Supplementary material

**Table of contents**

Supplementary material 1 – Study area and sample sizes 3

Figure S1 3

Table S1 4

Table S2 4

Table S3 5

Supplementary material 2 – SNP genotyping and Quality Control 6

Supplementary material 3 – GWAS and BLAST 6

Figure S2 6

Figure S3 7

Figure S4 7

Figure S5 8

Figure S6 9

Table S4 9

Supplementary material 4 – Genetic parentage analysis 9

Supplementary material 5 – Projection matrices 10

Table S5 10

Supplementary material 6 – Zero-inflation in fecundity variable 10

Figure S7 10

Supplementary material 7 – Selection 10

Table S6 12

Table S7 14

Table S8 15

Figure S8 16

Table S9 16

Table S10 17

Table S11 17

Figure S9 18

Table S12 18

Table S13 19

Supplementary material 8 – Details on significant GWAS SNPs 19

Table S14 19

Supplementary material 9 – Fitness GWAS 20

Figure S10 21

Supplementary material 10 – Gene analyses 22

Table S15: 22

Table S16 23

Table S17 24

Figure S11 26

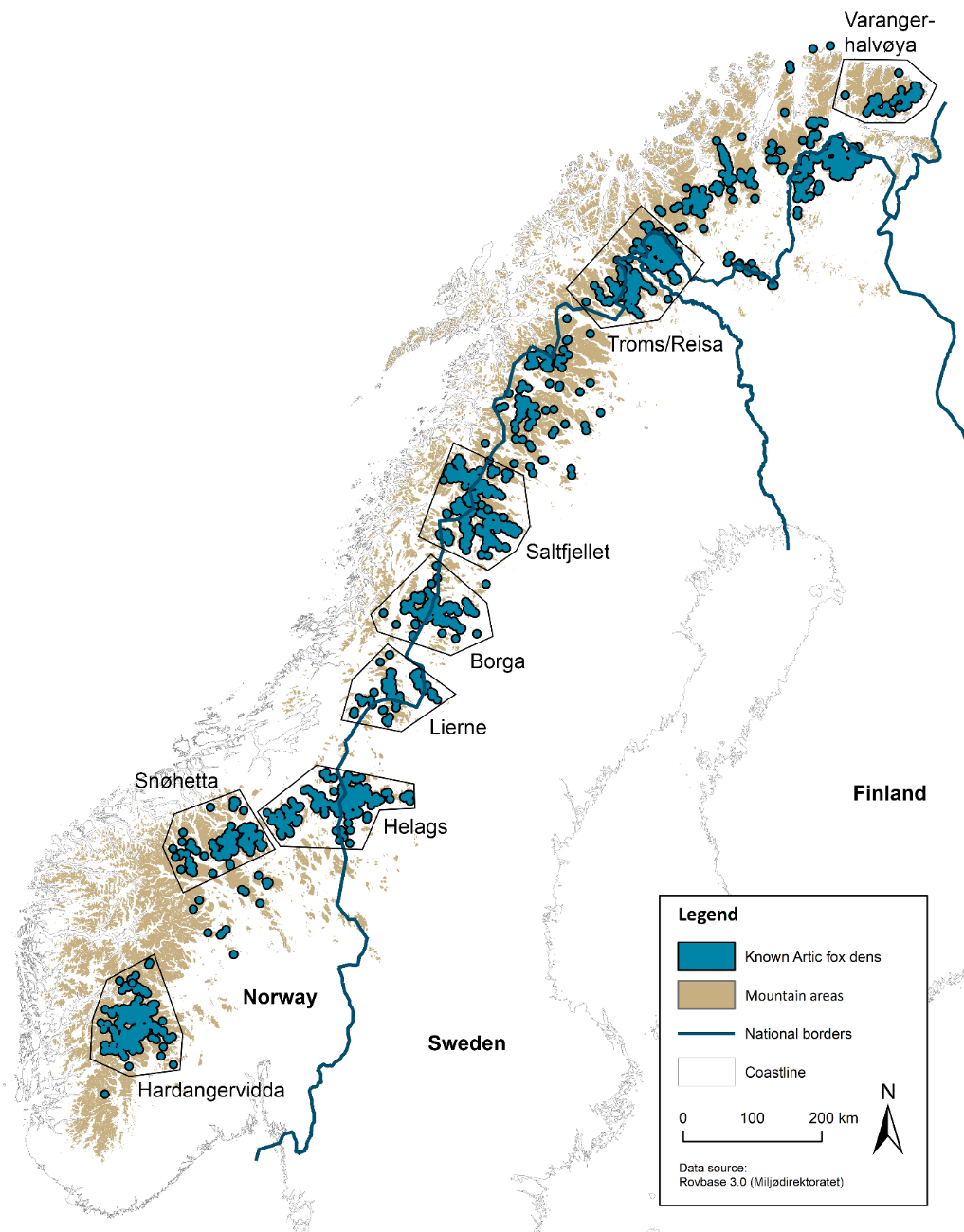
Supplementary material 11 – LD decay 27

Figure S12 27

Supplementary material 12 – MC1R sequence data from whole-genome 28

Supplementary material – References 29

Supplementary material 1 – Study area and sample sizes





CC

TC

TT

**Arctic fox fur colour genotype**

Figure S1: Overview over known historic arctic fox den sites in Fennoscandia. Polygons mark the subpopulations used in this study. Pie charts show the distribution of arctic fox fur colour genotypes in the different subpopulations.

Table S1: Overview over sample sizes of arctic foxes SNP-genotyped on custom Affymetrix and Fluidigm SNP-arrays across subpopulations. Columns give first the total number and then the number of fox individuals with recorded fur colour phenotype. The right side gives the sample sizes across subpopulations for the two genetics analyses conducted in this study (arctic fox fur colour GWAS and genome-wide heterozygosity). Effectively, sample sizes for these analyses consist of all fox individuals genotyped on the Affymetrix array and in the case of the GWAS, a subset of those individuals that have a fur colour phenotype recorded. DNA for the Affymetrix SNP genotyping was extracted from ear tissue using the Qiagen DNeasy 96 Blood & Tissue Kit (Qiagen GmbH, Hilden, Germany). DNA genotyped on the Fluidigm 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).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Only Affymetrix** | | **Only Fluidigm** | | **Both SNP arrays** | | **GWAS** | **GW heterozygosity** |
|  |  | *Total* | *Phenotype* | *Total* | *Phenotype* | *Total* | *Phenotype* |  |  |
| **Subpopulation** | **Varangerhalvøya** | 7 | 7 | 47 | 11 | 18 | 18 | 25 | 25 |
| **Troms/Reisa** | 1 | 1 | 15 | 0 | 0 | 0 | 1 | 1 |
| **Saltfjellet** | 117 | 114 | 79 | 23 | 6 | 6 | 120 | 123 |
| **Borga** | 32 | 32 | 40 | 0 | 0 | 0 | 32 | 32 |
| **Lierne** | 1 | 1 | 23 | 0 | 0 | 0 | 1 | 1 |
| **Helags** | 85 | 85 | 46 | 5 | 0 | 0 | 85 | 85 |
| **Snøhetta** | 138 | 138 | 408 | 340 | 27 | 27 | 165 | 165 |
| **Hardangervidda** | 131 | 131 | 139 | 65 | 45 | 45 | 176 | 176 |
| **Captivity** | 63 | 63 | 5 | 0 | 13 | 13 | 76 | 76 |
| **Unknown** | 5 | 0 | 1 | 0 | 0 | 0 | 0 | 5 |
|  | **Total** | **580** | **572** | **803** | **444** | **109** | **109** | **681** | **689** |

Table S2: Overview about the release of captive bred arctic fox individuals into wild Norwegian subpopulations as part of the Norwegian captive breeding programme.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Released foxes** | **Time period** |
| **Subpopulation** | **Varangerhalvøya** | 67 | 2017-2019 |
| **Troms/Reisa** | 0 |  |
| **Saltfjellet** | 63 | 2006-2015 |
| **Borga** | 0 |  |
| **Lierne** | 0 |  |
| **Helags** | 0 |  |
| **Snøhetta** | 93 | 2007-2010 |
| **Hardangervidda** | 195 | 2009-2018 |
|  | **Total** | **418** |  |

Table S3: Sample sizes for the selection analysis broken down into arctic fox fur colour genotype, sex and origin (CB = captive born; WB = wild born) across subpopulations (ordered after latitude). The right side of the table gives total sample sizes for genotype, sex and origin across subpopulations. Numbers in (a) are individuals, numbers in (b) are annual observations (i.e. one individual being observed for three years will result in three annual observations).

(a)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Genotype** | **CC** | | | | | **TC** | | | | **TT** | | | |  |  |  |  |  |  |  |  |
|  | **Sex** | Female | | Male | | NA | Female | | Male | | Female | | Male | |  | **Genotype** | | | **Sex** | | **Origin** | |
|  | **Origin** | *CB* | *WB* | *CB* | *WB* | *WB* | *CB* | *WB* | *CB* | *WB* | *CB* | *WB* | *CB* | *WB* | **Total** | **CC** | **TC** | **TT** | **Female** | **Male** | **CB** | **WB** |
| **Subpopulation** | **Varangerhalvøya** | 14 | 2 | 10 | 3 | --- | 6 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | **41** | 29 | 12 | 0 | 22 | 19 | 36 | 5 |
| **Troms/Reisa** | 0 | 4 | 0 | 1 | --- | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 1 | **9** | 5 | 3 | 1 | 6 | 3 | 0 | 9 |
| **Saltfjellet** | 27 | 19 | 30 | 15 | --- | 2 | 15 | 4 | 11 | 0 | 4 | 0 | 5 | **132** | 91 | 32 | 9 | 67 | 65 | 63 | 69 |
| **Borga** | 0 | 9 | 0 | 2 | --- | 0 | 6 | 0 | 8 | 0 | 0 | 0 | 1 | **26** | 11 | 14 | 1 | 15 | 11 | 0 | 26 |
| **Lierne** | 0 | 9 | 0 | 8 | --- | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | **19** | 17 | 2 | 0 | 10 | 9 | 0 | 19 |
| **Helags** | 2 | 18 | 3 | 23 | --- | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | **49** | 46 | 3 | 0 | 22 | 27 | 5 | 44 |
| **Snøhetta** | 36 | 198 | 50 | 173 | 2 | 3 | 52 | 6 | 72 | 0 | 1 | 0 | 2 | **595** | 459 | 133 | 3 | 290 | 303 | 95 | 500 |
| **Hardangervidda** | 64 | 53 | 85 | 51 | --- | 9 | 16 | 13 | 17 | 1 | 0 | 0 | 1 | **310** | 253 | 55 | 2 | 143 | 167 | 172 | 138 |
|  | **Total** |  | | | |  |  | | | |  | | | | **1181** | **911** | **254** | **16** | **575** | **604** | **371** | **810** |

(b)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Genotype** | **CC** | | | |  | **TC** | | | | **TT** | | | |  |  |  |  |  |  |  |  |
|  | **Sex** | Female | | Male | | NA | Female | | Male | | Female | | Male | |  | **Genotype** | | | **Sex** | | **Origin** | |
|  | **Origin** | *CB* | *WB* | *CB* | *WB* | *WB* | *CB* | *WB* | *CB* | *WB* | *CB* | *WB* | *CB* | *WB* | **Total** | **CC** | **TC** | **TT** | **Female** | **Male** | **CB** | **WB** |
| **Subpopulation** | **Varangerhalvøya** | 19 | 8 | 12 | 14 | --- | 9 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | **71** | 53 | 18 | 0 | 36 | 35 | 49 | 22 |
| **Troms/Reisa** | 0 | 12 | 1 | 4 | --- | 0 | 4 | 1 | 2 | 0 | 0 | 0 | 1 | **25** | 17 | 7 | 1 | 16 | 9 | 2 | 23 |
| **Saltfjellet** | 40 | 42 | 49 | 25 | --- | 6 | 41 | 9 | 28 | 0 | 11 | 0 | 7 | **258** | 156 | 84 | 18 | 140 | 118 | 104 | 154 |
| **Borga** | 0 | 17 | 0 | 10 | --- | 0 | 13 | 0 | 13 | 0 | 0 | 0 | 6 | **59** | 27 | 26 | 6 | 30 | 29 | 0 | 59 |
| **Lierne** | 0 | 41 | 0 | 39 | --- | 0 | 4 | 1 | 2 | 0 | 0 | 0 | 0 | **87** | 80 | 7 | 0 | 45 | 42 | 1 | 86 |
| **Helags** | 4 | 37 | 5 | 32 | --- | 0 | 6 | 12 | 2 | 0 | 0 | 0 | 0 | **98** | 78 | 20 | 0 | 47 | 51 | 21 | 77 |
| **Snøhetta** | 92 | 339 | 132 | 299 | 2 | 13 | 90 | 23 | 126 | 0 | 1 | 0 | 3 | **1120** | 864 | 252 | 4 | 535 | 583 | 260 | 860 |
| **Hardangervidda** | 134 | 85 | 193 | 83 | --- | 25 | 32 | 28 | 26 | 1 | 0 | 0 | 3 | **610** | 495 | 111 | 4 | 277 | 333 | 381 | 229 |
|  | **Total** |  | | | |  |  | | | |  | | | | **2328** | **1770** | **525** | **33** | **1126** | **1200** | **818** | **1510** |

Supplementary material 2 – SNP genotyping and Quality Control

The custom Affymetrix Axiom 702k SNP-array was designed as a two-species SNP array for arctic fox (*Vulpes lagopus*) and red fox (*Vulpes vulpes*), where around 500 000 SNPs were arctic fox specific and around 200 000 SNPs were red fox specific. The development of the array is described in Hagen *et al.*, (in prep.). 731 arctic fox individuals were genotyped on the custom Affymetrix Axiom 702k SNP-array at CIGENE (Ås, Norway). Of these, 701 individuals produced high quality genotypes, whereas 30 individuals failed. Only SNPs classified as high poly-resolution (classification performed by CIGENE) were used in downstream analysis.

The software PLINK 1.90 [1, 2] and *GenABEL* R package [3] were used for data quality control (QC). 1 632 SNPs and 12 individuals were removed due to high Mendelian errors (>10% and >5% error rate respectively). 448 SNPs were discarded due to low minor allele frequency (MAF < 0.01). No SNPs or individuals were excluded due to low call rate (<95%) or extremely high level of heterozygosity relative to HWE expectations (FDR<1%). Eight individuals were excluded due to unknown phenotype.

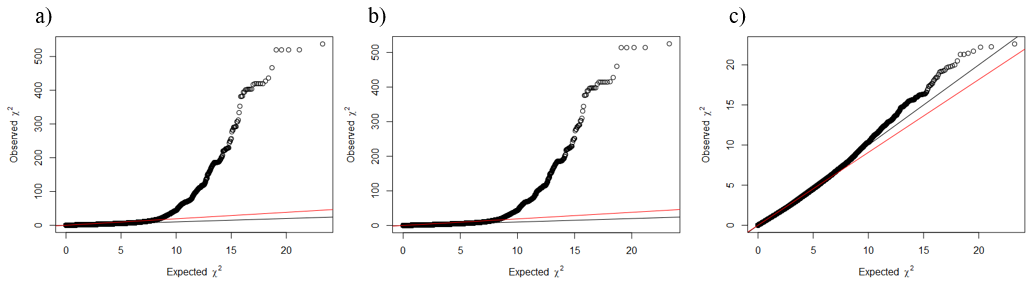
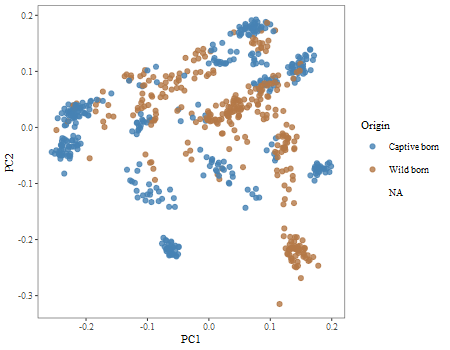
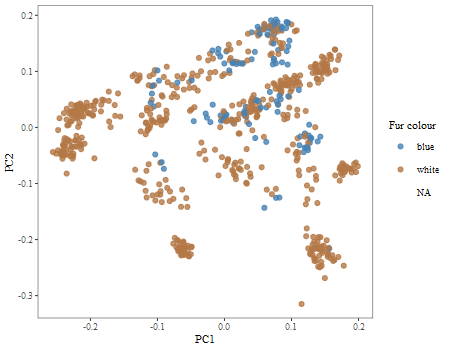
Supplementary material 3 – GWAS and BLAST

Figure S2: a) Quantile-quantile (Q-Q) plot of the p-value distribution obtained from the GWAS for arctic fox fur colour (λ=1.92). b) Q-Q plot for analysis including principal components (PCs) from multidimensional scaling (MDS, λ=1.902) c) Q-Q plot after scaffolds containing significant SNPs were removed (λ=0.91). Black lines show the 1 to 1 slopes expected under the null hypothesis of no genomic inflation. The red lines show the fitted slopes.



b)

a)

Figure S3: Cluster plot showing the two first principal component axes obtained through classical multidimensional scaling (MDS) of the data underlying the arctic fox fur colour GWAS. Different colours show a) the fur colour morphs and b) the origin of the arctic foxes.

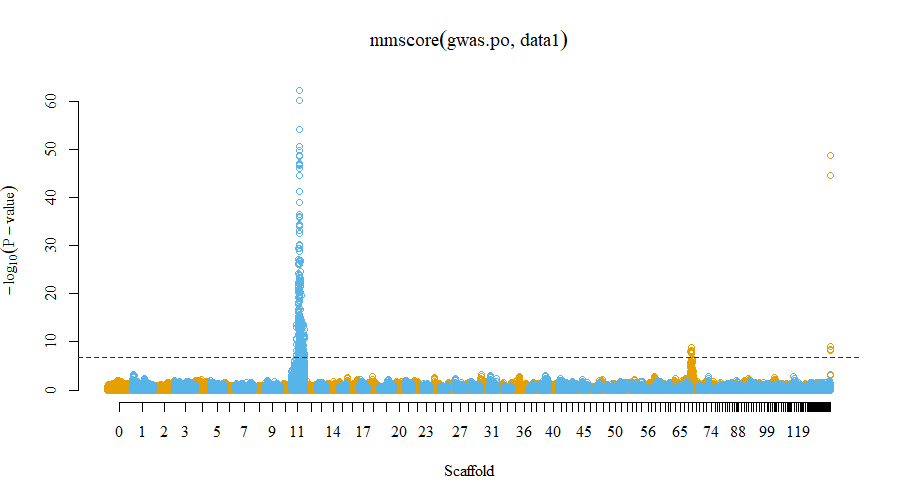


Figure S4: Manhattan plot showing results of genome-wide association study of arctic fox fur colour using information on 359 219 SNPs typed in 681 individuals. P-values are given on a negative log scale. The significance threshold after Bonferroni correction is shown with a dashed horizontal line. Because the arctic fox genome is not assembled into chromosomes, colours alternate between scaffolds and not chromosomes. 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.

For the BLAST 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 was likely to introduce any positional biases, positions of significant SNPs from the GWA analysis on the arctic fox scaffolds were compared to their respective positions in the dog genome *CanFam 3.1* based on the best BLAST hit. The SNPs occurred in the same order in both species (Figure S5) indicating a good fit between the two genomes. 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 (n=469; Table S14). SNPs on other scaffolds or that matched with different dog chromosomes would naturally appear off the diagonal in Figure S5*.*

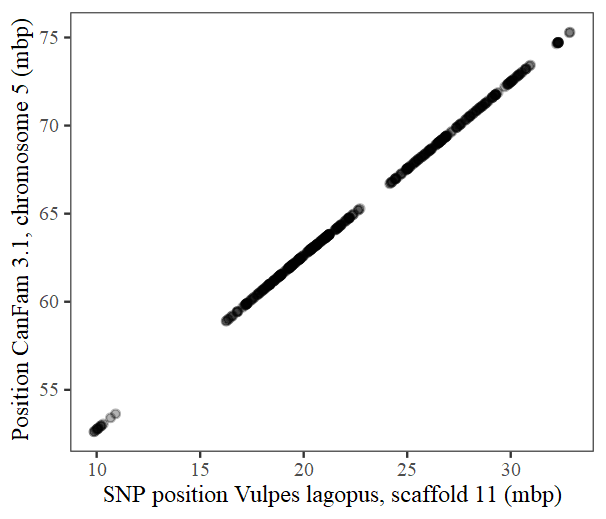
****

Figure S5: Comparison of positions of SNPs significantly associated with arctic fox fur colour on arctic fox scaffold 11 and dog chromosome 5.

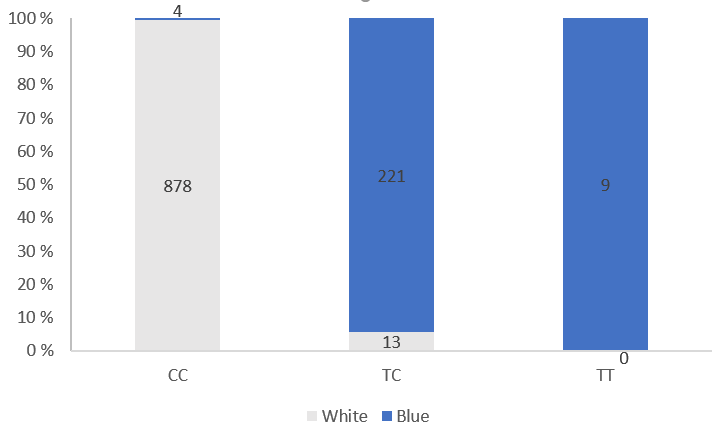


Figure S6: Comparison of MC1R genotypes (as represented by SNP AX-176934441; CC, TC and TT) and arctic fox colour morphs (white and blue). Absolute numbers of cases are given inside the bars. N=1125. This sample size is lower than for the selection analyses (n=1181; Table S3) since the selection analyses (in contrast to this comparison) did not require arctic fox fur colour phenotype to be recorded (i.e., individuals only recorded from scat samples will not have their fur colour phenotype recorded).

Table S4: 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 |

Supplementary material 4 – Genetic parentage analysis

In cases where the birth year of an individual was unknown (n=205), we 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, we assumed to be a juvenile born the same year. This threshold was chosen to coincide roughly with the emergence of pups from the den. Parentage was determined for 1 497 individuals with known or assumed birth year and genotype based on 85 autosomal SNPs that were typed using either our custom Affymetrix SNP-array or our custom Fluidigm SNP-array. Default settings were used in the pedigree construction in the R package *Sequoia* [4]*,* 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 pedigree was checked against known parent-offspring relationships in the captive breeding station. The correct parent pair was assigned for all 254 offspring where both parents had been genotyped. For 136 offspring where only one parent had been genotyped, the correct dummy parent was assigned in 127 cases (93.4 %) and the correct genotyped parent was assigned in 129 cases (94.9 %). Even in the few cases where none of the parents had been genotyped, the correct dummy parents were assigned in all seven cases (i.e. siblings cluster together with the same dummy parent(s). These figures show that the parentage analyses was of sufficient quality for reliable downstream analyses

Supplementary material 5 – Projection matrices

Table S5: Non-zero elements (fecundity and survival) of the projection matrices for females (*lf*) and males (*lm*).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Females (*lf*)** | | **Males (*lm*)** | |
| **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 6 – Zero-inflation in fecundity variable

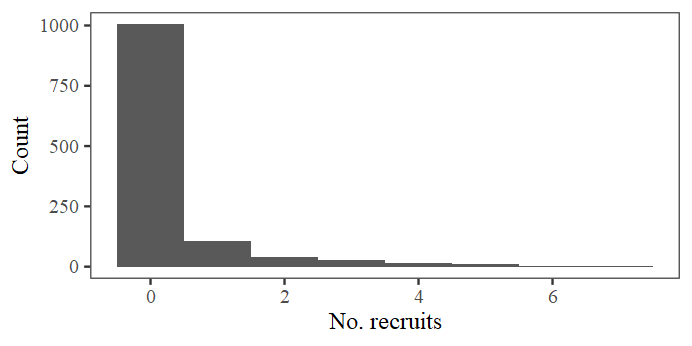


Figure S7: Histogram showing zero-inflation in the fecundity dataset where most individuals have zero fecundity (i.e. number of offspring that recruit into the adult population the next year).

Supplementary material 7 – Selection

*Analytical details*

To estimate age-specific reproductive values, the projection matrices for each sex (*lm* and *lf*, for males and females respectively,Table S5) were calculated using the R package *lmf* [5, 6]. The non-zero elements of *lm* and *lf* are fecundities in the first row and survival rates in the sub diagonal, estimated as the average for each age-class across years. The left and right eigenvectors of the projection matrices, scaled to *∑uv=1*, and *∑u=1,* give the reproductive values *v* and stable age distribution *u* at equilibrium. The reproductive value *vx is* interpreted as the expected contribution of an individual in age class *x* to the growth of the equilibrium population. The individual fitness for individual *i* independent of age can now be defined as [5, 7, 8]

, (Equation 1)

where *Wi* is the individual reproductive value, *Bi* is the number of recruits the individual has produced in year *t* and *Ji* is an indicator of survival from year *t* to year *t*+1, as described above.

In the GLM modelling used to estimate the relationships between fur colour genotype and individual fitness, the Poisson distribution requires the response variable to be integers. Hence, 2*Wi\** was used as response variable in place of *Λi*, where *Wi\** = (*Bi*/2 + *Ji*). To get correct parameter estimates and standard errors for *Λi* as the response variable, the model was fitted with an offset value log(*ci)* andweights *ωi = vx/ci,* where *vx* is theage‑specific reproductive values and *ci = 2Wi\*/Λi to* establish the relationship between 2*Wi*\*and *Λi*. When *Λi=0*, *ci* was set to 1*.*

Table S6: Overview over parameters used in modelling of arctic fox annual individual fitness and fitness components fecundity, adult survival, breeding probability and juvenile survival. Plus symbols (+) show variables included in the final models. Minus symbols (-) show variables that were tested but not included due to non-significance.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Fixed effects** | | | | | | **Random effects** | | | | **Model family** | **Results** |
| **Response** | **Genotype** | **Sex** | **Age** | **Age2** | **Genotype x sex** | **Genotype x age** | **Sub-population** | **Year** | **Birth year** | **Den** |  |  |
| **Annual individual fitness *lambda***  **(separate models for males and females)** | + |  |  |  |  |  | + | + |  |  | Poisson GLMM | Figure 1b  Table S8 |
| **Fecundity** | + | + | + | + | + | - | + | + |  |  | Zero-inflated Poisson GLMM | Figure 2a  Figure S8a  Table S9 |
| **Adult annual survival** | + | - | + | + | - | - | + | + |  |  | Binomial GLMM | Figure 2b  Figure S8b  Table S10 |
| **Adult breeding probability** | + | + | + | + | + | - | + | + |  |  | Binomial GLMM | Figure 2c  Figure S8c  Table S11 |
| **Recruitment probability (juv. survival)** | + | - |  |  | - |  | + |  | + | + | Binomial GLMM | Figure 2d |

*Environmental variables*

The abundance of small rodents usually varies in three to five year cycles. 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 [9]. Based on Angerbjörn *et al.* [10], 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 rodent phase was thus included as a fixed factor rather than a continuous covariate in the analyses.

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 [11] and was extracted for a 2.54 km buffer area around used arctic fox den sites in the study subpopulations (den site center, radius 2.54 km). This buffer area is an average of annual home ranges of resident arctic foxes presented in Landa *et al.* [12]. 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 any potential biases introduced by 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 across subpopulations and years before the analyses to create biologically meaningful intercepts (i.e. intercept represents the response variable at mean first or last day of snow fall).

Importantly, a measure of fitness in year *t* may be affected by the environment in the previous or subsequent year (e.g. fecundity in year *t* (the number of pups produced in year t that survived to year *t*+1)may be affected by rodent phase in year *t*, the year before (*t*-1) and/or the next year (*t*+1)).

To investigate if the effect of fur colour genotypes on individual fitness depended on the environmental variables or differed between origins, an interaction between *genotype* and the different variables were fitted. These interactions were added to models of individual fitness (estimated from Equation 1) and each of the fitness components, with one variable and interaction at a time (Table S7). LRTs between a model with the interaction and a model with only additive effects were performed to test whether there was support for the interaction between fur colour genotype and the variable of interest. In cases where *age* or *sex* was found to explain a significant proportion of the variance in the response variable (Table S6), these were included in both the additive and interaction model (Table S7).

Table S7: Overview over which predictor variables (and their temporal offsets; plus or minus one year) 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. Inclusion of variables were determined by prior knowledge of their ecological significance.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Response variable** | | | | |
| **Ecological predictor variable** | **Individual fitness (females and males separately)** | **Fecundity** | **Adult survival** | **Breeding probability** | **Recruitment 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** | + |  | + |  | + |

*Test for heterozygote advantage*

Differences in genome‑wide heterozygosity were modelled with a linear mixed effect model with a Gaussian error distribution. Fur colour *genotype* and *origin* (i.e. captive- or wild born) were included as fixed factor predictor variables. Random intercepts were fitted for *birth year* and *birth subpopulation* to account for interannual variation and variation between subpopulations. LRTs between models with and without the predictors were used to assess the effect of *genotype*. For 275 of these individuals, complete life histories were available. These individuals were used to investigate whether differences in genome‑wide heterozygosity affected the different fitness variables. For each fitness variable, an LRT was performed between the model described earlier and a model that in addition included *heterozygosity*, to assess the effect of heterozygosity.

*Further details on the results of the selection analyses*

Table S8: Parameter estimates and their 95% confidence intervals for GLMMs modelling individual fitness in female and male arctic foxes with *genotype* as predictor variable. Separate models were run for the two sexes, and parameter estimates and confidence intervals are given on log scale. Random intercepts were estimated for year and subpopulation. Estimates significant at the 0.05 significance level are given in bold. The parameter *TC* is given as the difference of TC from the intercept of CC individuals. Predictor variables are presented in normal font, random factors are written in *italic*.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **95% CI** | |
| **Variable** | **Estimate** | **Lower** | **Upper** |
| **Females** |  |  |  |
| CC (Intercept) | -0.131 | -0.270 | 0.006 |
| TC | 0.173 | -0.030 | 0.370 |
| *Year* | 0.0124 |  |  |
| *Subpopulation* | 0 |  |  |
| **Males** |  |  |  |
| CC (Intercept) | -0.133 | -0.266 | -0.004 |
| TC | 0.123 | -0.086 | 0.325 |
| *Year* | 0.0106 |  |  |
| *Subpoulation* | 0 |  |  |

Fecundity (i.e. number of recruiting offspring produced) was significantly affected by age and increased from age 1 to 4 before decreasing at age 5 (βage=1.478, βage^2=-0.203, LRT for age: χ2(2)=50.62, p<0.001; Figure S8a). The effect of genotype was independent of age (βTC x age=0.541±0.455, LRT for interaction: χ2(2)=1.69, p=0.429).

Survival probabilities of adult individuals increased from age 1 to age 2 and decreased for individuals older than 3 (βage=0.460±0.240, βage^2=‑0.093±0.041, LRT for age: χ2(2)=6.94, p=0.031; Figure S8b). The effect changed similarly with age for the two genotypes (βTC x age=0.606±0.559, LRT for interaction: χ2(2)=1.599, p=0.450).

Breeding probability increased with age until age 4, before levelling off at age 5 (βage=1.383±0.278, βage^2=‑0.150±0.046, LRT for age: χ2(2)=98.05, p<0.001; Figure S8c), an effect that was independent of genotype (βTC x age=-0.658±0.600, LRT for interaction: χ2(2)=4.78, p=0.091).

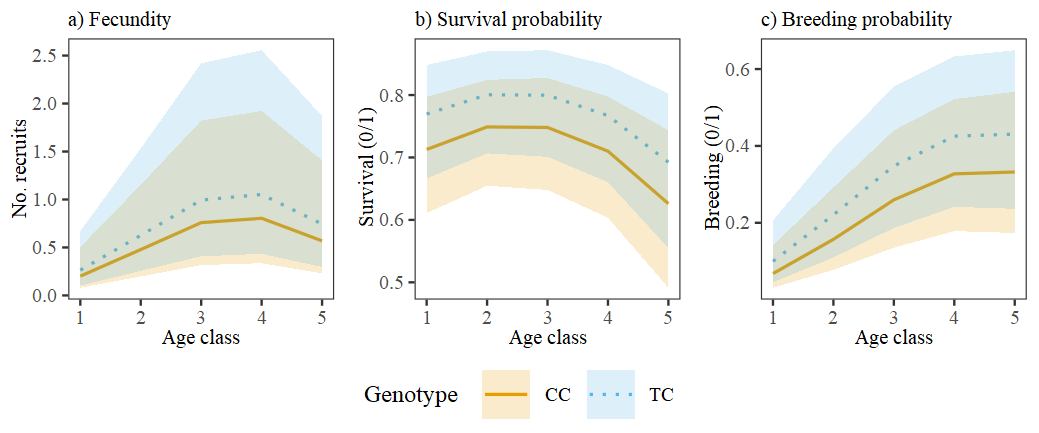


Figure S8: Predicted fecundity (a), adult survival probability (b) and breeding probability (c) for arctic fox fur colour genotypes CC and TC at different age classes. Shaded areas represent 95% confidence intervals. Predictions are based on additive GLMMs with *genotype* and *age* as predictor variables. Random intercepts were estimated for year and subpopulation. Age is included as a quadratic term to account for the non-linear relationship between age and the response variables.

Table S9: Parameter estimates and their 95% confidence intervals for GLMM modelling **fecundity** in arctic foxes. Random intercepts were estimated for year and subpopulation. Parameter estimates and confidence intervals are given on log scale. Estimates significant at the 0.05 significance level are given in bold. The parameter *TC* is given as the difference of TC from the intercept of CC individuals, and *Male* is the difference of males from the intercept of females. Random factors are given in *italic*.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **95% CI** | |
| **Parameter** | **Estimate** | **Lower** | **Upper** |
| Intercept (genotype CC; sex female) | -3.021 | -4.133 | -1.911 |
| **TC** | **0.497** | **0.180** | **0.814** |
| Male | 0.099 | -0.182 | 0.379 |
| **Age** | **1.526** | **1.072** | **1.980** |
| **Age2** | **-0.210** | **-0.283** | **-0.136** |
| **TC:Male** | **-0.528** | **-1.001** | **-0.055** |
| *Year* | 1.506 |  |  |
| *Subpopulation* | 0.097 |  |  |

Table S10: Parameter estimates and their 95% confidence intervals for GLMM modelling **adult survival** in arctic foxes. Random intercepts were estimated for year and subpopulation. Parameter estimates and confidence intervals are given on logit scale. Estimates significant at the 0.05 significance level are given in bold. The parameter TC is given as the difference of TC from the intercept of CC individuals. Random factors are given in *italic*.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **95% CI** | |
| **Parameter** | **Estimate** | **Lower** | **Upper** |
| Intercept (genotype CC) | 0.538 | -0.134 | 1.237 |
| TC | 0.296 | -0.008 | 0.608 |
| Age | 0.466 | -0.006 | 0.938 |
| **Age2** | **-0.094** | **-0.176** | **-0.013** |
| *Year* | 0.230 |  |  |
| *Subpopulation* | 0.111 |  |  |

Table S11: Parameter estimates and their 95% confidence intervals for GLMM modelling **breeding probability** in arctic foxes. Random intercepts were estimated for year and subpopulation. Parameter estimates and confidence intervals are given on logit scale. Estimates significant at the 0.05 significance level are given in bold. The parameter TC is given as the difference of TC from the intercept of CC individuals. Random factors are given in *italic*.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **95% CI** | |
| **Parameter** | **Estimate** | **Lower** | **Upper** |
| Intercept (genotype CC; sex female) | -3.889 | -5.046 | -2.861 |
| **TC** | **0.785** | **0.319** | **1.248** |
| Male | -0.041 | -0.386 | 0.304 |
| **Age** | **1.434** | **0.886** | **1.993** |
| **Age2** | **-0.158** | **-0.251** | **-0.067** |
| **TC:Male** | **-0.745** | **-1.417** | **-0.081** |
| *Year* | 0.939 |  |  |
| *Subpopulation* | 0.345 |  |  |

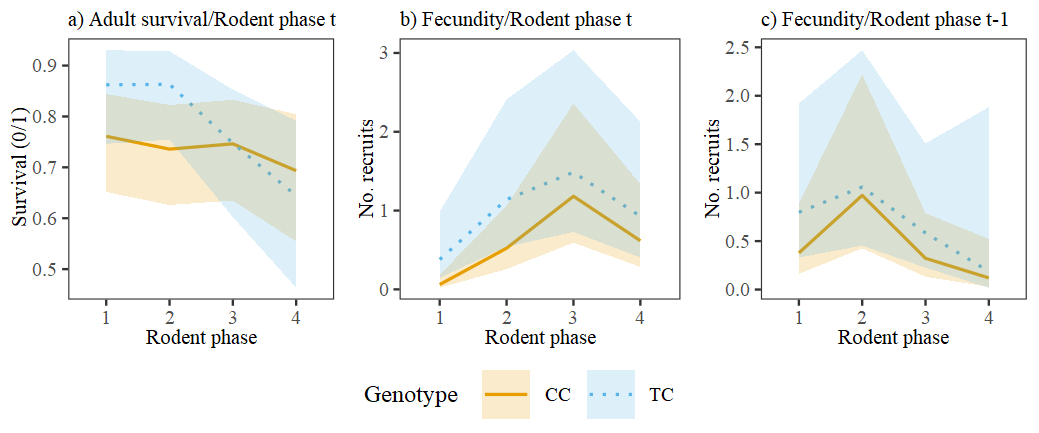


Figure S9: Predictions for adult survival (a), fecundity (b, c) of arctic fox fur colour genotypes CC and TC in different rodent phases. (c) shows rodent phase the year before breeding. Shaded areas represent 95% confidence intervals. Predictions are based on GLMMs with the following predictor variables: (a) *genotype*, *age*, *rodent phase* and *genotype*:*rodent phase* interaction; (b) *genotype*, *sex*, *age*, *rodent phase*, *sex*:*genotype* interaction and *genotype:rodent phase* interaction; (c) *genotype*, *sex*, *age*, *rodent phase t-1*, *sex*:*genotype* interaction and *genotype:rodent phase t-1* interaction. Age is included as a quadratic term to account for the non-linear relationship between age and the response variables. All models included *year* and *subpopulation* as random factors.

Table S12: Chi-square test statistics and p values for likelihood ratio tests conducted between additive (i.e. *variable* + *genotype*) and interaction models (i.e. *variable \* genotype*) 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** | **χ2 (df)** | **p value** | **Variable** | **χ2 (df)** | **p value** |
| *Fitness females* |  |  | *Recruitment probability* |  |  |
| Origin | 1.07 (1) | 0.302 | Rodent phase t | 0.76 (3) | 0.859 |
| Rodent phase t | 1.70 (3) | 0.637 | Rodent phase t+1 | 0.76 (3) | 0.859 |
| Rodent phase t+1 | 2.39 (3) | 0.495 | First snow fall t | 0.88 (1) | 0.350 |
| Rodent phase t-1 | 2.17 (3) | 0.538 | Last snow fall t+1 | 1.16 (1) | 0.282 |
| First snow fall t | 0.26 (1) | 0.612 | *Adult survival* |  |  |
| First snow fall t-1 | 0.01 (1) | 0.941 | Origin | 1.64 (1) | 0.201 |
| Last snow fall t | 0.13 (1) | 0.715 | **Rodent phase t** | **7.36 (3)** | **0.061** |
| Last snow fall t+1 | 0.21 (1) | 0.647 | Rodent phase t+1 | 5.38 (3) | 0.146 |
| *Fitness males* |  |  | First snow fall t | 0.48 (1) | 0.488 |
| Origin | 0.23 (1) | 0.634 | Last snow fall t+1 | 0.01 (1) | 0.917 |
| Rodent phase t | 1.06 (3) | 0.787 | *Fecundity* |  |  |
| Rodent phase t+1 | 0.22 (3) | 0.975 | Origin | 0.01 (1) | 0.926 |
| Rodent phase t-1 | 0.50 (3) | 0.919 | **Rodent phase t** | **9.32 (3)** | **0.025** |
| First snow fall t | 0.02 (1) | 0.899 | Rodent phase t-1 | 6.31 (3) | 0.097 |
| First snow fall t-1 | 0.01 (1) | 0.905 | First snow fall t-1 | 1.00 (1) | 0.318 |
| Last snow fall t | 0.19 (1) | 0.666 | Last snow fall t | 0.57 (1) | 0.450 |
| Last snow fall t+1 | 0.30 (1) | 0.582 | *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 |

Table S13: Chi-square test statistics and p values for likelihood ratio tests conducted between models containing *genotype* and, if significant, covariates *age* and *sex*, and models including *genome-wide heterozygosity* in addition to the other predictors. Parameter estimates and standard errors for genome-wide heterozygosity are given. One test is performed per response variable (i.e. fitness variable). Sample size n is given for observations (individual-years) and individuals.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fitness variable** | **p** | **Estimate±SE** | **n (individual-years)** | **n (individuals)** |
| Lambda females | 0.73 | 0.82±2.33 | 341 | 129 |
| Lambda males | 0.39 | 2.31±2.68 | 408 | 146 |
| Survival | 0.18 | 4.74±3.66 | 749 | 275 |
| Fecundity | 0.37 | 2.58±2.91 | 749 | 275 |
| Breeding probability | 0.71 | -1.63±4.37 | 749 | 275 |
| Recruitment probability | 0.76 | 3.28±10.34 | 130 | 130 |

Supplementary material 8 – Details on significant GWAS SNPs

Table S14: Distribution of SNPs significantly associated with arctic fox fur colour across scaffolds in the arctic fox genome. For each scaffold, the total number of SNPs included in the GWA-analysis, number of significant SNPs, to which dog chromosomes the scaffold blasts to, and number of intragenic SNPs are given.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Scaffold** | **No. SNPs** | **No. significant SNPs** | **Dog chromosomes (No. SNPs)** | **No. intragenic SNPs** |
| 11 | 6769 | 477 | Chr 5 (469)  Chr 17 (1)\*  Chr 27 (1)\* | 421 |
| 68 | 1636 | 13 | Chr 5 (13) | 13 |
| 1772 | 4 | 3 | Chr 5 (3) | 3 |
| 2224 | 2 | 2 | Chr 5 (1)  Chr 27 (1)\* | 1 |
| **Total** | **8411** | **495** | **489** | **438** |

*\* SNPs also had strong BLAST hits (e-value < 0.001 and query coverage > 70%) on dog chromosome 5, which suggests that these SNPs most likely also reside on dog chromosome 5. These SNPs were however conservatively discarded from further analysis.*

*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, χ2 test statistics and p‑values for all SNPs.

Supplementary material 9 – Fitness GWAS

We performed a candidate region GWAS for individual fitness in the area on arctic fox scaffold 11 where significant SNPs were found in the fur colour GWAS (9 872 872 – 32 864 328 bp, see results section). This region included 4 025 SNPs that passed genomic data quality control (see methods section), all these SNPs were included in this analysis. Arctic fox individuals included in this analysis needed to a) be genotyped on the Affymetrix SNP-array to assure sufficient genomic coverage, and b) included in the selection analysis (i.e. have complete known life history). 275 individuals matched both criteria. For these individuals, the same measures of annual individual fitness (lambda) as presented in the main text were used. The GWA analysis was performed using the R package *RepeatABEL* [13]which allows the use of repeated measures which was necessary due to the nature of lambda (annual measures). Individual fitness (lambda) was used as the response variable, sex, year, origin and subpopulation were included in the model as fixed factors. Lastly, individual ID was included as a random factor to correct for repeated measures. To account for relatedness between individuals, the genetic relatedness matrix (GRM) that we also used in the fur colour GWAS was included in the model (see methods section). To account for multiple testing, p-values were corrected using Bonferroni correction, where the significance level (α=0.05) was divided by the number of SNPs included in the analysis.

For the SNP with the lowest p-value, a BLAST [14] search was performed as presented for the fur colour SNPs in the methods section and Supplementary material 3. Gene functions were investigated using the UniProt knowledgebase [15].

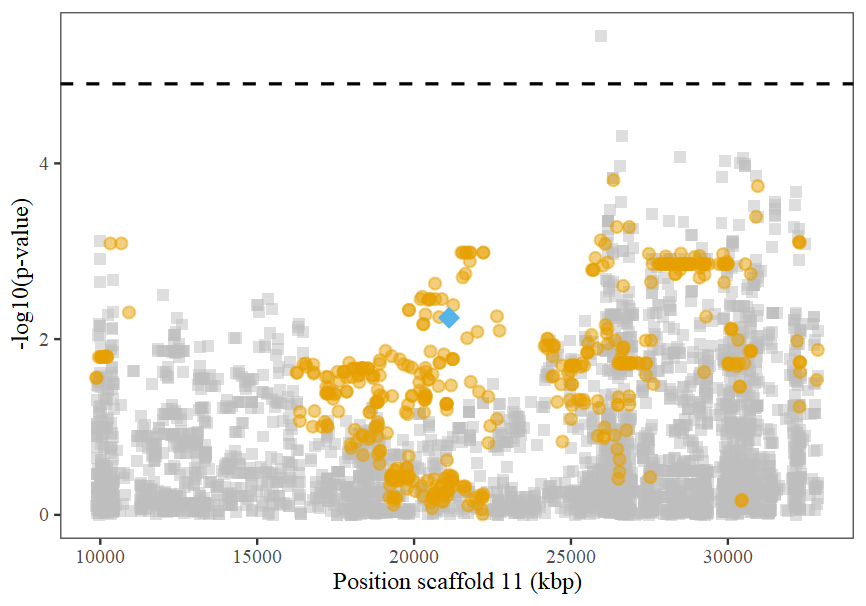


Figure S10: Manhattan plot showing results of candidate GWAS of arctic fox individual fitness (lambda) using information of 4 025 SNPs on arctic fox scaffold 11 typed in 275 individuals. P-values are given on a negative log scale. The significance threshold after Bonferroni correction is shown with a dashed horizontal line (the significance threshold is different than in the fur colour GWAS [Figure S4] due to the lower amount of SNPs included in the analysis). SNPs significantly associated with arctic fox fur colour are shown with orange circles. The blue diamond shows the position of the diagnostic SNP for arctic fox fur colour (AX-176934441, p=0.0021).

Supplementary material 10 – Gene analyses

Table S15: Names of 154 genes that are located close (<20 kbp) to a SNP significantly associated with arctic fox fur colour. The assumed causal gene MC1R is given in bold.

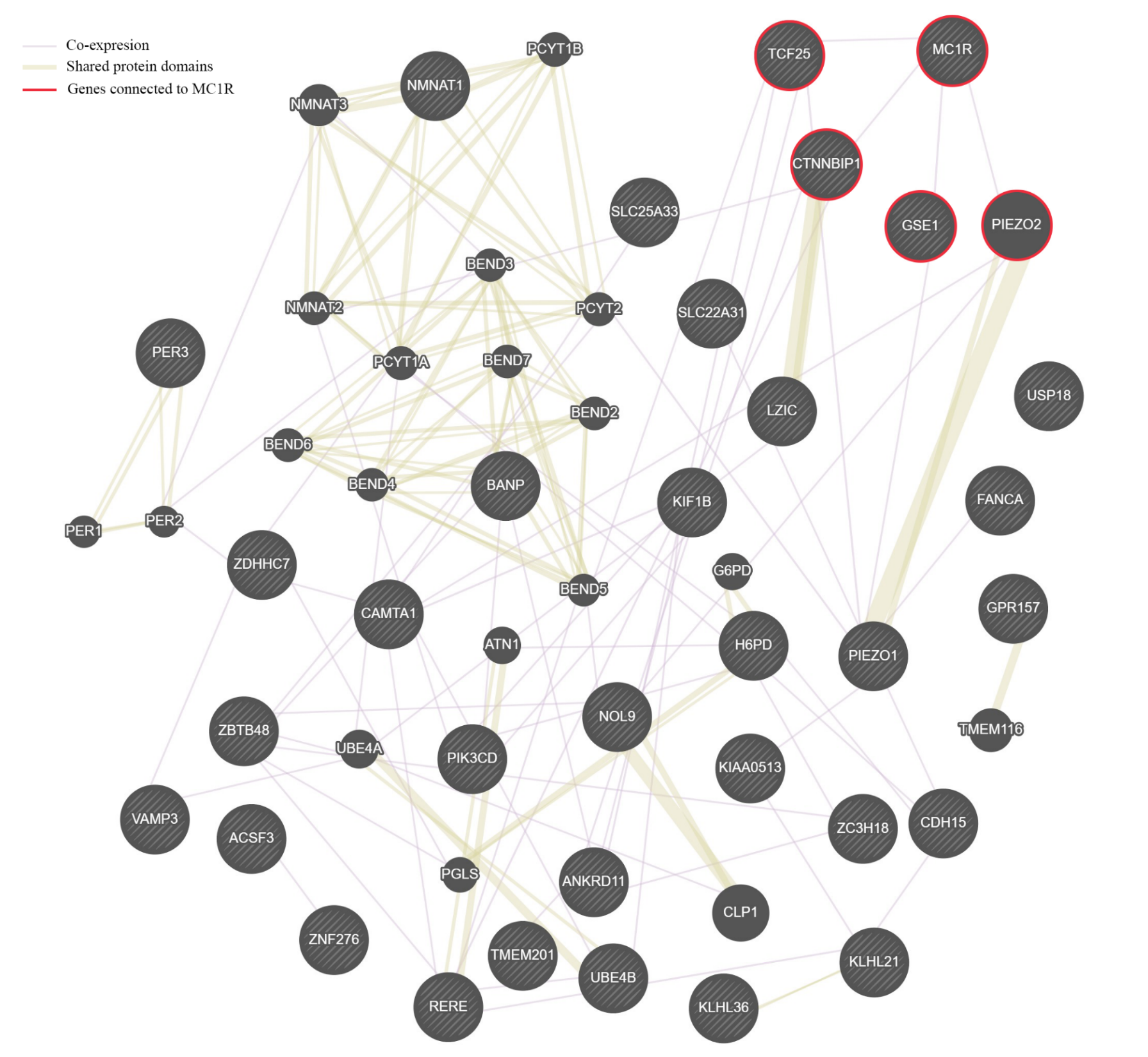
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Distance to closest significant SNP*** | | | | |
| **Intragenic** | **Intragenic** | **Intragenic** | **<10 kbp** | **<20 kbp** |
| ACOT7 | HSD17B2 | PER3 | CBFA2T3 | APRT |
| ACSF3 | HSDL1 | PIEZO1 | CDH15 | ATP2C2 |
| AJAP1 | ICMT | PIK3CD | CENPBD1 | 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 | VAMP3 | **MC1R** |  |
| ERRFI1 | LOC489707 | VAT1L | NECAB2 |  |
| FANCA | LZIC | WFDC1 | PABPN1L |  |
| FUK | MBTPS1 | WWOX | SCAMPER |  |
| GAN | MPHOSPH6 | ZBTB48 | SDR42E1 |  |
| GINS2 | NMNAT1 | ZC3H18 | SLC22A31 |  |
| GPR157 | NOL9 | ZDHHC7 | TMEM231 |  |
| GSE1 | NPHP4 |  | TRAPPC2L |  |
| H6PD | OSGIN1 |  | UTS2 |  |
|  |  |  | ZNF276 |  |

Table S16: 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 | PFDR 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] |

Table S17: Summary of gene functions of 41 genes located closer than 10 kbp to SNP that is significantly associated with arctic fox fur colour and in high LD (r2 >= 0.5) with the SNP most associated with arctic fox fur colour. The assumed causal gene MC1R is given in bold. Gene functions were retrieved from UniProt Knowledgebase (UniProtKB) [15]

|  |  |  |
| --- | --- | --- |
| Gene | Function | Review status |
| 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 |
| CTNNBIP1 | Negative regulator of Wnt signalling pathway | Dog |
| 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 |
| MC1R | **Receptor for MSH (alpha, beta) and ACTH. Involved in the melanocortin system that regulates melanin-based colouration** | **Dog** |
| 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 |
| SCAMPER | Calcium regulation | Dog |
| 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 |
| 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 |

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 (r2 >= 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 [16].

Supplementary material 11 – LD decay

Linkage disequilibrium (LD) decay was investigated in arctic fox scaffold 11 since this was the scaffold of largest interest in this study. Pairwise LD (r2) was calculated for all SNPs on scaffold 11 using PLINK 1.90 [1, 2]. Mean LD was then calculated for 5 kbp bins for SNPs closer than 100 kbp (Fig. S12a) and for 100 kbp bins spanning the complete scaffold (Fig. S12b). Data from all fox individuals included in the fur colour GWAS were included in this LD decay analysis, including individuals from different subpopulations and close relatives.

Mean LD (r2) decreased quickly from 0.33 (SNP distance up to 5 kbp) to 0.26 (SNP distance 5-10 kbp) (Figure S12a). At a SNP distance of roughly 10 000 kbp (i.e. 10 Mbp), r2 values below 0.05 are reached (Figure S12b).

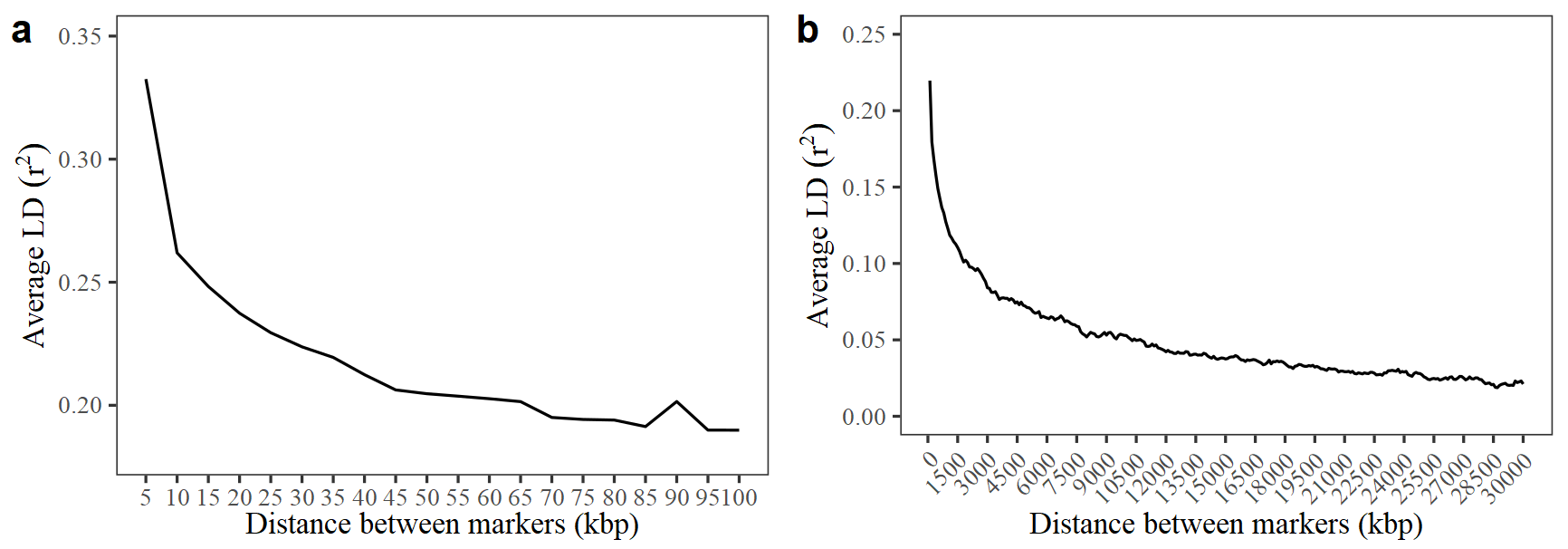


Figure S12: Linkage disequilibrium (r2) decay in arctic fox scaffold 11. (a) shows mean LD between SNPs closer than 100 kbp based on 5 kbp bins. (b) shows mean LD for SNPs spanning the entire scaffold based on 100 kbp bins.

Supplementary material 12 – 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 in main text; Figure 1) 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.* [17] 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.* [17]. 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.* [17] 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 – References

[1] Purcell, S. & Chang, C.C. 2019 PLINK (v.1.90) [www.cog-genomics.org/plink/1.9/](file:///C:\D\Forskning\Fjellrev\Manus\Colour%20morph%20genetics\MS-20210807\www.cog-genomics.org\plink\1.9\).

[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).

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

[4] Huisman, J. 2017 Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol Ecol Resour* **17**, 1009-1024. (doi:10.1111/1755-0998.12665).

[5] Engen, S., Sæther, B.-E., Kvalnes, T. & Jensen, H. 2012 Estimating fluctuating selection in age-structured populations. *J Evol Biol* **25**, 1487-1499. (doi:10.1111/j.1420-9101.2012.02530.x).

[6] Kvalnes, T. 2013 lmf: Functions for estimation and inference of selection in age-structured populations. R package version 1.2. (

[7] Engen, S., Kvalnes, T. & Sæther, B.E. 2014 Estimating phenotypic selection in age‐structured populations by removing transient fluctuations. *Evolution* **68**, 2509-2523. (doi:10.1111/evo.12456).

[8] Engen, S., Lande, R., Sæther, B.-E. & Dobson, F.S. 2009 Reproductive value and the stochastic demography of age-structured populations. *The American Naturalist* **174**, 795-804. (doi:10.1086/647930).

[9] Framstad, E. 2017 Terrestrisk naturovervåking i 2016: Markvegetasjon, epifytter, smågnagere og fugl. Sammenfatning av resultater. *NINA Rapport* **1376**. (doi:<http://hdl.handle.net/11250/2452619>).

[10] Angerbjörn, A., Eide, N.E., Dalén, L., Elmhagen, B., Hellström, P., Ims, R.A., Killengreen, S., Landa, A., Meijer, T., Mela, M., et al. 2013 Carnivore conservation in practice: replicated management actions on a large spatial scale. *J Appl Ecol* **50**, 59-67. (doi:10.1111/1365-2664.12033).

[11] Saloranta, T. 2014 New version (v.1.1.1) of the seNorge snow model and snow maps for Norway. In *Rapport* (Oslo, Norwegian Water Resources and Energy Directorate.

[12] 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. *Can J Zool* **76**, 448-457. (doi:10.1139/z97-209).

[13] Rönnegård, L., McFarlane, S.E., Husby, A., Kawakami, T., Ellegren, H. & Qvarnström, A. 2016 Increasing the power of genome wide association studies in natural populations using repeated measures: evaluation and implementation. *Methods Ecol Evol* **7**, 792-799. (doi:10.1111/2041-210X.12535).

[14] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990 Basic local alignment search tool. *J Mol Biol* **215**, 403-410. (doi:10.1016/S0022-2836(05)80360-2).

[15] The UniProt Consortium. 2018 UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **46**, 2699-2699. (doi:10.1093/nar/gky092).

[16] Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. 2010 The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* **38**, W214-W220. (doi:10.1093/nar/gkq537).

[17] 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**, 1814-1817. (doi:10.1016/j.peptides.2004.11.040).