC1R is shown with an orange dot (note that the y-axis values do not apply for genes). The dashed horizontal line shows the significance threshold of the GWAS.

There was a near-perfect Mendelian relationship between genotypes at MC1R and fur colour phenotypes, where the C allele represents a recessive white fur colour allele, and T a dominant blue fur colour allele. The MC1R genotypes agree with a simple Mendelian inheritance of fur colour phenotype for 98.4% of the 681 arctic foxes that were SNP genotyped at the Affymetrix SNP-array (Table 3). Furthermore, genotyping of 444 arctic foxes with fur colour phenotype on MC1R using a Fluidigm SNP-array confirmed this result: genotypes of 98.6% individuals

were following simple Mendelian inheritance (Table 3). Analysis of 12 whole-genome sequenced arctic fox individuals found the same base-pair mutations in MC1R that were presented by Våge et al. (2005), in the one blue individual available. The eleven white individuals did not show these mutations. All other SNPs found in MC1R had the same genotype in one or more white foxes and the blue fox. See Supplementary material 8 for detailed information on this analysis.

Table 3: Relationship between arctic fox fur colour phenotypes and MC1R genotypes (as represented by SNP AX-176934441) from genotyping on either an Affymetrix (N=681) or a Fluidigm (N=444) SNP-array Expected phenotypes based on the assumption of simple Mendelian inheritance at one locus is given as well.

	Affymetrix		**Fluidigm**		
Genotype / Phenotype	**White**	**Blue**	**White**	**Blue**	**Expected phenotype**
CC	554	3	324	1	White
TC	8	113	5	108	Blue
TT	0	3	0	6	Blue

ESTIMATING SELECTION

Due to low sample size, TT individuals had to be excluded from the analyses of individual fitness ($n_{females}$=2, n_{males}=5). In females, heterozygous (TC) individuals have significantly higher fitness than homozygotes (CC) at the white allele ($\chi^2(1)$=4.32, p=0.038) (Figure 4, Table 4). Likewise, in males, TC individuals appear to have higher fitness than CC individuals, but the difference is not statistically significant ($\chi^2(1)$=1.64, p=0.201) (Figure 4, Table 4).

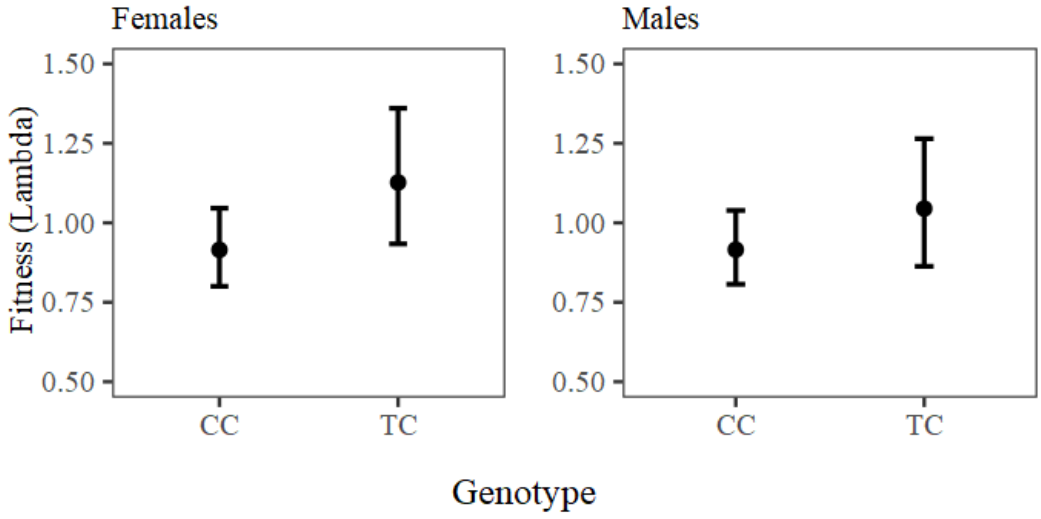


Figure 4: Predicted fitness (Lambda) of the arctic fox fur colour morph genotypes CC and TC. Whiskers represent 95% confidence intervals. Predictions are based on additive GLMMs with *genotype* as predictor variable.

Table 4: Model estimates and their 95% confidence intervals for GLMMs modelling fitness with *genotype* as predictor variable. Model estimates and confidence intervals are given on log scale. Estimates significant at the 0.05 significance level are given in bold.

		95% CI	
Predictor variable	**Estimate**	**Lower**	**Upper**
Females			
CC (Intercept)	-0.089	-0.237	0.055
TC	**0.208**	**0.012**	**0.400**
Males			
CC (Intercept)	-0.089	-0.227	0.045
TC	0.132	-0.071	0.329

DECOMPOSING FITNESS COMPONENTS

Fecundity

Fecundity was significantly affected by age and increased from age 1 to 4 before decreasing at age 5 (β_{age}=1.478, $\beta_{age\wedge2}$=-0.203, $\chi^2(2)$=50.62, p<0.001; Figure 5b). Heterozygous individuals had higher fecundity than CC individuals (β_{TC}=0.497±0.162, $\chi^2(1)$=4.54, p=0.033), however, the effect was more pronounced in females ($\beta_{TC \times Male}$=-0.528±0.241, $\chi^2(2)$=5.47, p=0.065; Figure 5a). The effect of genotype was independent of age ($\beta_{TC \times age}$=0.541±0.455, $\chi^2(2)$=1.69, p=0.429).

Breeding probability

There was a significant difference between genotypes on the probability of breeding that differed between the sexes ($\beta_{TC \times Male}$=-0.644±0.319, $\chi^2(1)$=4.09, p=0.043; Figure 5c). Heterozygote females had a higher probability of breeding than individuals homozygous at the fur colour locus, while there was no clear pattern in males (Figure 5c). Breeding probability increased with age until age 4, before levelling off at age 5 (β_{age}=1.383±0.278, $\beta_{age\wedge2}$=-0.150±0.046, $\chi^2(2)$=98.05, p<0.001; Figure 5d), an effect that was independent of genotype ($\beta_{TC \times age}$=-0.658±0.600, $\chi^2(2)$=4.78, p=0.091).

Recruitment probability (Juvenile survival)

The recruitment probability (juvenile survival) was found to be independent of genotype (β_{TC}=0.026±0.237, $\chi^2(1)$=0.012, p=0.914, Figure 5e), suggesting that the tendency for higher fecundity for heterozygote individuals origins from the probability of breeding or the number of pups born. Neither was recruitment probability affected by sex (β_{male}=0.013±0.201, $\chi^2(1)$=0.004, p=0.948) nor an interaction between sex and genotype ($\beta_{TC \times male}$=-0.713±0.457, $\chi^2(1)$=2.403, p=0.121).

Adult survival

Survival probabilities increased from age 1 to age 2 and decreased for individuals older than 3 (β_{age}=0.460±0.240, $\beta_{age\wedge2}$=-0.093±0.041, $\chi^2(2)$=6.94, p=0.031; Figure 5f). Besides, there was a marginal significantly higher probability of survival for heterozygote compared to homozygote individuals (β_{TC}=0.296±0.157, $\chi^2(1)$=3.632, p=0.057, Figure 5f). This effect was not affected by sex ($\beta_{TC\ x\ male}$=0.217±0.308, $\chi^2(1)$=0.491, p=0.483) and changed with age similarly for the two genotypes ($\beta_{TC\ x\ age}$=0.606±0.559, $\chi^2(2)$=1.599, p=0.450).

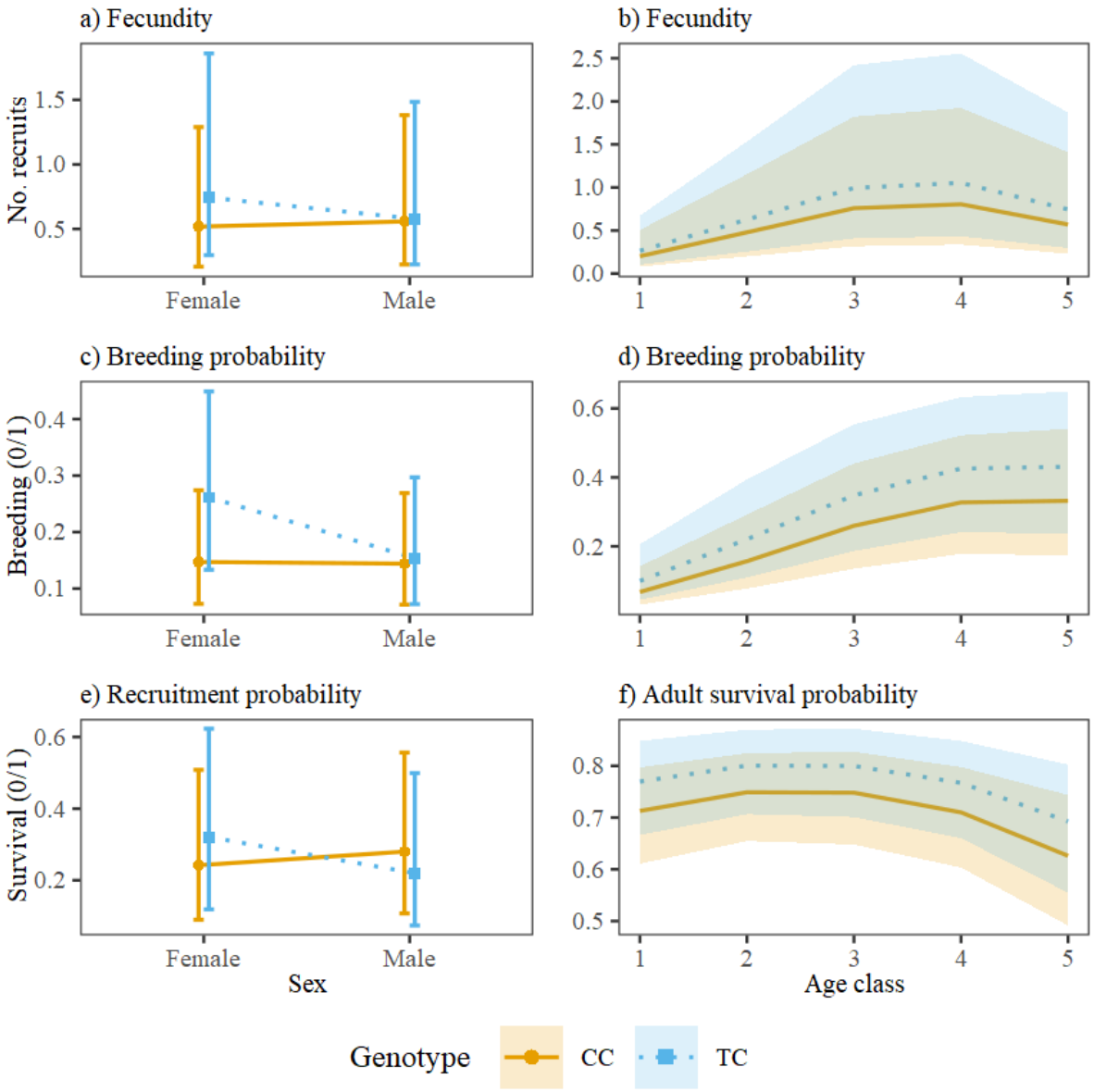


Figure 5: Predicted fecundity (a, b), breeding probability (c, d), recruitment probability (e) and adult survival probability (f) for arctic fox fur colour genotypes CC and TC at different age classes and both sexes. Whiskers and shaded areas represent 95% confidence intervals. Predictions are based on additive GLMMs with *genotype* and *sex*, as well as *genotype* and *age* as predictor variables. *Age* is included as a quadratic term to account for the non-linear relationship between age and the response variables.

EFFECT OF ENVIRONMENTAL VARIABLES

Different fur colour genotypes had the tendency to be affected differently by rodent phase in terms of adult survival ($\chi^2(3)=7.36$, $p=0.061$) and fecundity ($\chi^2(3)=9.32$, $p=0.025$). Also, rodent phase the year before breeding showed the tendency to affect fur colour genotypes differently with regards to fecundity ($\chi^2(3)=6.31$, $p=0.097$). In years of low (1) or increasing rodent phase (2), adult TC individuals seemed to have higher survival probabilities than CC individuals (Figure 6a), whereas there were no differences in peak (3) or crash (4) years. In general, fecundity was higher in TC individuals across all rodent phases, but the difference appeared to be strongest in increase years (2; Figure 6b). Fecundity was higher in years following an increase phase (2) in which case there was no difference between fur colour genotypes (Figure 6c). In years following a low phase (1), TC individuals seem to have somewhat higher fecundity than CC individuals (Figure 6c). None of the other studied environmental factors affected fur colour genotypes differently (Table S9).

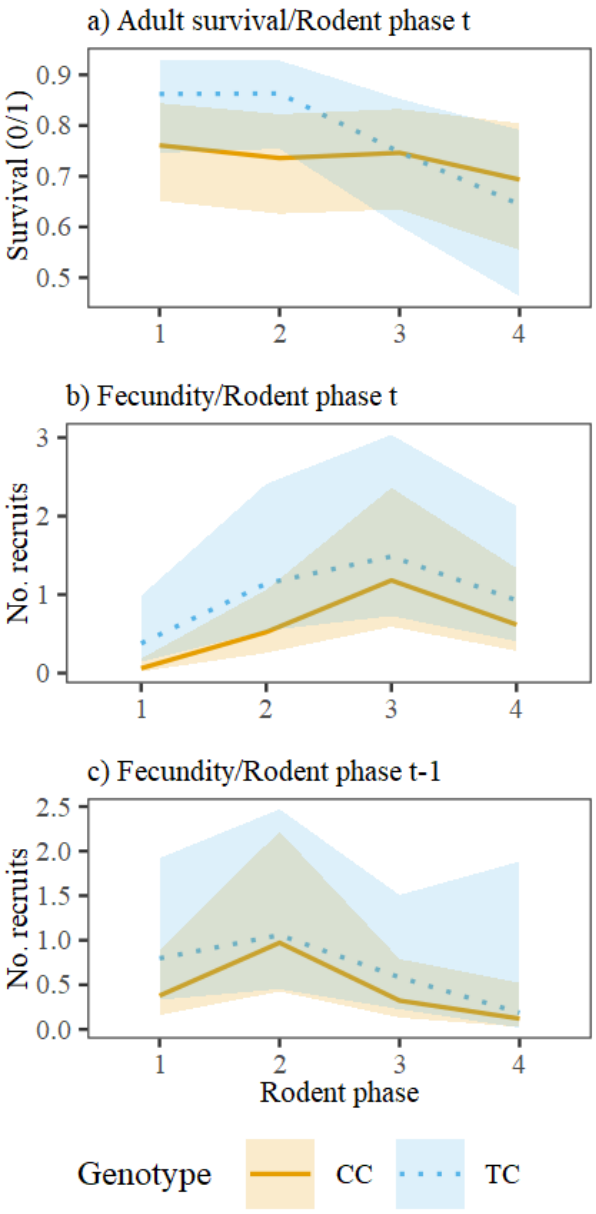


Figure 6: Predictions for adult survival (a), fecundity (b, c) of arctic fox fur colour genotypes in different rodent phases. c) shows rodent phase the year before breeding.

TEMPORAL CHANGES IN GENOTYPE FREQUENCIES

To get an idea whether the observed fitness differences can be seen in the wild subpopulations, and whether the release of individuals from the captive breeding programme might affect those subpopulations, I have analysed how fur colour genotype frequencies have changed over time in the two subpopulations Hardangervidda and Snøhetta. These two subpopulations were chosen because they are two of the largest and most important in Norway and individuals were released over several years. The subpopulations consisted of over 90% CC individuals at the

beginning of the study period and the proportion of these individuals gradually decreased to 73% and 74% in 2019, respectively (Figure 7). Since the proportion of TT individuals is negligible, the proportion of heterozygotes (TC) increases in accordance with the decrease of CC individuals. During the study period, 208 (180 CC, 27 TC, 1 TT) and 100 (88 CC, 12 TC) captive-born individuals were released into the Hardangervidda and Snøhetta subpopulations, respectively. According to the census year used in this study (1st of April – 31st of March the following year), foxes that were released early in the year are counted as released the year before. If a released fox recruited to the adult population, it will thus show up the year after release (albeit technically, release and recruitment happened in the same calendar year).

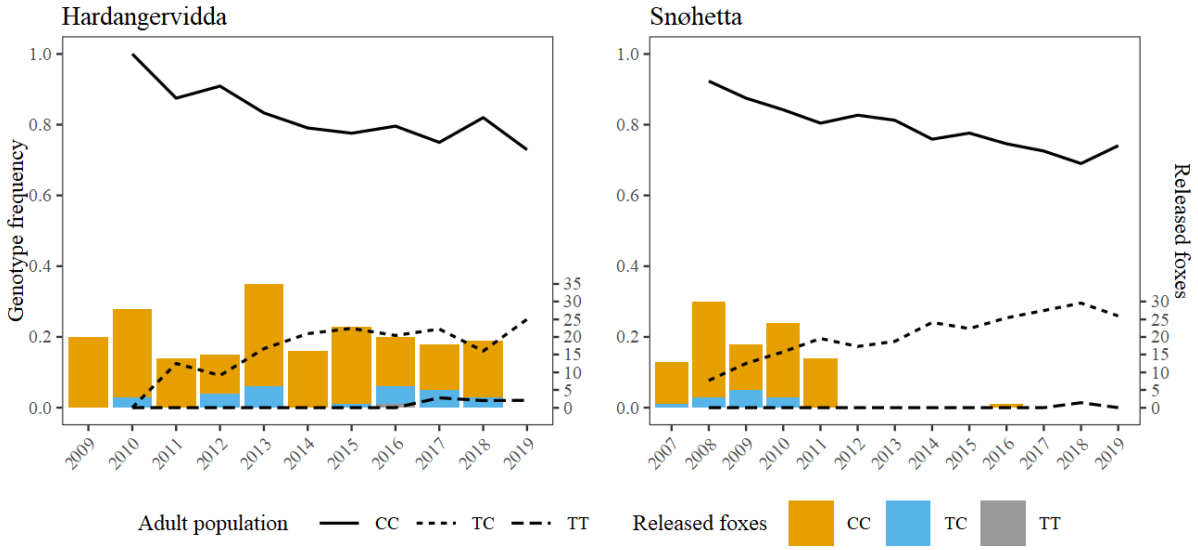


Figure 7: Temporal changes in MC1R genotype frequencies in the subpopulations Hardangervidda and Snøhetta where large numbers of arctic foxes have been released during the captive breeding and release program. Genotype frequencies (left y-axis) are calculated for the adult population each year. Number of released foxes of each genotype is given on the right y-axis. Note that released foxes are counted towards census year (1st of April – 31st of March), i.e. foxes released before 1st of April are counted the calendar year before.

GENE ONTOLOGY

As presented earlier, many genes are located close to and in LD with MC1R (Figure 3). Consequently, changes in MC1R allele- and genotype frequencies (Figure 7) will lead to changes in frequencies of variants at other genes as well. To gather insight on what functions these genes have and how they might be expressed in arctic foxes, I have conducted some preliminary gene ontology investigations. For 132 genes that were found to be less than 10 kbp away from a SNP significantly associated with arctic fox fur colour, a gene ontology (GO) term enrichment analysis showed overrepresentation of 33 GO terms (Table 5). Many of these GO terms represent fundamental biological functions (e.g. Cytoplasm, Intracellular or Organelle). Eight of the 33 overrepresented GO terms are involved in metabolic processes, six of them in

lipid metabolism (Table 5). Other GO terms are involved in developmental processes (Developmental processes, Regulation of Wnt signalling pathway).

Table 5: Gene ontology (GO) terms that were overrepresented among 132 genes located closer than 10 kbp to a SNP significantly associated with arctic fox fur colour. P-values are corrected for multiple testing using false discovery rate (FDR). Column *Gene count* shows how many of the 132 genes are part of the GO terms. Numbers in brackets show the total number of genes in each GO term.

GO ID	Gene ontology term description	P_{FDR} value	Gene count
GO:0044424 [Consider GO:0005622]	Obsolete intracellular part	0.00128	50 [12958]
GO:0043231	Intracellular membrane-bounded organelle	0.00128	38 [8824]
GO:0043227	Membrane-bounded organelle	0.00128	38 [8827]
GO:0005737	Cytoplasm	0.00338	33 [7482]
GO:0005622	Intracellular	0.00669	52 [14906]
GO:0043229	Intracellular organelle	0.0172	40 [10763]
GO:0043226	Organelle	0.0172	40 [10768]
GO:0032502	Developmental process	0.0172	18 [3347]
GO:0008013	Beta-catenin binding	0.0172	2 [9]
GO:0048523	Negative regulation of cellular process	0.0175	10 [1137]
GO:0005515	Protein binding	0.0177	35 [9005]
GO:0048519	Negative regulation of biological process	0.0199	10 [1182]
GO:0016043	Cellular component organization	0.0252	17 [3277]
GO:0009416	Response to light stimulus	0.0433	3 [81]
GO:0007275	Multicellular organism development	0.0447	13 [2299]
GO:0044464	Cell part	0.0495	63 [21746]
GO:0006629	Lipid metabolic process	0.0495	8 [946]
GO:0044255	Cellular lipid metabolic process	0.0495	7 [768]
GO:0006512	Obsolete ubiquitin cycle	0.0495	6 [549]
GO:0008202	Steroid metabolic process	0.0495	4 [223]
GO:0009314	Response to radiation	0.0495	3 [101]
GO:0030111	Regulation of Wnt signalling pathway	0.0495	2 [27]
GO:0047936	Glucose 1-dehydrogenase [NAD(P)] activity	0.0495	1 [1]
GO:0004671	Protein C-terminal S-isoprenylcysteine carboxyl O-methyltransferase activity	0.0495	1 [1]
GO:0050201	Fucokinase activity	0.0495	1 [1]
GO:0055098 [Replaced by GO:0071404]	Cellular response to low-density lipoprotein particle stimulus	0.0495	1 [1]
GO:0002040	Sprouting angiogenesis	0.0495	1 [1]
GO:0055094	Response to lipoprotein particle	0.0495	1 [1]
GO:0055095	Lipoprotein particle mediated signalling	0.0495	1 [1]
GO:0008267	Poly-glutamine tract binding	0.0495	1 [1]
GO:0030223	Neutrophil differentiation	0.0495	1 [1]
GO:0055096	Low-density lipoprotein particle mediated signalling	0.0495	1 [1]
GO:0043890	N-acetylgalactosamine-6-sulfatase activity	0.0495	1 [1]

To limit the analysis to genes that likely are highly associated with MC1R genotype, I looked for genes closer than 10 kbp to a SNP that is a) significantly associated with arctic fox fur colour and b) in high LD (r^2 >= 0.5) with the SNP most associated with arctic fox fur colour. Here, 42 genes were found, and their functions according to UniProtKB are summarised in Table S10. Only three of these genes were Swiss-Prot reviewed for dogs, MC1R being one of them. For several of the genes listed here, important functions are known. These include regulation of the Wnt signalling pathway (CTNNBIP1), DNA reparation (FANCA), glucose metabolism (H6PD), development (RERE) and immune response (PIK3CD, USP18, BANP). The same 42 genes were included in the GeneMANIA analysis that showed co-expression of MC1R with four genes: CTNNBIP1, GSE1, PIEZO2, TCF25 (Figure S11). PIEZO2 itself was included in the analysis by the software without being among the input genes. It is closely related to and shares protein domains with PIEZO1 that was in the input. GeneMANIA did not present coexpression for any of the other input genes.

GENOME-WIDE HETEROZYGOSITY

To investigate whether observed patterns in selection could be influenced by heterozygote advantage, I analysed genome-wide heterozygosity in individuals of the different fur colour genotypes. Genome-wide heterozygosity was significantly affected by both genotype ($\chi^2(1)=5.96$, p=0.015) and origin ($\chi^2(1)=5.34$, p=0.021). The interaction between origin and genotype was not significant ($\chi^2(1)=0.14$, p=0.710). Genome-wide heterozygosity was 0.02±0.007 lower in wild-born foxes (n=314) compared to foxes born at the breeding station (Figure 8). Furthermore, foxes heterozygous at MC1R (n=124) showed 0.006±0.002 higher level of genome-wide heterozygosity than foxes with the CC genotype (Figure 8). Due to low sample size (n=3), no meaningful comparison could be made with genotype TT individuals.

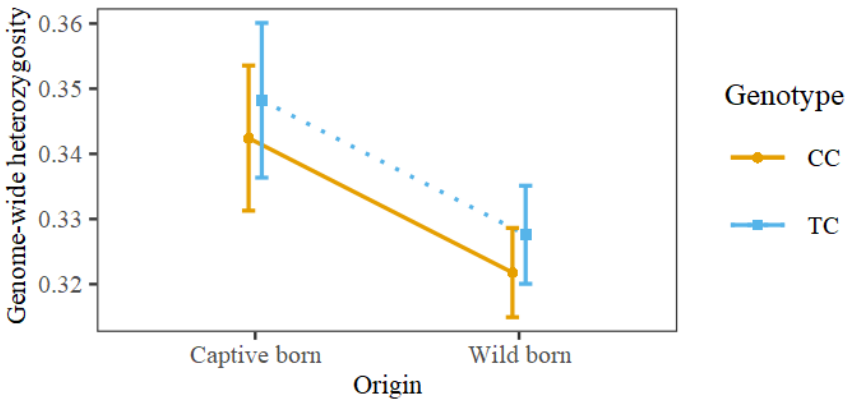


Figure 8: Model predictions for genome-wide heterozygosity for captive-born and wild-born individuals of fur colour genotypes CC and TC. Predictions are based on an additive GLMM with *origin* and *genotype* as the predictor variables.

DISCUSSION

In this study, I have investigated the genetic basis and architecture of arctic fox fur colour. My results represent strong evidence that MC1R is the only causal gene underlying the white and blue fur colour morphs in the arctic fox. Additionally, I have quantified selection on the colour morphs and found a fitness advantage of heterozygous individuals at the MC1R gene that appeared to be similar across most environmental conditions. This fitness advantage was stronger in females than in males, and different genotypes were partly affected differently by food access (rodent cycle). The gene ontology analyses showed that several genes involved in developmental and metabolic processes covary with arctic fox fur colour genotype.

My GWA analysis identified many SNPs with significant association to arctic fox fur colour (Figure 2). BLAST results showed that all but three of the significant SNPs were spread around MC1R in the region from 52 to 77 mbp on dog chromosome 5 (Figure 3). Together with MC1R sequence data from 12 whole-genome sequences fox (Supplementary) and the near-perfect association between arctic fox fur colour and the top SNP genotype (Table 3), these GWA results support the hypothesis by Adalsteinsson et al. (1987) that fur colour morphs in arctic fox is determined by a single Mendelian allele and the results of Våge et al. (2005) which suggested MC1R as the sole causal gene behind the distinct arctic fox fur colour morphs. While Våge et al. (2005) found strong indications that MC1R is involved in arctic fox fur colouration, one has to be cautious when investigating the genetic basis of traits based on a candidate gene approach, as this approach may have some pitfalls as outlined by Brown et al. (2013) and Slate

(2015). Additionally, the original study based their results on very few individuals (Våge et al., 2005). The genome-wide scale of this study confirms that no other areas in the genome explained variation in fur colour and provides evidence of MC1R's role based on much larger sample size. While the candidate-gene approach can work (and in fact, has worked in this case), large-scale genome scans should be preferred and used to verify causal genes due to their unbiased approach.

The quantification of selection showed that foxes heterozygous at the fur colour locus have higher fitness than individuals that are homozygous for the white allele C (Figure 4). My modelling revealed significant differences only for females, but the trend for males mirrored that of females (Figure 4). Decomposition of fitness into different components strengthened the selection results and showed that TC individuals scored better in both fecundity, breeding probability and adult survival probability than CC individuals, with the effect being reduced to female foxes (Figure 5). Only recruitment probability (juvenile survival) was not affected by genotype per se. However, female TC foxes had slightly higher recruitment probability than female CC individuals, whereas it was the opposite way in males (Figure 5e). Unfortunately, blue homozygotes (TT) are rare in the study area and quantifying potential differences between blue homozygotes and the other genotypes with confidence would require a larger sample size, hence these individuals were not included.

Ultimate causes (i.e. evolutionary mechanisms) underlying arctic fox fur colouration are understudied. It has been hypothesized that the main difference between the distinct colour morphs is their camouflage value for different habitats, a hypothesis that is mostly based on the small scale distribution patterns observed in Iceland (Hersteinsson, 1989) and our general knowledge about animal colouration and its function as anti-predation camouflage (Caro, 2005). Predation (e.g. by golden eagles *Aquila chrysaetos*, red foxes *Vulpes vulpes* or wolverines *Gulo gulo*) is hard to quantify in a wild alpine species and we do not have empirical data on predation pressure in the Norwegian population. Some of the known predators of arctic foxes do in addition not occur regularly in the distribution area of the Norwegian arctic fox (Brown bears *Ursus arctos*, snowy owls *Bubo scandiacus*, wolves *Canis lupus*). However, predation in the study area does not seem to be large enough to inflict a decrease in survival probability for the (at least in the winter) more conspicuous blue colour morph, as observed here (Figure 5f). This is in accordance with recent results from Di Bernadi et al. (in prep.) where blue arctic foxes, in fact, had marginally higher survival probabilities compared to white

foxes. Furthermore, there are no obvious reasons for why predation pressure should act differently on sexes as shown in my results (Figure 5e)

A second route of direct selection on arctic fox colouration may be due to thermoregulation features associated with different fur colours. Arctic foxes live in harsh environments and experience severe conditions in terms of cold and food scarcity during the winter months. It seems reasonable to assume that selection could work through the response to these climatic conditions. My results show that individuals of different fur colour genotypes were somewhat affected differently by access to food, albeit no very clear pattern was obvious (rodent phase; Figure 6). The different fur colour genotypes were not affected differently by snow season (first or last day of snowfall; Table S9). These results indicate that the adaptations the arctic fox has to withstand the arctic winter (Prestrud, 1991; Prestrud & Nilssen, 1992), are mostly independent of colour genotype. The general perception that camouflage is a more potent driver of animal colouration than thermoregulation is also in line with these findings (Caro, 2005; Zimova et al., 2018). However, while the insulative properties of the white winter fur have been quantified (Underwood & Reynolds, 1980; Prestrud, 1991), potential differences in the insulative properties of the white and blue winter fur have not been investigated. Melanin-coloured feathers are shown to have higher resistance than other feathers (Bonser, 1995), a property that can lead to birds spending less time on preening behaviour (Roulin, 2007). It remains unknown whether a similar difference could apply for melanin-based fur as well.

Since support for the two most likely routes of direct selection on arctic fox fur colour is weak or missing in my results, it seems reasonable to explore potential routes of indirect selection. Pleiotropic interactions in the melanocortin complex, which MC1R is part of, are well known and reviewed in Ducrest et al. (2008). Many of the differences found between darker, melanin-coloured individuals and their lighter conspecifics are indeed favourable and potentially fitness-enhancing for the dark morph. Both experimental and observational studies have shown the large variety of traits that are affected by the melanocortin system and thus covary with melanin-based colouration (Ducrest et al., 2008). Many of these traits have the potential to play vital roles for a wild species living in a harsh climate such as resistance to stressors and enhanced immune response. Also, behavioural traits such as aggressiveness are affected by pleiotropy in the melanocortin system and could impact foxes with genotypes for the blue colour morph positively in terms of getting access to good den sites and chasing away both conspecifics as well as competitors (e.g. red foxes). The last group of traits affected by pleiotropy in the melanocortin system is sexual traits, and both sexes can be affected positively

in terms of sexual receptivity and fertility (Ducrest et al., 2008). One could also expect higher fertility in male blue foxes based on findings that male blue foxes had higher concentration of spermatozoa in their ejaculates compared to white foxes (Stasiak et al., 2019). Yet, my findings do not concur as I did not find a difference in fecundity or breeding probability between male foxes with CC (i.e. white) and TC (i.e. blue) genotype (Figure 5a, c), indicating that the difference in spermatozoa concentration does not translate into higher reproduction in wild arctic foxes.

MC1R is located in a region with numerous other genes and I found several genes close to SNPs that were significantly associated with arctic fox fur colour (Figure 3). Based on an analysis of linkage disequilibrium (Figure 3, Figure S7), it is certain that some of these genes covary with fur colour genotype. Both, the GO term analysis (Table 5) and the analysis of genes close SNPs significantly associated with arctic fox fur colour (Table S10) show that genes expected to covary with MC1R genotypes are involved in important processes. As for all species enduring harsh winter conditions, the ability to control metabolism is relevant and potentially vital for arctic foxes in enduring cold climate and food scarcity. Eight of the overrepresented GO terms were related to lipid and steroid metabolism, making this an interesting pathway for future studies to investigate. Regulation of the Wnt signalling pathway showed up in my results as an enriched GO term (Table 5), as well as a single gene in form of CTNNBIP1. This pathway plays significant roles in organism development (Komiya & Habas, 2008) and inhibition can lead to severe and potentially fatal effects (Satoh et al., 2004). Several other overrepresented GO terms were also part of developmental processes (Table 5). In addition, the gene RERE that plays a role in developmental processes was found among the genes likely to covary with MC1R. Another three of these genes are involved in immune response (BANP, PIK3CD and USP18), a trait that also is part of the pleiotropic melanocortin system. Precisely how these genotypes are expressed phenotypically and whether these phenotypes can affect viability in the arctic fox remains to be seen. However, all these genes provide good examples of covarying genes that may (or may not) have serious implications for individual arctic fox fitness and should be investigated in more detail in future research on this topic.

A last potential explanation of higher fitness in blue heterozygotes compared to white homozygotes is heterozygote advantage (Sellis et al., 2011). I have found significantly higher genome-wide heterozygosity in individuals heterozygous at the fur colour locus compared to those homozygous for the C allele (Figure 8), indicating that there are many loci where the

heterozygote advantage could work. Although there has been some doubt previously about how many loci are affected by heterozygote advantage (Hedrick, 2012), it has been shown that heterozygote advantage indeed can play a role in fitness-related traits in wild populations (Johnston et al., 2014). Due to the lack of data on individuals homozygous for the T allele, I could not show heterozygote advantage at the fur colour locus itself.

It was also shown that foxes born at the captive breeding station had higher genome-wide heterozygosity than individuals born in the wild (Figure 8). This could indicate a lower degree of inbreeding in captive-born foxes, which albeit seeming contradicting at first glance, makes sense as breeding pairs in the breeding station are chosen to maintain genetic diversity. Hasselgren et al. (2018) presented a good example of the genetic rescue effect where blue arctic foxes enjoyed great reproductive success in an inbred population in Sweden. It is possible that we can see a light genetic rescue effect in this study as well, and that the observed reproductive advantages of heterozygous individuals (Figure 5) might be the result of genetic rescue by the release of TC individuals from the breeding station.

In the two largest Norwegian subpopulations, Snøhetta and Hardangervidda, I observed a gradual frequency increase of heterozygotes during the last ten years (Figure 7). This increase can be explained by a variety of mechanisms outlined in this study. It could be the result of the apparent fitness advantage of the TC genotype in terms of both survival and annual reproduction. Also, a general heterozygote advantage can contribute to explaining the frequency increase. As outlined before, the release of captive-born individuals has the potential of affecting genotype frequencies in the wild populations and since the subpopulations consisted of very few blue individuals before the release, the release of blue foxes has likely contributed to an increase of the T allele and consequently blue fur colour genotypes in the wild population. Lastly, effective population sizes are still relatively small in all Norwegian subpopulations and genetic drift will thus most likely influence genotype frequencies over time. A more detailed analysis would be needed to completely differentiate the observed changes in genotype frequencies between genetic drift and selection.

My study adds to the body of research that has identified major genes underlying traits with fitness implications for a wild animal species through genetic mapping (Johnston et al., 2011; Ayllon et al., 2015; Barson et al., 2015; Lawson & Petren, 2017; Sinclair-Waters et al., 2020). However, my results also reveal the large potential for interesting genetic interactions that are hidden behind the seemingly simple trait of arctic fox fur colour. Covariation between colour

and other phenotypic traits is well documented (Ducrest et al., 2008; San-Jose & Roulin, 2017) and I argue that it is likely that such covariation may be the driver of selection on fur colour also in the Scandinavian arctic fox population. This demonstrates the need to look further than the most apparent phenotype when attempting to understand the mechanisms of selection in wild populations. More specifically, it becomes apparent that gene mapping can provide valuable insight into the genetic architecture of adaptive traits and other linked traits. Also, when the linked gene that actually affects individual fitness cannot be identified, genetic findings (e.g. linked genes, pleiotropy, genome-wide heterozygosity) can be used to determine knowledge gaps and areas of interest for future research. In the case of my study species, one major issue is the lack of data on other differences between the colour morphs. I have identified several traits that may be affected by the fur colour morph, however, data on differences between the fur colour morphs (e.g. behaviour, metabolism, energy expenditure or immune response) is lacking. Future research may use such traits as a starting point for gaining more insight into selection processes that occur in the arctic fox.

CONCLUSION

In this study, I have successfully mapped the genetic architecture of arctic fox fur colour, a well-known phenotypic trait in a wild non-model species and identified the sole causal gene, MC1R. Using analyses of selection, I have demonstrated that there is selection for the heterozygote blue fur colour genotype that showed through higher survival and reproduction in females. Additionally, I have shown a gradual increase in blue fur colour genotypes in two large Norwegian subpopulations. Even though I was not able to directly identify other traits linked to arctic fox fur colour, the genome-wide scale of the study has provided detailed information on linkage disequilibrium between the fur colour gene and several genes that are likely to segregate with the causative colour morph gene, as well as traits likely linked to MC1R through pleiotropy. This also shows the large potential of genome-wide analyses when studying adaptive traits in wild populations. In the case examined here, it has become clear that there might be phenotypic differences between arctic fox colour morphs that have not yet been discovered. Knowledge about the genetic architecture and genes linked to MC1R allows us to better understand the genetic mechanisms behind eco-evolutionary changes in Scandinavian arctic fox populations.

REFERENCES

Adalsteinsson, S., Hersteinsson, P. & Gunnarsson, E. (1987). Fox colors in relation to colors in mice and sheep. *Journal of Heredity, 78*(4), 235-237. doi:10.1093/oxfordjournals.jhered.a110373

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410. doi:10.1016/S0022-2836(05)80360-2

Andersen, R., Linnell, J. D. C., Landa, A. & Strand, O. (2003). Fjellrev i Norge 2003 - Overvåkingsrapport. [Arctic fox in Norway 2003 - Monitoring report]. *NINA Minirapport, 037*, 15.

Andrew, R. L., Bernatchez, L., Bonin, A., Buerkle, C. A., Carstens, B. C., Emerson, B. C., Garant, D., Giraud, T., Kane, N. C. & Rogers, S. M. (2013). A road map for molecular ecology. *Molecular Ecology, 22*(10), 2605-2626. doi:10.1111/mec.12319

Angerbjörn, A., Eide, N. E., Dalén, L., Elmhagen, B., Hellström, P., Ims, R. A., Killengreen, S., Landa, A., Meijer, T., Mela, M., Niemimaa, J., Norén, K., Tannerfeldt, M., Yoccoz, N. G. & Henttonen, H. (2013). Carnivore conservation in practice: replicated management actions on a large spatial scale. *Journal of Applied Ecology, 50*(1), 59-67. doi:10.1111/1365-2664.12033

ArtDatabanken. (2015). *Rödlistade arter i Sverige 2015*. Retrieved from ArtDatabanken SLU, Uppsala

Atmeh, K., Andruszkiewicz, A. & Zub, K. (2018). Climate change is affecting mortality of weasels due to camouflage mismatch. *Scientific Reports, 8*(1), 7648. doi:10.1038/s41598-018-26057-5

Audet, A. M., Robbins, C. B. & Larivière, S. (2002). *Alopex lagopus*. *Mammalian species*, 1-10. doi:10.1644/0.713.1

Aulchenko, Y. S., Ripke, S., Isaacs, A. & Van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics, 23*(10), 1294-1296. doi:10.1093/bioinformatics/btm108

Ayllon, F., Kjærner-Semb, E., Furmanek, T., Wennevik, V., Solberg, M. F., Dahle, G., Taranger, G. L., Glover, K. A., Almén, M. S., Rubin, C. J., Edvardsen, R. B. & Wargelius, A. (2015). The vgll3 Locus Controls Age at Maturity in Wild and Domesticated Atlantic Salmon (*Salmo salar* L.) Males. *PLOS Genetics, 11*(11), e1005628-e1005628. doi:10.1371/journal.pgen.1005628

Barabash-Nikiforov, I. (1938). Mammals of the Commander Islands and the surrounding sea. *Journal of Mammalogy, 19*(4), 423-429. doi:10.2307/1374226

Barson, N. J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G. H., Fiske, P., Jacq, C., Jensen, A. J., Johnston, S. E., Karlsson, S., Kent, M., Moen, T., Niemelä, E., Nome, T., Næsje, T. F., Orell, P., Romakkaniemi, A., Sægrov, H., Urdal, K., Erkinaro, J., Lien, S. & Primmer, C. R. (2015). Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature, 528*(7582), 405-408. doi:10.1038/nature16062

Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., Mahoney, L., Wood, D., Alperin, E. S. & Rosyara, U. R. (2015). Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry Fragaria× ananassa. *BMC Genomics, 16*(1), 155. doi:10.1186/s12864-015-1310-1

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Beddard, F. E. (1892). *Animal coloration: An account of the principal facts and theories relating to the colours and markings of animals*. London: S. Sonnenschein & Company.

Beißbarth, T. & Speed, T. P. (2004). GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics, 20*(9), 1464-1465. doi:10.1093/bioinformatics/bth088

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8*, 3-62.

Bonser, R. H. (1995). Melanin and the abrasion resistance of feathers. *The Condor, 97*(2), 590-591. doi:10.2307/1369048

Braestrup, F. W. (1941). *A study on the Arctic Fox in Greenland: immigrations, fluctuations in numbers, based mainly on trading statistics*: Reitzel.

Brommer, J. E., Ahola, K. & Karstinen, T. (2005). The colour of fitness: plumage coloration and lifetime reproductive success in the tawny owl. *Proceedings of the Royal Society B: Biological Sciences, 272*(1566), 935-940. doi:10.1098/rspb.2005.3052

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M. & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal, 9*(2), 378-400. doi:10.3929/ethz-b-000240890

Brown, E. A., Pilkington, J. G., Nussey, D. H., Watt, K. A., Hayward, A. D., Tucker, R., Graham, A. L., Paterson, S., Beraldi, D., Pemberton, J. M. & Slate, J. (2013). Detecting genes for variation in parasite burden and immunological traits in a wild population: testing the candidate gene approach. *Molecular Ecology, 22*(3), 757-773. doi:10.1111/j.1365-294X.2012.05757.x

Bush, W. S. & Moore, J. H. (2012). Genome-wide association studies. *PLoS computational biology, 8*(12), e1002822. doi:10.1371/journal.pcbi.1002822
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics, 10*, 421. doi:10.1186/1471-2105-10-421
Caro, T. (2005). The adaptive significance of coloration in mammals. *BioScience, 55*(2), 125-136. doi:10.1641/0006-3568(2005)055[0125:TASOCI]2.0.CO;2
Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience, 4*, 7. doi:10.1186/s13742-015-0047-8
Chesemore, D. L. (1968). Distribution and movements of white foxes in northern and western Alaska. *Canadian Journal of Zoology, 46*(5), 849-854. doi:10.1139/z68-121
Chesemore, D. L. (1970). Notes on the pelage and priming sequence of arctic foxes in northern Alaska. *Journal of Mammalogy, 51*(1), 156-159. doi:10.2307/1378547
Comeault, A. A., Soria-Carrasco, V., Gompert, Z., Farkas, T. E., Buerkle, C. A., Parchman, T. L. & Nosil, P. (2014). Genome-wide association mapping of phenotypic traits subject to a range of intensities of natural selection in *Timema cristinae*. *Am Nat, 183*(5), 711-727. doi:10.1086/675497
Direktoratet for naturforvaltning. (2003). Handlingsplan for fjellrev. [Action plan for the arctic fox]. *Rapport 2003-2*.
Ducrest, A.-L., Keller, L. & Roulin, A. (2008). Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends in Ecology & Evolution, 23*(9), 502-510. doi:10.1016/j.tree.2008.06.001
Eide, N. E., Elmhagen, B., Norén, K., Killengreen, S. T., Wallén, J., Ulvind, K., Landa, A., Ims, R. A., Flagstad, Ø., Ehrich, D. & Angerbjörn, A. (2017). Handlingsplan for fjellrev (*Vulpes lagopus*), Norge-Sverige 2017-2021. [Norwegian-Swedish action plan for the arctic fox (*Vulpes lagopus*), 2017-2021]. *Report M-794 | 2017*.
Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution, 29*(1), 51-63. doi:10.1016/j.tree.2013.09.008
Engen, S., Kvalnes, T. & Sæther, B. E. (2014). Estimating phenotypic selection in age-structured populations by removing transient fluctuations. *Evolution, 68*(9), 2509-2523. doi:10.1111/evo.12456
Engen, S., Sæther, B.-E., Kvalnes, T. & Jensen, H. (2012). Estimating fluctuating selection in age-structured populations. *Journal of Evolutionary Biology, 25*(8), 1487-1499. doi:10.1111/j.1420-9101.2012.02530.x
Fetherston, K. (1947). Geographic variation in the incidence of occurrence of the blue phase of the arctic fox in Canada. *Canadian Field-Naturalist, 61*(1), 15-18.
Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford, England: Clarendon Press.
Framstad, E. (2017). Terrestrisk naturovervåking i 2016: Markvegetasjon, epifytter, smågnagere og fugl. Sammenfatning av resultater. [Terrestrial Ecosystems Monitoring in 2016: Ground vegetation, epiphytes, small mammals and birds. Summary of results.]. *NINA Rapport, 1376*. doi:http://hdl.handle.net/11250/2452619
Gratten, J., Wilson, A., McRae, A., Beraldi, D., Visscher, P., Pemberton, J. & Slate, J. (2008). A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science, 319*(5861), 318-320. doi:10.1126/science.1151182
Hadfield, J. D., Nutall, A., Osorio, D. & Owens, I. P. F. (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of Evolutionary Biology, 20*(2), 549-557. doi:10.1111/j.1420-9101.2006.01262.x
Hasselgren, M., Angerbjörn, A., Eide, N. E., Erlandsson, R., Flagstad, Ø., Landa, A., Wallén, J. & Norén, K. (2018). Genetic rescue in an inbred Arctic fox *Vulpes lagopus* population. *Proceedings of the Royal Society B: Biological Sciences, 285*(1875), 20172814. doi:10.1098/rspb.2017.2814
Hecht, B. C., Campbell, N. R., Holecek, D. E. & Narum, S. R. (2013). Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology, 22*(11), 3061-3076. doi:10.1111/mec.12082
Hedrick, P. W. (2012). What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution, 27*(12), 698-704. doi:10.1016/j.tree.2012.08.012
Henriksen, S. & Hilmo, O. (Eds.). (2015). *Norsk rødliste for arter 2015*. Norway: Artsdatabanken.
Hersteinsson, P. (1989). Population genetics and ecology of different colour morphs of arctic foxes *Alopex lagopus* in Iceland. *Finnish Game Research, 46*, 64-78.
Hersteinsson, P., Angerbjörn, A., Frafjord, K. & Kaikusalo, A. (1989). The arctic fox in Fennoscandia and Iceland: management problems. *Biological Conservation, 49*(1), 67-81. doi:10.1016/0006-3207(89)90113-4

Hoeppner, M. P., Lundquist, A., Pirun, M., Meadows, J. R. S., Zamani, N., Johnson, J., Sundström, G., Cook, A., FitzGerald, M. G., Swofford, R., Mauceli, E., Moghadam, B. T., Greka, A., Alföldi, J., Abouelleil, A., Aftuck, L., Bessette, D., Berlin, A., Brown, A., Gearin, G., Lui, A., Macdonald, J. P., Priest, M., Shea, T., Turner-Maier, J., Zimmer, A., Lander, E. S., di Palma, F., Lindblad-Toh, K. & Grabherr, M. G. (2014). An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*, *9*(3), e91172-e91172. doi:10.1371/journal.pone.0091172

Huisman, J. (2017). Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, *17*(5), 1009-1024. doi:10.1111/1755-0998.12665

Jensen, H., Szulkin, M. & Slate, J. (2014). Molecular quantitative genetics. In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild* (pp. 209-227). Oxford: Oxford University Press.

Johnson, R. C., Nelson, G. W., Troyer, J. L., Lautenberger, J. A., Kessing, B. D., Winkler, C. A. & O'Brien, S. J. (2010). Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics*, *11*(1), 724. doi:10.1186/1471-2164-11-724

Johnston, S. E., Mcewan, J. C., Pickering, N. K., Kijas, J. W., Beraldi, D., Pilkington, J. G., Pemberton, J. M. & Slate, J. (2011). Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, *20*(12), 2555-2566. doi:10.1111/j.1365-294X.2011.05076.x

Johnston, S. E., Orell, P., Pritchard, V. L., Kent, M. P., Lien, S., Niemelä, E., Erkinaro, J. & Primmer, C. R. (2014). Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (Salmo salar). *Molecular Ecology*, *23*(14), 3452-3468. doi:10.1111/mec.12832

Kardos, M., Husby, A., McFarlane, S. E., Qvarnström, A. & Ellegren, H. (2016). Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. *Molecular Ecology Resources*, *16*(3), 727-741. doi:10.1111/1755-0998.12498

Komiya, Y. & Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis*, *4*(2), 68-75. doi:10.4161/org.4.2.5851

Korte, A. & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, *9*(1), 29. doi:10.1186/1746-4811-9-29

Kvalnes, T. (2013). lmf: Functions for estimation and inference of selection in age-structured populations. R package version 1.2. Retrieved from https://CRAN.R-project.org/package=lmf

Landa, A., Flagstad, Ø., Areskoug, V., Linnell, J. D. C., Strand, O., Ulvund, K. R., Thierry, A.-M., Rød-Eriksen, L. & Eide, N. E. (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., Strand, O., Linnell, J. D. C. & Skogland, T. (1998). Home-range sizes and altitude selection for arctic foxes and wolverines in an alpine environment. *Canadian Journal of Zoology*, *76*(3), 448-457. doi:10.1139/z97-209

Lawson, L. P. & Petren, K. (2017). The adaptive genomic landscape of beak morphology in Darwin's finches. *Molecular Ecology*, *26*(19), 4978-4989. doi:10.1111/mec.14166

Linnell, J. D. C., Strand, O. & Landa, A. (1999a). Use of dens by red *Vulpes vulpes* and arctic *Alopex lagopus* foxes in alpine environments: Can inter-specific competition explain the non-recovery of Norwegian arctic fox populations? *Wildlife Biology*, *5*(3), 167-176. doi:10.2981/wlb.1999.021

Linnell, J. D. C., Strand, O., Loison, A., Solberg, E. J. & Jordhøy, P. (1999b). Har fjellreven en framtid i Norge? Statusrapport og forslag til forvaltningsplan. [A future for the Arctic fox in Norway? A status report and action plan]. *NINA oppdragsmelding*, *575*, 1-37.

Lundregan, S. L., Hagen, I. J., Gohli, J., Niskanen, A. K., Kemppainen, P., Ringsby, T. H., Kvalnes, T., Pärn, H., Rønning, B. & Holand, H. (2018). Inferences of genetic architecture of bill morphology in house sparrow using a high-density SNP array point to a polygenic basis. *Molecular Ecology*, *27*(17), 3498-3514. doi:10.1111/mec.14811

Lüdecke, D. (2018). ggeffects: Create Tidy Data Frames of Marginal Effects for 'ggplot' from Model Outputs. *Journal of Open Source Software*, *3*(26). doi:10.21105/joss.00772

Meijer, T., Norén, K., Hellström, P., Dalén, L. & Angerbjörn, A. (2008). Estimating population parameters in a threatened arctic fox population using molecular tracking and traditional field methods. *Animal Conservation*, *11*(4), 330-338. doi:10.1111/j.1469-1795.2008.00188.x

Nachman, M. W., Hoekstra, H. E. & D'Agostino, S. L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences*, *100*(9), 5268-5273. doi:10.1073/pnas.0431157100

Norwegian Polar Institute. Arctic fox (*Vulpes lagopus*). Retrieved from https://www.npolar.no/en/species/arctic-fox/

Poulton, E. B. (1890). *The colours of animals: their meaning and use, especially considered in the case of insects*: D. Appleton.

Prestrud, P. (1991). Adaptations by the arctic fox (*Alopex lagopus*) to the polar winter. *Arctic*, 132-138. doi:10.14430/arctic1529
Prestrud, P. & Nilssen, K. (1992). Fat Deposition and Seasonal Variation in Body Composition of Arctic Foxes in Svalbard. *The Journal of wildlife management, 56*(2), 221-233. doi:10.2307/3808816
Purcell, S. & Chang, C. C. (2019). PLINK v1.90. Retrieved from www.cog-genomics.org/plink/1.9/
Robbins, L. S., Nadeau, J. H., Johnson, K. R., Kelly, M. A., Roselli-Rehfuss, L., Baack, E., Mountjoy, K. G. & Cone, R. D. (1993). Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell, 72*(6), 827-834. doi:10.1016/0092-8674(93)90572-8
Robinson, M. R., Santure, A. W., DeCauwer, I., Sheldon, B. C. & Slate, J. (2013). Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Molecular Ecology, 22*(15), 3963-3980. doi:10.1111/mec.12375
Roff, D. & Fairbairn, D. (2013). The costs of being dark: the genetic basis of melanism and its association with fitness-related traits in the sand cricket. *Journal of Evolutionary Biology, 26*(7), 1406-1416. doi:doi:10.1111/jeb.12150
Roulin, A. (2007). Melanin pigmentation negatively correlates with plumage preening effort in barn owls. *Functional Ecology, 21*(2), 264-271. doi:10.1111/j.1365-2435.2006.01229.x
Roulin, A., Dijkstra, C., Riols, C. & Ducrest, A.-L. (2001). Female- and male-specific signals of quality in the barn owl. *Journal of Evolutionary Biology, 14*(2), 255-266. doi:10.1046/j.1420-9101.2001.00274.x
Roulin, A. & Ducrest, A.-L. (2013). Genetics of colouration in birds. *Seminars in Cell & Developmental Biology, 24*(6), 594-608. doi:10.1016/j.semcdb.2013.05.005
Saloranta, T. (2014). *New version (v.1.1.1) of the seNorge snow model and snow maps for Norway*. Retrieved from Norwegian Water Resources and Energy Directorate, Oslo
San-Jose, L. M., Ducrest, A.-L., Ducret, V., Béziers, P., Simon, C., Wakamatsu, K. & Roulin, A. (2015). Effect of the MC1R gene on sexual dimorphism in melanin-based colorations. *Molecular Ecology, 24*(11), 2794-2808. doi:10.1111/mec.13193
San-Jose, L. M. & Roulin, A. (2017). Genomics of coloration in natural animal populations. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1724), 20160337. doi:doi:10.1098/rstb.2016.0337
Santure, A. W., De Cauwer, I., Robinson, M. R., Poissant, J., Sheldon, B. C. & Slate, J. (2013). Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. *Molecular Ecology, 22*(15), 3949-3962. doi:10.1111/mec.12376
Santure, A. W. & Garant, D. (2018). Wild GWAS—association mapping in natural populations. *Molecular Ecology Resources, 18*(4), 729-738. doi:10.1111/1755-0998.12901
Santure, A. W., Poissant, J., De Cauwer, I., van Oers, K., Robinson, M. R., Quinn, J. L., Groenen, M. A. M., Visser, M. E., Sheldon, B. C. & Slate, J. (2015). Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. *Molecular Ecology, 24*(24), 6148-6162. doi:10.1111/mec.13452
Santure, A. W., Stapley, J., Ball, A. D., Birkhead, T. R., Burke, T. & Slate, J. (2010). On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology, 19*(7), 1439-1451. doi:10.1111/j.1365-294X.2010.04554.x
Satoh, K., Kasai, M., Ishidao, T., Tago, K., Ohwada, S., Hasegawa, Y., Senda, T., Takada, S., Nada, S., Nakamura, T. & Akiyama, T. (2004). Anteriorization of neural fate by inhibitor of β-catenin and T cell factor (ICAT), a negative regulator of Wnt signaling. *Proceedings of the National Academy of Sciences of the United States of America, 101*(21), 8017-8021. doi:10.1073/pnas.0401733101
Schmutz, S. M. & Berryere, T. G. (2007). Genes affecting coat colour and pattern in domestic dogs: a review. *Animal genetics, 38*(6), 539-549. doi:doi:10.1111/j.1365-2052.2007.01664.x
Sellis, D., Callahan, B. J., Petrov, D. A. & Messer, P. W. (2011). Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences, 108*(51), 20666-20671. doi:10.1073/pnas.1114573108
Silva, C. N. S., McFarlane, S. E., Hagen, I. J., Rönnegård, L., Billing, A. M., Kvalnes, T., Kemppainen, P., Rønning, B., Ringsby, T. H., Sæther, B. E., Qvarnström, A., Ellegren, H., Jensen, H. & Husby, A. (2017). Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity, 119*(3), 197-205. doi:10.1038/hdy.2017.29
Sinclair-Waters, M., Ødegård, J., Korsvoll, S. A., Moen, T., Lien, S., Primmer, C. R. & Barson, N. J. (2020). Beyond large-effect loci: large-scale GWAS reveals a mixed large-effect and polygenic architecture for age at maturity of Atlantic salmon. *Genetics Selection Evolution, 52*(1), 9. doi:10.1186/s12711-020-0529-8
Slagsvold, P. (1949). Nedarving av den blå og hvite farge hos polarreven (*Alopex lagopus*). [Inheritance of the blue and white colour in the arctic fox (*Alopex lagopus*)]. *Nord. Vet. Medicin, 1*, 429-441.

Slate, J. (2015). Why I'm wary of candidate gene studies. Retrieved from http://jon-slate.staff.shef.ac.uk/why-im-wary-of-candidate-gene-studies/
Spencer, H. (1896). *The principles of biology* (Vol. 1): D. Appleton.
Stasiak, K., Kondracki, S. & Iwanina, M. (2019). Assessment of chosen semen characteristics of two colour morphs of the Arctic fox *Alopex lagopus* L. *Animal Science Journal, 90*(9), 1120-1126. doi:10.1111/asj.13257
Stinchcombe, J. R. & Hoekstra, H. E. (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity, 100*(2), 158-170. doi:10.1038/sj.hdy.6800937
The UniProt Consortium. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res, 46*(5), 2699-2699. doi:10.1093/nar/gky092
Ulvund, K., Flagstad, Ø., Sandercock, B., Kleven, O., Landa, A. & Eide, N. E. (2019). Fjellrev i Norge 2019. Resultater fra det nasjonale overvåkingsprogrammet for fjellrev. [The arctic fox in Norway 2019. Results from the national monitoring program for the arctic fox]. *NINA Rapport 1737*. doi:hdl.handle.net/11250/2643424
Underwood, L. S. & Reynolds, P. (1980). Photoperiod and fur lengths in the arctic fox (*Alopex lagopus* L.). *International Journal of Biometeorology, 24*(1), 39-48. doi:10.1007/BF02245540
van den Brink, V., Dolivo, V., Falourd, X., Dreiss, A. N. & Roulin, A. (2011). Melanic color-dependent antipredator behavior strategies in barn owl nestlings. *Behavioral Ecology, 23*(3), 473-480. doi:10.1093/beheco/arr213
Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A. & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics, 101*(1), 5-22. doi:10.1016/j.ajhg.2017.06.005
Våge, D. I., Fuglei, E., Snipstad, K., Beheim, J., Landsem, V. M. & Klungland, H. (2005). Two cysteine substitutions in the MC1R generate the blue variant of the arctic fox (*Alopex lagopus*) and prevent expression of the white winter coat. *Peptides, 26*(10), 1814-1817. doi:10.1016/j.peptides.2004.11.040
Våge, D. I., Lu, D., Klungland, H., Lien, S., Adalsteinsson, S. & Cone, R. D. (1997). A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes*. *Nature Genetics, 15*(3), 311-315. doi:10.1038/ng0397-311
Wallén, J., Ulvund, K., Sandercock, B. & Eide, N. E. (2020). Inventering av fjällräv 2019/Overvåking av fjellrev 2019 [Monitoring of the arctic fox 2019]. *Beståndsstatus för fjällräv i Skandinavien/Bestandsstatus for fjellrev i Skandinavia, 1-2020*, 36.
Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D. & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res, 38*(suppl_2), W214-W220. doi:10.1093/nar/gkq537
Wenzel, M. A., James, M. C., Douglas, A. & Piertney, S. B. (2015). Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. *Molecular Ecology, 24*(16), 4175-4192. doi:10.1111/mec.13313
Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer. Retrieved from https://ggplot2.tidyverse.org
Zimova, M., Hackländer, K., Good, J. M., Melo-Ferreira, J., Alves, P. C. & Mills, L. S. (2018). Function and underlying mechanisms of seasonal colour moulting in mammals and birds: what keeps them changing in a warming world? *Biological Reviews, 93*(3), 1478-1498. doi:10.1111/brv.12405
Zimova, M., Mills, L. S., Lukacs, P. M. & Mitchell, M. S. (2014). Snowshoe hares display limited phenotypic plasticity to mismatch in seasonal camouflage. *Proceedings of the Royal Society of London B: Biological Sciences, 281*(1782), 20140029. doi:10.1098/rspb.2014.0029

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY MATERIAL 1

Table S1: Overview over sample sizes of SNP genotyped arctic foxes on Affymetrix and Fludigim SNP arrays and the different analyses conducted in this study.

		Number of genotyped foxes						Analyses				
								GWAS	Genome-wide heterozygosity	Selection analysis		
Subpopulation	**Released foxes (time period)**	**Affymetrix**		**Fluidigm**		**Both SNP arrays**				**Total**	**Affymetrix**	**Fluidigm**
		Total	*Phenotype*	*Total*	*Phenotype*	*Total*	*Phenotype*					
Borga		0	0	18	0	0	0	0	0	26	1	27
Hardangervidda	195 (2009-2018)	144	144	102	65	45	45	189	189	310	189	167
Helags		8	8	29	5	0	0	8	8	49	13	42
Lierne		0	0	18	0	0	0	0	0	19	1	22
Saltfjellet	63 (2006-2015)	57	57	54	23	7	7	64	64	132	65	77
Snøhetta	93 (2007-2010)	159	159	394	340	35	35	194	194	595	197	440
Troms/Reisa		0	0	8	0	0	0	0	0	9	2	11
Varangerhalvøya	67 (2017-2019)	6	6	17	11	18	18	24	24	41	24	35
Other populations (e.g. Swedish subpopulations)		206	198	163	0	4	4	202	210	0	0	0
Total		**580**	**572**	**803**	**444**	**109**	**109**	**681**	**689**	**1181**	**492**	**821**

SUPPLEMENTARY MATERIAL 2

Positions of significant SNPs from the GWA analysis in the arctic fox genome were compared to positions of the best BLAST hit in the dog genome *CanFam 3.1* to assess the fit between the two genomes (Figure S2). The analysis was restrained to SNPs lying on arctic fox scaffold 11 and that matched with a position on dog chromosome 5 during the BLAST. SNPs on other scaffolds or that matched with different dog chromosome would naturally appear as outliers in the Figure.

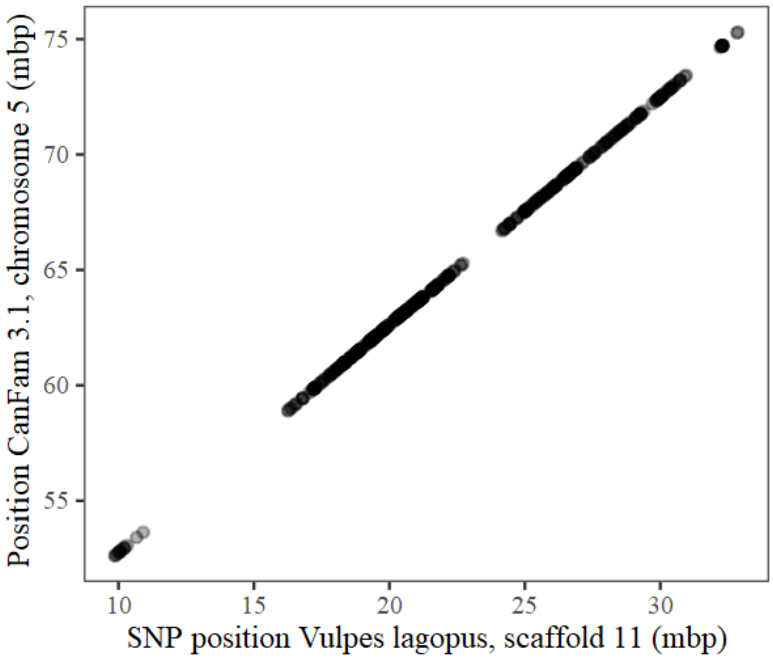


Figure S2: Comparison of positions of SNPs significantly associated with arctic fox fur colour on arctic fox scaffold 11 and dog chromosome 5.

SUPPLEMENTARY MATERIAL 3

Table S3: Non-zero elements (Fecundity and survival) of the projection matrices for females (l_f) and males (l_m)

	Females (l_f)		**Males (l_m)**	
Age class	**Fecundity**	**Survival**	**Fecundity**	**Survival**
1	0.140351	0.688596	0.105485	0.672269
2	0.289855	0.702899	0.253425	0.727891
3	0.731183	0.720430	0.562500	0.765306
4	0.703125	0.750000	0.628571	0.722222
5+	0.430769	0.569231	0.521127	0.506667

SUPPLEMENTARY MATERIAL 4

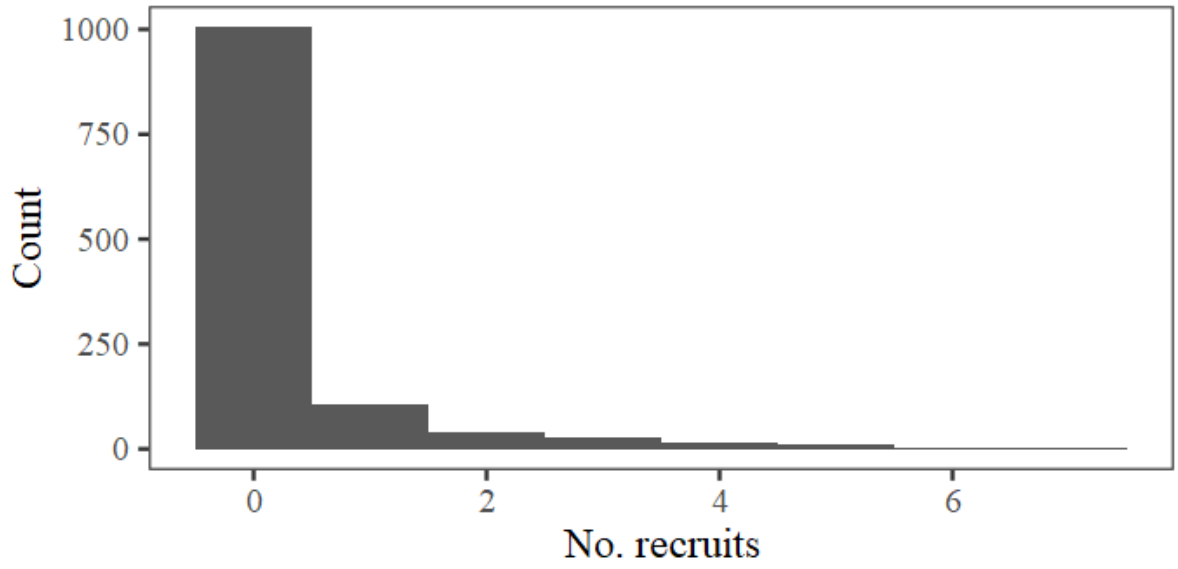


Figure S4: Histogram showing zero-inflation in the fecundity dataset where most observations have zero in fecundity (number of recruits).

SUPPLEMENTARY MATERIAL 5

See attached Excel-file. Detailed Information about all 495 SNPs significantly associated with arctic fox fur colour according to the GWA analysis. Information includes sample size, effect size, minor and major allele, χ^2 test statistics and p-values for all SNPs.

SUPPLEMENTARY MATERIAL 6

Table S6: Names of 155 genes that are located close (<20 kbp) to a SNP significantly associated with arctic fox fur colour.

Distance to closest significant SNP				
Intragenic	**Intragenic**	**Intragenic**	**<10 kbp**	**<20 kbp**
ACOT7	HSD17B2	PER3	CBFA2T3	APRT
ACSF3	HSDL1	PIEZO1	CDH15	ATP2C2
AJAP1	ICMT	PIK3CD	CENPBD1	CDT1
ANKRD11	IL34	PKD1L2	CENPN	CFDP1
BANP	IRF8	PLCG2	CTU2	CYBA
C5H16orf74	KCNAB2	PPAP2B	GALNS	ENO1
C5H1orf168	KIAA0513	PRKAA2	GAS8	FAM92B
C8A	KIF1B	RBP7	IL17C	GPR153
CA5A	KLHL21	RERE	LOC100682843	HES2
CA6	KLHL36	RNF207	LOC100683814	HES3
CAMTA1	LOC100682766	SEMA4F	LOC100684376	HSBP1
CDH13	LOC100683117	SF3B3	LOC100688050	LOC102153218
CDYL2	LOC102152359	SLC25A33	LOC100688505	LOC102153763
CHD5	LOC102154063	SLC38A8	LOC102151626	LOC102155731
CMC2	LOC102154525	SLC45A1	LOC102152301	LOC102155776
CMIP	LOC102155075	SPIRE2	LOC102152587	LOC102156020
CNTNAP4	LOC102155268	SPSB1	LOC102153284	MLYCD
COG4	LOC102156069	TCF25	LOC102154293	PARK7
COTL1	LOC102156119	TLDC1	LOC102154784	PHF13
CRISPLD2	LOC102156165	TMEM170A	LOC102156208	RNF166
CTNNBIP1	LOC102156251	TMEM201	LOC102156493	RPL22
DBNDD1	LOC102156622	TNFRSF9	LOC102157088	TAF1C
DEF8	LOC479600	UBE4B	LOC489633	TUBB3
DNAAF1	LOC489640	USP10	LOC489638	
DNAJC11	LOC489647	USP18	MC1R	
ERRFI1	LOC489707	VAMP3	NECAB2	
FANCA	LZIC	VAT1L	PABPN1L	
FUK	MBTPS1	WFDC1	SCAMPER	
GAN	MPHOSPH6	WWOX	SDR42E1	
GINS2	NMNAT1	ZBTB48	SLC22A31	
GPR157	NOL9	ZC3H18	TMEM231	
GSE1	NPHP4	ZDHHC7	TRAPPC2L	
H6PD	OSGIN1		UTS2	
			ZNF276	

SUPPLEMENTARY MATERIAL 7

Linkage disequilibrium (LD) decay was investigated in arctic fox scaffold 11 since this was the scaffold of largest interest in this study. Pairwise LD (r^2) was calculated for all SNPs on scaffold 11 using PLINK 1.90 (Chang et al., 2015; Purcell & Chang, 2019). Mean LD was then calculated for 5 kbp bins for SNPs closer than 100 kbp and for 100 kbp bins spanning the complete scaffold.

Mean LD (r^2) decreased quickly from 0.33 (SNP distance up to 5 kbp) to 0.26 (SNP distance 5-10 kbp) (Figure S7a). At a SNP distance of roughly 10 000 kbp, r^2 values below 0.05 are reached (Figure S7b).

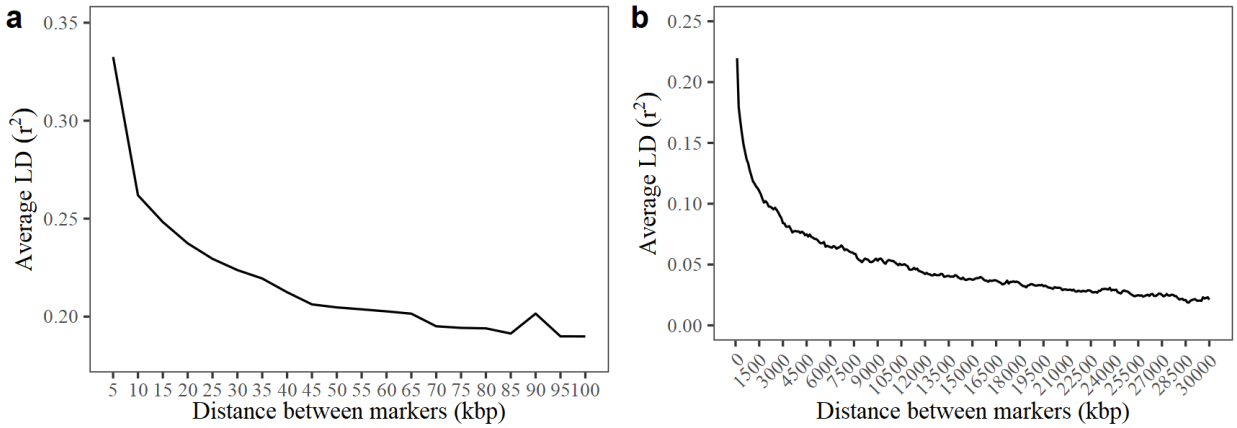


Figure S7: Linkage disequilibrium (r^2) decay in arctic fox scaffold 11. (a) shows mean LD between SNPs closer than 100 kbp based on 5 kbp bins. (b) shows mean LD for SNPs spanning the entire scaffold based on 100 kbp bins.

SUPPLEMENTARY MATERIAL 8

MC1R sequence data from whole-genome sequenced individuals

In order to provide additional support for MC1R being the causative gene for Arctic fox fur colour, we examined the sequence information for MC1R of 12 whole-genome sequenced individuals. These whole-genome sequences were sampled from across the global arctic fox distribution and were used in the development of the custom Affymetrix Axiom 702k SNP-array with 507 000 arctic fox specific single nucleotide polymorphisms (SNPs) (Hagen et al. (in prep)). Of the 12 sequenced individuals, 11 were known to be of the white colour morph and one individual was known to be of the blue colour morph. Using BLAST information on position of Arctic fox SNPs in the dog reference genome (see methods above) and corresponding position on Arctic fox scaffold 11 (see GWA results presented below) in addition to position of MC1R in the dog reference genome and distance in base pairs from SNPs to MC1R, we located the position of the MC1R gene and the two causative SNPs as described in Våge et al. (2005) in scaffold 11 of the already developed Arctic fox sequence mappings (mapping parameters described in Hagen et al., in prep). The MC1R sequence of all 12 individuals were scrutinized for SNPs along the length of the gene.

Only the one individual with the blue morph was found to have the alleles that produce the blue morph caused by a glycine to cysteine substitution in position 5 of the MC1R protein and a phenylalanine to cysteine substitution in position 280 of the MC1R protein as described in Våge et al. (2005). The individual was heterozygous for the two SNPs. Several other synonymous (not affecting the protein sequence) SNPs were found in the MC1R sequence of the 12 individuals but only the SNPs described in Våge et al. (2005) had the property of having a different genotype in the one individual with the blue morph compared to the 11 individuals with the white morph. All other SNPs in the MC1R gene had the same genotype in one or more white foxes and in the blue fox.

SUPPLEMENTARY MATERIAL 9

Table S9: Chi-square test statistics and p values for likelihood ratio tests conducted between additive and interaction models including different environmental variables. Response variables are given in italic. Interactions significant or near significant at the 0.05 significance level are given in bold.

Variable	**χ^2 (df)**	**p value**	**Variable**	**χ^2 (df)**	**p value**
Fitness females			*Recruitment probability*		
Origin	1.10 (1)	0.294	Rodent phase t	0.76 (3)	0.859
Rodent phase t	1.96 (3)	0.581	Rodent phase t+1	0.76 (3)	0.859
Rodent phase t+1	2.73 (3)	0.435	First snow fall t	0.88 (1)	0.350
Rodent phase t-1	2.62 (3)	0.454	Last snow fall t+1	1.16 (1)	0.282
First snow fall t	0.15 (1)	0.698	*Adult survival*		
First snow fall t-1	0.01 (1)	0.932	Origin	1.64 (1)	0.201
Last snow fall t	0.20 (1)	0.658	**Rodent phase t**	**7.36 (3)**	**0.061**
Last snow fall t+1	0.35 (1)	0.555	Rodent phase t+1	5.38 (3)	0.146
Fitness males			First snow fall t	0.48 (1)	0.488
Origin	0.35 (1)	0.552	Last snow fall t+1	0.01 (1)	0.917
Rodent phase t	1.07 (3)	0.784	*Fecundity*		
Rodent phase t+1	0.49 (3)	0.921	Origin	0.01 (1)	0.926
Rodent phase t-1	0.65 (3)	0.885	**Rodent phase t**	**9.32 (3)**	**0.025**
First snow fall t	0.14 (1)	0.705	**Rodent phase t-1**	**6.31 (3)**	**0.097**
First snow fall t-1	0.002 (1)	0.962	First snow fall t-1	1.00 (1)	0.318
Last snow fall t	0.15 (1)	0.702	Last snow fall t	0.57 (1)	0.450
Last snow fall t+1	0.60 (1)	0.437	*Breeding probability*		
			Origin	0.89 (1)	0.345
			Rodent phase	2.25 (3)	0.522
			Rodent phase t-1	2.07 (3)	0.559
			First snow fall t-1	0.20 (1)	0.658
			Last snow fall t	0.02 (1)	0.888

SUPPLEMENTARY MATERIAL 10

Table S10: Summary of gene functions of 42 genes located closer than 10 kbp to SNP that is significantly associated with arctic fox fur colour and in high LD (r^2 >= 0.5) with the SNP most associated with arctic fox fur colour. Gene functions were retrieved from UniProt Knowledgebase (UniProtKB) (The UniProt Consortium, 2018)

Gene	Function	Review status
CTNNBIP1	Negative regulator of Wnt signalling pathway	Dog
MC1R	Receptor for MSH (alpha, beta) and ACTH. Involved in the melanocortin system that regulates melanin-based colouration	Dog
SCAMPER	Calcium regulation	Dog
ACSF3	Catalyzes the initial reaction in intramitochondrial fatty acid synthesis	Human
ANKRD11	Chromatine regulator which modulates histone acetylation and gene expression in neural precursor cells	Human
BANP	Involved in T-cell development and cell cycle arrest.	Human
CAMTA1	Transcriptional activator. May act as a tumor suppressor.	Human
CDH15	Calcium-dependent cell adhesion proteins	Human
FANCA	DNA repair protein that may operate in a postreplication repair or a cell cycle checkpoint function.	Human
GPR157	Contributes to neurogenesis	Human
GSE1	Genetic suppressor element 1. Function largely unknown, found in breast cancer tissue	Human
H6PD	Glucose metabolic process	Human
KIAA0513	Uncharacterized protein	Human
KIF1B	Motor for anterograde transport of mitochondria	Human
KLHL21	Involved in efficient chromosome alignment and cytokinesis	Human
KLHL36	Probable substrate-specific adapter of an E3 ubiquitin-protein ligase complex which mediates the ubiquitination and subsequent proteasomal degradation of target proteins	Human
LOC100682766	No information available	NA
LOC100683814	No information available	NA
LOC100684376	No information available	NA
LOC102151626	No information available	NA
LOC102153284	No information available	NA
LOC102155268	No information available	NA
LOC479600	No information available	NA
LOC489638	No information available	NA
LZIC	Beta-catenin binding	Human
NMNAT1	Catalyses formation and pyrophosphorolytic cleavage of NAD^+. Involved in ATP synthesis in nucleus.	Human
NOL9	rRNA processing	Human
PER3	Part of the circadian clock. Not essential for the circadian rhythms maintenance. important role in sleep-wake timing and sleep homeostasis probably through the transcriptional regulation of sleep homeostasis-related genes, without influencing circadian parameters.	Human
PIEZO1	Component of a mechanosensitive channel required for rapidly adapting mechanically activated (MA) currents	Human
PIK3CD	Phosphorylation of PtdIns(4,5)P2 to PIP3. Involved in immune responses. Plays role in B-cell development and function. Mediates TCR signalling events at the immune synapse.	Human

RERE	Plays a role as a transcriptional repressor during development. May play a role in the control of cell survival.	Human
SLC22A31	Organic anion transporter that mediates the uptake of ions	Human
SLC25A33	Mitochondrial transporter. Participates in mitochondrial genome maintenance.	Human
TCF25	May play a role in cell death control. Acts as a transcriptional repressor.	Human
TMEM201	Involved in nuclear movement during fibroblast polarization and migration.	Human
UBE4B	E3 ligase. May function as E4 ligase mediating assembly of polyubiquitin chains. May regulate myosin assembly in striated muscles.	Human
USP18	Involved in the regulation of inflammatory response to interferon type 1	Human
VAMP3	Vesicular transport from the late endosomes to the trans-golgi network	Human
ZBTB48	Regulator of telomere length	Human
ZC3H18	Zinc finger CCCH domain-containing protein 18	Human
ZDHHC7	Palmitoyltransferase with broad specificity	Human
ZNF276	May be involved in transcriptional regulation.	Human

SUPPLEMENTARY MATERIAL 11

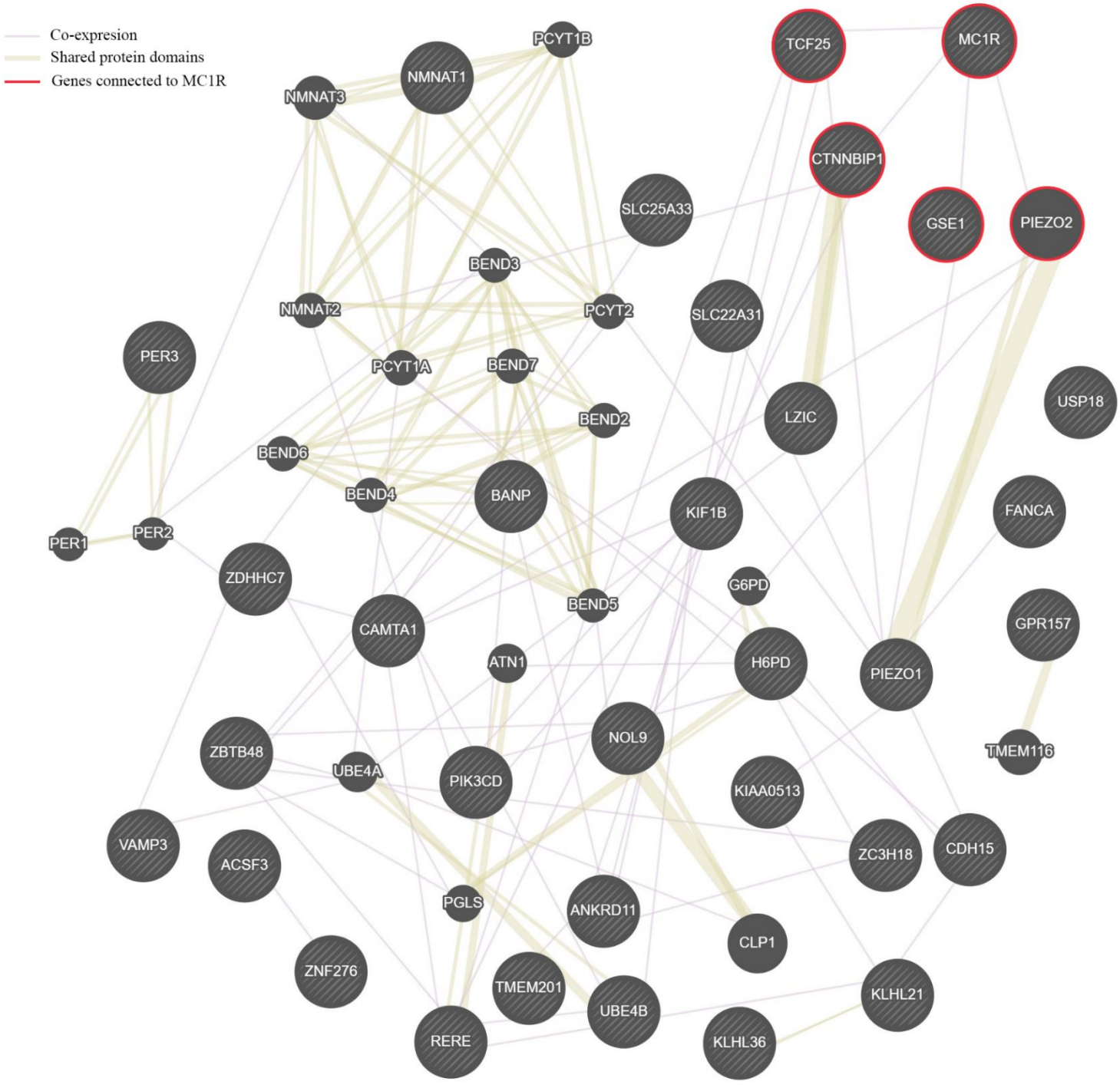
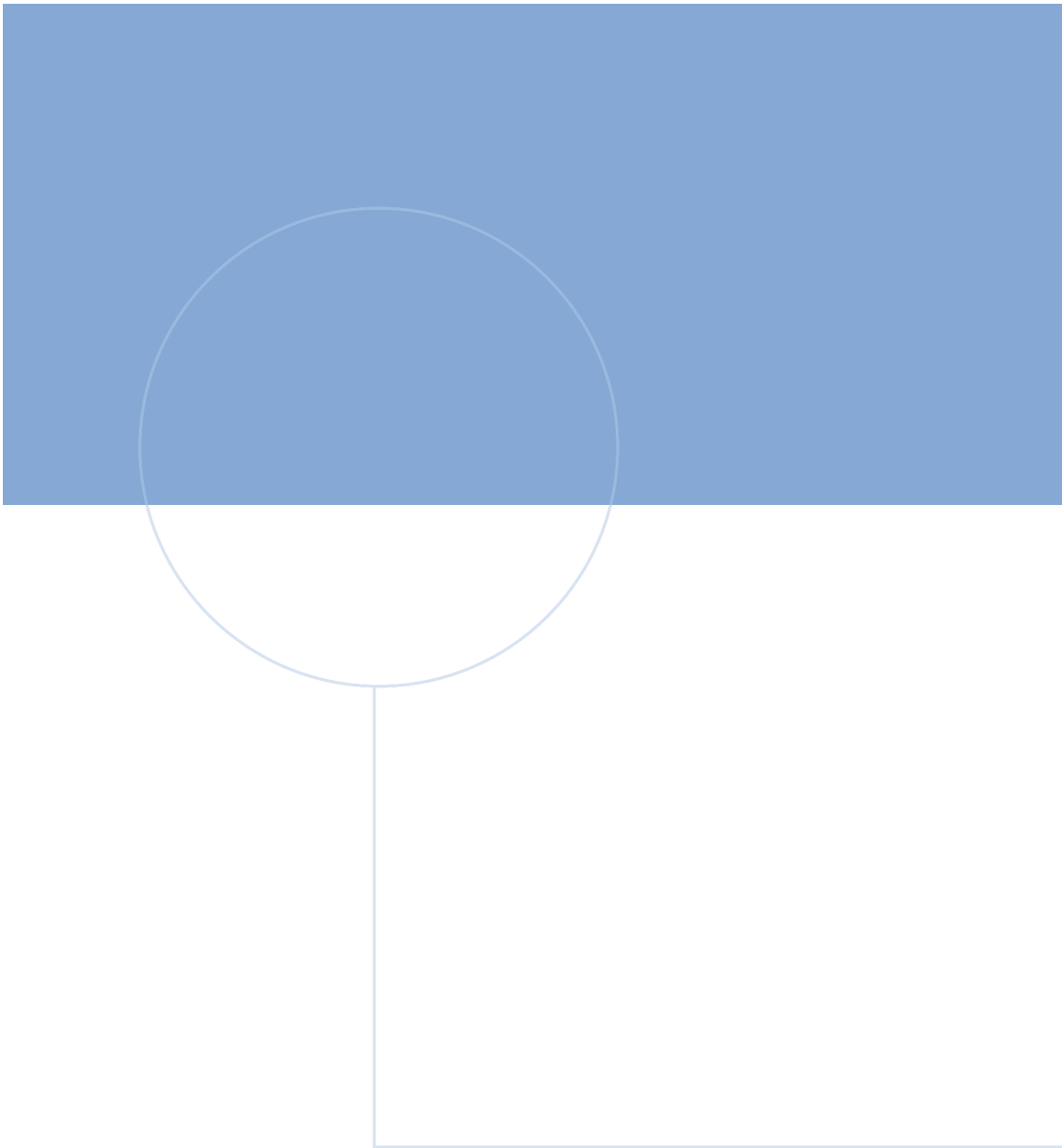


Figure S11: Network of co-expression (thin purple lines) and shared protein domains (thick brown lines) for 30 genes that were found within 10 kbp of a SNP significantly associated with arctic fox fur colour and in high LD (r^2 >= 0.5) with the SNP most associated with arctic fox fur colour. Genes connected to MC1R are positioned in the upper right corner and marked with a red ring. Network produced with GeneMANIA (Warde-Farley et al., 2010).



NTNU

Kunnskap for en bedre verden