ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12476

Genetic diversity analyses reveal first insights into breed-specific selection signatures within Swiss goat breeds

A. Burren*[1], M. Neuditschko[†1], H. Signer-Hasler*, M. Frischknecht*, I. Reber[‡], F. Menzi[‡], C. Drögemüller[‡] and C. Flury*

*School of Agricultural, Forest and Food Sciences HAFL, Bern University of Applied Sciences, Länggasse 85, 3052 Zollikofen, Switzerland. [†]Swiss National Stud Farm, Agroscope Research Station, Les Longs-Prés, 1580 Avenches, Switzerland. [‡]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109, 3001 Bern, Switzerland.

Summary

We used genotype data from the caprine 50k Illumina BeadChip for the assessment of genetic diversity within and between 10 local Swiss goat breeds. Three different cluster methods allowed the goat samples to be assigned to the respective breed groups, whilst the samples of Nera Verzasca and Tessin Grey goats could not be differentiated from each other. The results of the different genetic diversity measures show that Appenzell, Toggenburg, Valais and Booted goats should be prioritized in future conservation activities. Furthermore, we examined runs of homozygosity (ROH) and compared genomic inbreeding coefficients based on ROH (F_{ROH}) with pedigree-based inbreeding coefficients (F_{PED}). The linear relationship between F_{ROH} and F_{PED} was confirmed for goats by including samples from the three main breeds (Saanen, Chamois and Toggenburg goats). F_{ROH} appears to be a suitable measure for describing levels of inbreeding in goat breeds with missing pedigree information. Finally, we derived selection signatures between the breeds. We report a total of 384 putative selection signals. The 25 most significant windows contained genes known for traits such as: coat color variation (*MITF*, *KIT*, *ASIP*), growth (*IGF2*, *IGF2R*, *HRAS*, *FGFR3*) and milk composition (*PITX2*). Several other putative genes involved in the formation of populations, which might have been selected for adaptation to the alpine environment, are highlighted. The results provide a contemporary background for the management of genetic diversity in local Swiss goat breeds.

Keywords inbreeding coefficient, pedigree, runs of homozygosity, selection signals, SNP

Introduction

In 2011, the International Goat Genome Consortium reported genome-wide SNP data at ~50 000 markers in 281 goats from 10 geographically and biologically diverse breeds (Tosser-Klopp *et al.* 2014). Through these activities, the caprine 50k Illumina BeadChip became available for commercial use. This BeadChip was used in different studies after its release, including in the investigations of genes for coat color variation (Becker *et al.* 2014), wattles (Reber *et al.* 2015) and polledness (Kijas *et al.* 2013). In addition, this caprine 50k Illumina BeadChip was used to evaluate genomic selection programs in dairy goats. However, so far, only attempts in France (Carillier *et al.* 2013, 2014) and the UK (Mucha *et al.* 2015) are being undertaken to implement genomic selection in dairy goat breeding programs, as the global breeding industry is not as well established when compared with that for dairy cattle. Furthermore, Kijas *et al.* (2013) have demonstrated that SNP array technology offers great potential to enhance the understanding of domestication and genetic diversity in goats. Based on genome-wide SNP information, three Australian goat breeds clustered separately, although model-based clustering suggested that admixture has occurred and that (historical) genetic links exist between the breeds (Kijas *et al.* 2013). Nicoloso *et al.* (2015) investigated the genetic diversity of 14 Italian goat breeds based on 50k genotypes and reported a north–south geographic pattern in present-day genetic diversity. Both SNP-based goat diversity studies highlighted the high levels of polymorphisms and concluded that goats contain more polymorphic sites than do other livestock species (Kijas *et al.* 2013; Nicoloso *et al.* 2015).

Address for correspondence

A. Burren, School of Agricultural, Forest and Food Sciences HAFL, Bern University of Applied Sciences, Länggasse 85, 3052 Zollikofen.
E-mail: alexander.burren@bfh.ch

[1]These authors contributed equally to this work.

Accepted for publication 16 June 2016

Switzerland has 10 local goat breeds: Appenzell (APP), Booted (SGB), Chamois colored (CHA), Grisons striped (GST), Nera Verzasca goat (NVE), Peacock (PEA), Saanen (SAA), Tessin grey (TGR), Toggenburg (TOG) and Valais (VAG). There are two different coat color strains in the Valais goat breed: one with a black neck ('black-neck'; Glowatzki-Mullis *et al.* 2008) and a rare one with a *TYRP1*-associated brown neck ('copper-neck'; Becker *et al.* 2014). To date, 43 microsatellite markers have been used to determine genetic diversity measures for these breeds (Glowatzki-Mullis *et al.* 2008). At this stage, it has not been possible to separate three breeds from each other, namely, PEA, NVE and TGR (Glowatzki-Mullis *et al.* 2008), and samples of the copper-neck strain of VAG were not available.

With the availability of genome-wide SNP data, there is an increased interest in estimating genomic inbreeding through genome-wide runs of homozygosity (ROH) (Curik *et al.* 2014). So far, the suitability of ROH for the estimation of inbreeding coefficients has been demonstrated for cattle (Ferenčaković *et al.* 2011, 2013a,b; Purfield *et al.* 2012) and pigs (Herrero-Medrano *et al.* 2013; Silió *et al.* 2013). Kim *et al.* (2015) found a higher proportion of individuals lacking long stretches in the Barki goat breed – an indigenous goat breed from a hot arid environment – when compared with the Boer goat breed and other exotic breeds. Furthermore, the same authors derived selection signatures for the Barki goat and the Barki sheep and concluded that these signatures identified genes underlying local adaptation. Benjelloun *et al.* (2015) used whole genome sequencing data from 44 individuals from three phenotypically distinct goat populations in Morocco and presented several genomic regions that are influenced by positive selection.

In the current study, we used 50k Illumina BeadChip genotypes from the aforementioned Swiss goat breeds for the re-assessment of population structure. Furthermore, we derived ROH and compared marker-based measures of inbreeding with pedigree-based inbreeding coefficients. Finally, selection signatures were calculated and potential loci leading to population differentiation are presented. The results are put forward as a contemporary background for the management of genetic diversity in local Swiss goat breeds.

Materials and methods

Sample collection and genotypes

In the context of this and other projects (Becker *et al.* 2014; Reber *et al.* 2015), 473 animals from the 10 Swiss goat breeds were genotyped for the caprine 50k Illumina BeadChip. PLINK v1.07 (Purcell *et al.* 2007) was used for the derivation of genomic relationships by using the --genome function, as described in Burren *et al.* (2014). After exclusion of closely related individuals (genomic relationship >0.30), the final data set (Data1) included 284 individuals from 10 breeds. The number of genotypes per breed is given in Table 1. Pedigree information was obtained for 249 animals from the corresponding herd books (Table 1). We incorporated closely related individuals into the analysis for the derivation of ROH. This involved the additional inclusion of 189 genotypes from highly related animals (Data2), all with known pedigree information (Table 1). For each breed, phenotypic characteristics related to coat color, hair length, polledness, body size, performance and its classification as a main or rare breed are summarized in Table S1.

Data editing

PLINK ped and map files (v1.7/v1.9; Purcell *et al.* 2007; Chang *et al.* 2015) were prepared based on the TOP format from the available Illumina raw data files. The map file was created by downloading SNP positions on the Chinese goat assembly CHI_1.0 (Dong *et al.* 2013) from the SNPchiMp database (Nicolazzi *et al.* 2014). A total of 3406 sex-chromosomal and non-annotated SNPs were removed. Data1 contained 48 019 SNPs after filtering for minor allele frequency (>0.01), for missing genotypes at individual and marker levels (<0.1) as well as for SNPs deviating from Hardy-Weinberg equilibrium ($P > 0.0001$). Across all autosomal caprine chromosomes (CHI) and informative markers, a mean gap of 49.996 kb (minimum, 0.029 kb; maximum, 475.015 kb) was observed between adjacent SNPs, with a mean r^2 value of 0.089 (Table S2). The 48 019 SNPs in Data1 were also considered to be informative for Data2. Data2 is available as part of a Dryad data package (doi: 10.5061/dryad.q1cv6).

Population structure

Various parameters related to genetic diversity and population structure were assessed based on Data1. Expected

Table 1 Overview of sampled individuals per breed in Data1 (genetic diversity and selection signature analyses) and Data2 (ROH analysis). The number of animals with pedigree information are indicated in the second column for Data1 and in the fourth column for Data2.

	Data1		Data2	
Breed	Total animals	With pedigree	Total animals	With pedigree
Appenzell goat, APP	21	18	29	26
Grisons striped goat, GST	26	24	49	47
Tessin grey goat, TGR	27	19	37	29
Chamois colored, CHA	61	61	124	124
Valais goat, VAG	24	15	43	34
Nera Verzasca goat, NVE	29	19	42	32
Peacock goat, PEA	22	22	31	31
Saanen goat, SAA	34	34	64	64
Booted goat, SGB	16	13	23	20
Toggenburg goat, TOG	24	24	31	31
Total	284	249	473	438

(H_E) and observed (H_O) heterozygosity were calculated with the command --het, as implemented in PLINK v1.7 (Purcell *et al.* 2007). ADZE software (Szpiech *et al.* 2008) was utilized for the derivation of allelic richness and private allelic richness, using a standardized sample size for each breed. The SNEP package was used to calculate estimates of linkage disequilibrium (LD) and effective population sizes (N_e) for each breed (Barbato *et al.* 2015). The program ADMIXTURE (Alexander *et al.* 2009) was employed to determine the optimal number of k clusters and to assign individuals to their true clusters. We determined the optimal number of clusters by adding the cv-flag (Alexander *et al.* 2009). Within this analysis, the number of clusters was increased from 1 to 11, and the k with the lowest cross-validation error was used for the selection of the optimal number of clusters for the set of genotypes under investigation. The software DISTRUCT (Rosenberg 2004) was used for the graphical presentation of each cluster assignment, increasing k from 2 to 11. Apart from the model-based cluster analysis, we further investigated population structure using principal components analysis (PCA), a non-parametric approach that utilizes pairwise relationships between individuals for the final visualization of genome-wide population structures.

High definition network visualization for the available SNP genotypes was used for the detection of fine-scale population structures within and between breeds (Neuditschko *et al.* 2012). The five distinct components of the so-called NETVIEW approach (Neuditschko *et al.* 2012) are described in Burren *et al.* (2014). In the current study, we set the number of nearest neighbors equal to 10.

Pairwise F_{ST} values were calculated among the 10 goat breeds using the R package GENELAND (Guillot *et al.* 2005) to evaluate the general hierarchical population structure. The phylogenetic relationships between breeds were visualized using the commonly applied neighbor-joining method, as implemented in the program SPLITSTREE 4 (Huson & Bryant 2006).

Runs of homozygosity

For the 473 individual genotypes in Data2, ROH were derived by using PLINK v1.7 (Purcell *et al.* 2007) and the following settings: minimum SNP density was set to one SNP every 70 kb, with a maximum gap length of 1 Mb. For each ROH, one heterozygote and two missing genotypes were permitted.

The total number of ROH, length of ROH (in Mb) and the sum of all ROH segments (in Mb) were calculated for all animals, separated by breed and ROH length category. The relationship between the number of ROH and the sum of ROH was investigated using Spearman's rank correlation coefficient (r_S) (McQuillan *et al.* 2008; Ferenčaković *et al.* 2011). The ROH segments were allocated to the following seven classes to compare the ROH segment lengths between breeds: 1–5, >5–10, >10–15, >15–20, >20–25, >25–30 and >30 Mb, identified as $ROH_{1–5Mb}$, $ROH_{5–10Mb}$, $ROH_{10–15Mb}$, $ROH_{15–20Mb}$, $ROH_{20–25Mb}$, $ROH_{25–30Mb}$ and $ROH_{>30Mb}$ respectively. The genomic inbreeding coefficients (F_{ROH}) were calculated using the method published in McQuillan *et al.* (2008):

$$F_{ROH} = \sum \frac{L_{ROH}}{L_{AUTO}}$$

where L_{AUTO} is the length of the autosomal genome for the SNPs. In the current study, L_{AUTO} was set to 2399.4 Mb (Table S2). All available ancestors of the 438 animals with known pedigree (Data2) were considered for the derivation of pedigree-based inbreeding coefficients. The pedigrees included totals of 725 (APP), 546 (BST), 140 (TGR), 3064 (CHA), 112 (VAG), 499 (NVE), 432 (PEA), 2795 (SAA), 377 (SGB) and 1737 (TOG) animals. The inbreeding coefficient (F_{PED}) was calculated with the software CFC v1.0 (Sargolzaei *et al.* 2005, 2006), using Wright's (1922) method of path coefficients and a mean pedigree completeness index for five generations (MacCluer *et al.* 1983). F_{PED} and F_{ROH} were compared using linear regression and Pearson's correlation coefficients (r_p). In a first step, all animals with pedigree information were considered, irrespective of the mean pedigree completeness index for the first five generations ($n = 438$). In a second step, only animals with a pedigree completeness index >90% ($n = 330$) were included in the linear regression of F_{PED} on F_{ROH}.

Selection signatures

Data1 was used for the identification of selection signatures. Due to recent results indicating a lack of distinction between TGR, NVE and PEA (Glowatzki-Mullis *et al.* 2008), these breeds were considered as one population (abbreviated as TGR/NVE/PEA), resulting in eight breed groups for the investigation of selection signatures. Wright's (1943) F_{ST} values were calculated for all of the 28 breed pairs by using PLINK v1.9 (--fst). Following this, d_i values were calculated for each breed and for all 48 019 SNPs according to the formula published by Akey *et al.* (2010):

$$d_i = \sum_{j \neq i} \frac{F_{ST}^{ij} - E[F_{ST}^{ij}]}{\mathrm{sd}[F_{ST}^{ij}]}$$

where $E[F_{ST}^{ij}]$ and $\mathrm{sd}[F_{ST}^{ij}]$ denote the expected value and standard deviation of F_{ST} between breeds i and j, calculated from all 48 019 SNPs. The d_i values were averaged for SNPs in non-overlapping windows of 500 kb. Windows with fewer than four SNPs were discarded, resulting in 4792 informative windows covering a mean number of 10.01 SNPs (maximum = 33 SNPs). The 48 windows exceeding the 99th percentile of the breed-specific empirical distribution of d_i were considered as putative selection signals for

each breed. Genes in the region of the 25 windows with the highest d_i values were identified breed-wise with the NCBI MAPVIEWER (http://www.ncbi.nlm.nih.gov/projects/mapview/). All genes found 1 Mb up- and downstream of the middle position of the 25 top significant windows were listed (Table S4). Knowledge about breed-specific characteristics (Table S1) and insights from the literature were combined to select the candidate genes given in Table 4.

Results

Population structure

The results of the different indices of genetic diversity are given in Table 2. The proportion of polymorphic SNPs (P_N) was comparable among the breeds and ranged from 0.958 (TOG) to 0.962 (TGR, NVE and SGB). The observed heterozygosity varied between 0.369 (APP and TOG) and 0.401 (GST and PEA). Allelic richness was highest in the CHA, at 1.914. The private allelic richness was similar and at a low level among breeds (range: 0.002–0.003).

Estimates of recent effective population sizes (13 generations ago) below 90 were found for APP, GST, VAG, SGB and TOG. For TGR, CHA, NVE and SAA, recent effective population sizes (13 generations ago) were >110 (Table 2). Breed-specific decay in LD and estimated effective population sizes for all generations are given in Fig. S1 and Table S3. Mean genomic relationships ranged from 0.041 (TGR) to 0.160 (APP). A boxplot showing the distribution of genomic relationships for all breeds can be found in Fig. S2.

The first principal component (PC1) explained 13% (Fig. 1), the second (PC2) 9% and the third (PC3) 8% of the observed variation. Contrasting PC1 vs. PC2 and PC1 vs. PC3 resulted in clear separation of APP, GST, CHA, VAG, SAA, SGB and TOG. However, a clear distinction between TGR, NVE and PEA was not possible based on the first three components. In addition, a degree of overlap was observed between GST and TGR (Fig. 1).

The graphical visualization of the results of the cluster analysis of the 284 animals for k ranging from two to 11 clusters is shown in Fig. 2. Based on the cross-validation error, we identified an optimal value of $k = 9$ clusters (Fig. S3). The results from the cluster analysis at the optimal number of $k = 9$ were further included in the high-resolution population network illustration given in Fig. S4. In addition, F_{ST} distances between breeds are graphically displayed as a neighbor-joining tree (Fig. S5).

Runs of homozygosity

In total, 5803 ROHs of varying lengths ranging between 3.663 and 103.745 Mb were identified based on Data2. Out of the 473 individuals, 461 individuals had at least one ROH. On average, this resulted in 13 ROHs per individual, ranging from seven (TGR and PEA) to 23 (APP). The proportion of individuals lacking ROH was zero for APP, GST, VAG, NVE, PEA and TOG. Non-zero values were found for TGR (21.6%), SGB (4.3%), SAA (3.1%) and CHA (0.8%). There was a correlation between the number of ROHs (nROH) and the sum of all ROH segments ($r_s = 0.95$; Fig. S6). The distribution of F_{ROH} and the corresponding breed is given as the box plot in Fig. 3, where the lower end of the box stands for the first quartile (Q1) followed by the median and the third quartile (Q3). The whiskers are defined as $Q1 - 1.5\ (Q3 - Q1)$ and $Q3 + 1.5\ (Q3 - Q1)$ respectively. The highest median value was reported in APP and the lowest in TGR. For all 10 breeds investigated here at least one up to three outliers (i.e. values greater than the upper whisker) with unexpected high F_{ROH} were identified (dots in Fig. 3).

The distribution of the relative numbers of ROH in different length classes and the 10 breeds is shown in Fig. 4. For all breeds, the majority of ROHs were in the class >5–10 Mb. However, the proportions in this major class varied between breeds: at least 60% of the ROHs were in this length class for APP, CHA, VAG, PEA, SAA and TOG,

Table 2 Indices of genetic diversity (Data1; 48 019 SNPs and 284 individuals).

Breed	n	P_N	H_O	H_E	A_R	pA_R	Ne_{13Gen}	Ø gRel
APP	21	0.960	0.369	0.360	1.803	0.002	65	0.160
GST	26	0.961	0.401	0.389	1.897	0.003	86	0.063
TGR	27	0.962	0.396	0.393	1.921	0.002	116	0.041
CHA	61	0.959	0.394	0.389	1.941	0.003	157	0.063
VAG	24	0.960	0.371	0.366	1.900	0.003	83	0.123
NVE	29	0.962	0.386	0.387	1.839	0.002	119	0.046
PEA	22	0.959	0.401	0.385	1.901	0.002	80	0.082
SAA	34	0.960	0.386	0.379	1.866	0.002	120	0.092
SGB	16	0.962	0.391	0.374	1.531	0.002	52	0.115
TOG	24	0.958	0.369	0.362	1.812	0.002	78	0.154

N, number of individuals tested per breed; P_N, the proportion of SNPs that displayed polymorphism; H_O, observed heterozygosity; H_E, expected heterozygosity or gene diversity; A_R, allelic richness; pA_R, private allelic richness; Ne_{13Ge}, recent effective population size; Ø gRel, average genomic relationships within breed; APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois Colored; VAG, Valais Goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

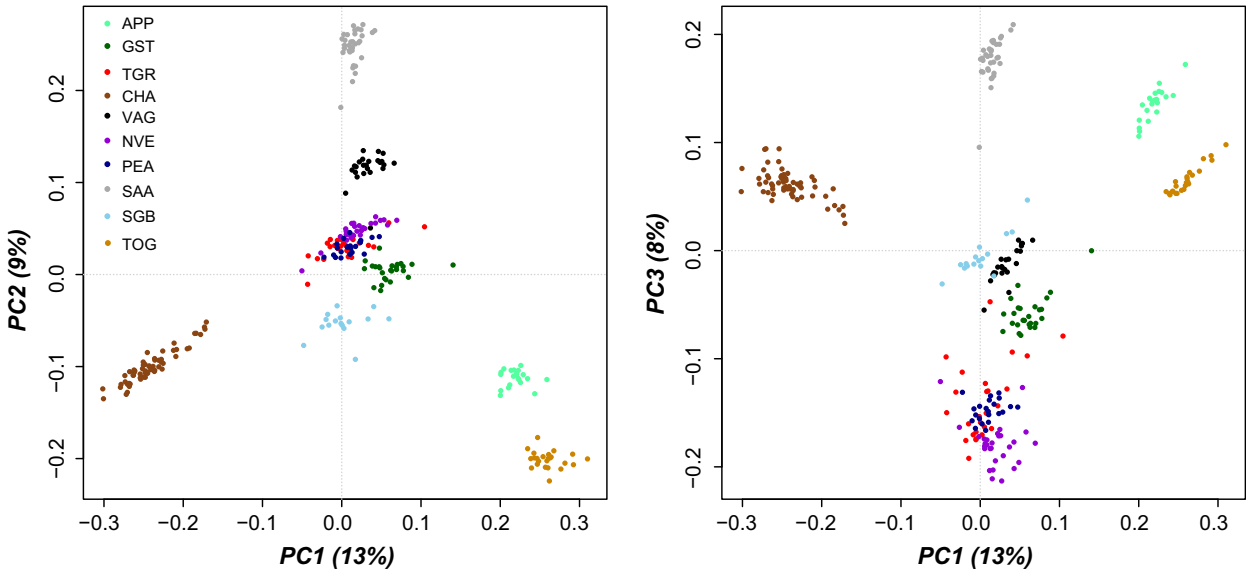

Figure 1 PCA plot for the first three components PC1, PC2 and PC3 (and the respective variation explained in brackets) and for the 10 Swiss goat breeds. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

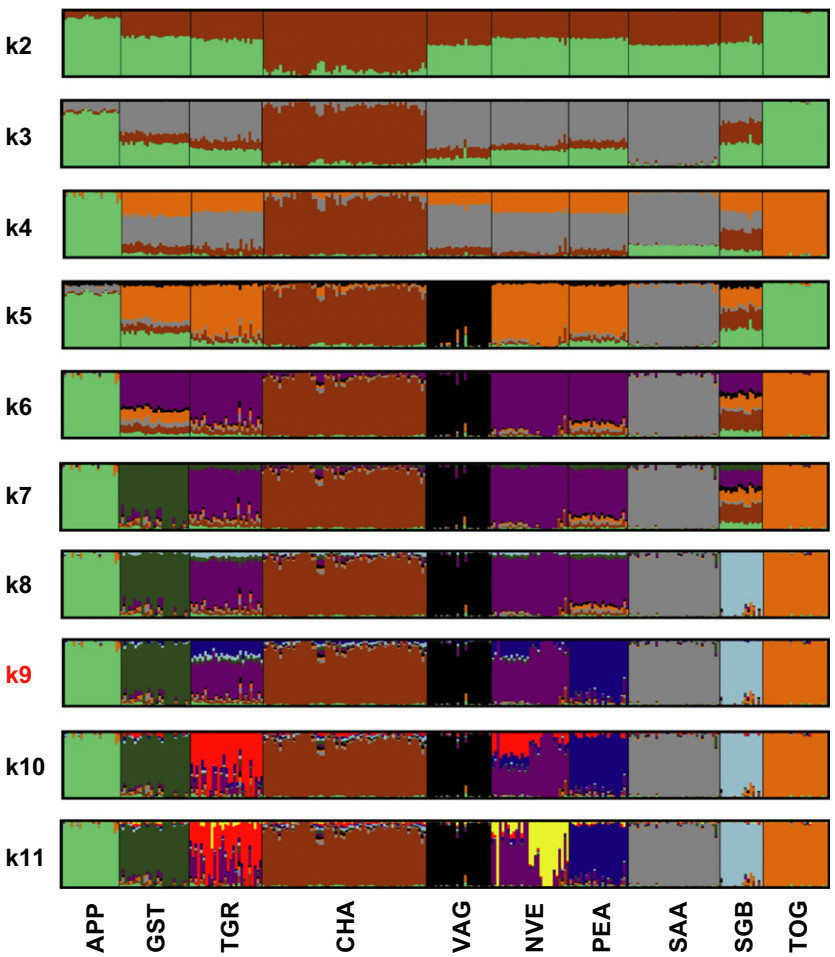

Figure 2 Admixture results for $k = 2$–11. The optimal number of clusters ($k = 9$), according to the cross-validation analysis, is indicated in red. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

whereas the proportion was 54% for GST and SGB and less than 50% for TGR and NVE. A comparison among the breeds revealed that TGR and NVE had the highest frequencies in the length classes greater than 20 Mb, which is more indicative for recent inbreeding (Purfield *et al.* 2012).

F_{ROH} values were compared for the 438 pedigreed individuals with their pedigree-based coefficients of inbreeding to assess the utility of F_{ROH} as an indicator of individual levels of inbreeding. Mean pedigree completeness, mean F_{ROH} (minimum and maximum), mean F_{PED} (minimum and maximum) and R^2 from the linear regression of F_{ROH} on F_{PED} for each breed are summarized in Table 3. The linear regression plot of F_{ROH} on F_{PED} for all breeds is shown in Fig. 5a. The correlation coefficient, r_p, was 0.519 ($R^2 = 0.269$). This relationship is highly dependent on pedigree completeness, which was generally lower in rare breeds when compared with the three main breeds (CHA, SAA, TOG) (Table 3). By excluding animals with a pedigree completeness <90%, r_p increased to 0.599 ($R^2 = 0.359$ (Fig. S7a). In Fig. 5b, the correlation between F_{ROH} and F_{PED} was restricted to the 219 animals in the three main breeds, CHA, SAA and TOG, resulting in a correlation of 0.810 ($R^2 = 0.652$).

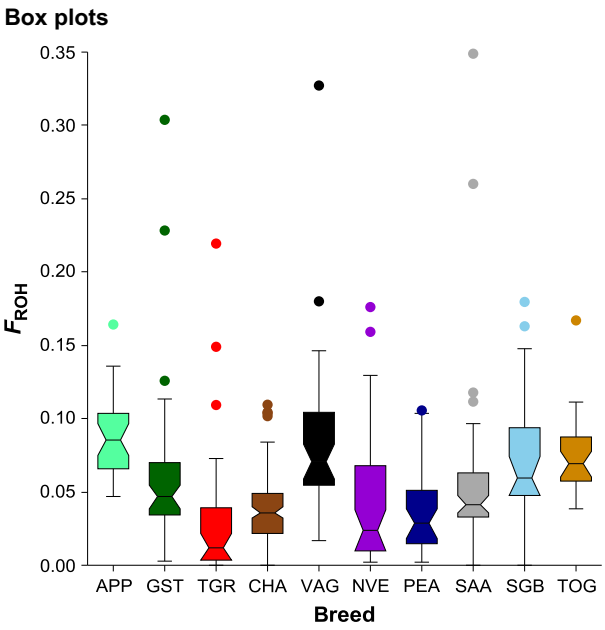

Figure 3 Runs of homozygosity in 10 Swiss goat breeds. Box plots of within-breed F_{ROH} calculated across all the 473 animals. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

Selection signatures

For each breed, 48 windows (1% of the empirical distribution) were considered as putative selection signals, leading to a total of 384 significant windows across all breeds. Of these, 270 (70.3%) windows were significant in one breed. The remaining 114 (29.7%) windows were significant in two or more breeds, with one extreme window on CHI9 showing significant d_i values in five breeds/groups (APP, GST, TGR/NVE/PEA, SAA, TOG; Table 4). The most significant signals were located on CHI6, followed by CHI9, CHI13 and CHI5. The frequency of putative signals observed in one or more (two, three, four or five) breeds and per chromosome is shown in Fig. S8. The genome-wide distribution of the d_i statistic in the eight goat breed groups is given in Fig. 6, where the 99th percentile is indicated by a red dashed line.

The genomic coordinates of the 25 windows with the highest d_i values are listed in Table 4. The window with the highest d_i value (20.29) was found on CHI6 (12.31 Mb) in the APP breed. The same window was also significant in VAG. *PITX2*, a gene that has been reported to influence milk fat, lactose and solid content in dairy goats, is located within this window (Zhao *et al.* 2013). The window with the second highest d_i value was located on CHI2 in SAA and VAG. In total, 21 genes are annotated in this region (22.6–24.4 Mb), without any putative candidate gene (Table S4). On CHI6 (113.3 Mb), the window with the third highest d_i

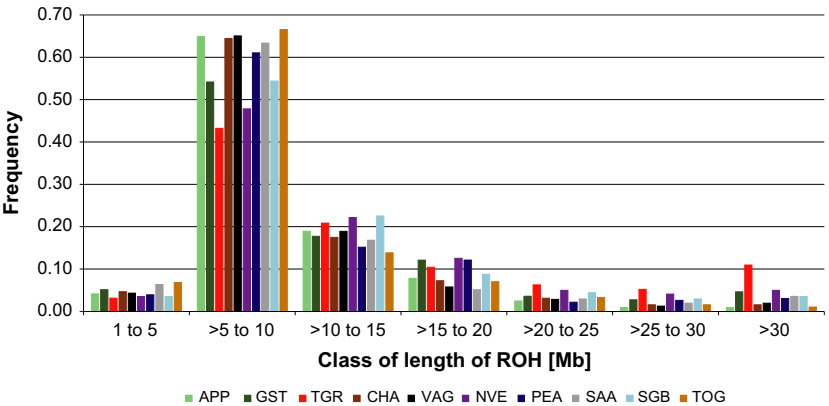

Figure 4 Frequency distribution of the number of ROH in different length classes and for each of the 10 Swiss goat breeds. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

Table 3 Number of animals, mean (min–max) of pedigree- (F_{PED}) and ROH-based (F_{ROH}) inbreeding coefficients and R^2 from linear regressions for all breeds.

Breed	*n*	Ø pedigree completeness	F_{ROH}		F_{PED}		R^2
			Mean	Range	Mean	Range	
APP	26	0.878	0.090	0.047–0.164	0.045	0.000–0.265	0.009
GST	47	0.898	0.059	0.002–0.304	0.039	0.000–0.268	0.877
TGR	29	0.402	0.034	0.000–0.219	0.016	0.000–0.250	0.002
CHA	124	0.974	0.038	0.000–0.110	0.031	0.000–0.092	0.491
VAG	34	0.295	0.082	0.017–0.327	0.015	0.000–0.250	0.594
NVE	32	0.872	0.036	0.002–0.176	0.062	0.000–0.209	0.057
PEA	31	0.936	0.033	0.002–0.106	0.028	0.002–0.073	0.000
SAA	64	0.985	0.055	0.000–0.349	0.044	0.000–0.298	0.672
SGB	20	0.963	0.074	0.024–0.180	0.078	0.020–0.299	0.072
TOG	31	0.995	0.074	0.038–0.167	0.058	0.027–0.160	0.567
Total	438	0.568	0.053	0.000–0.349	0.038	0.000–0.299	0.269

APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

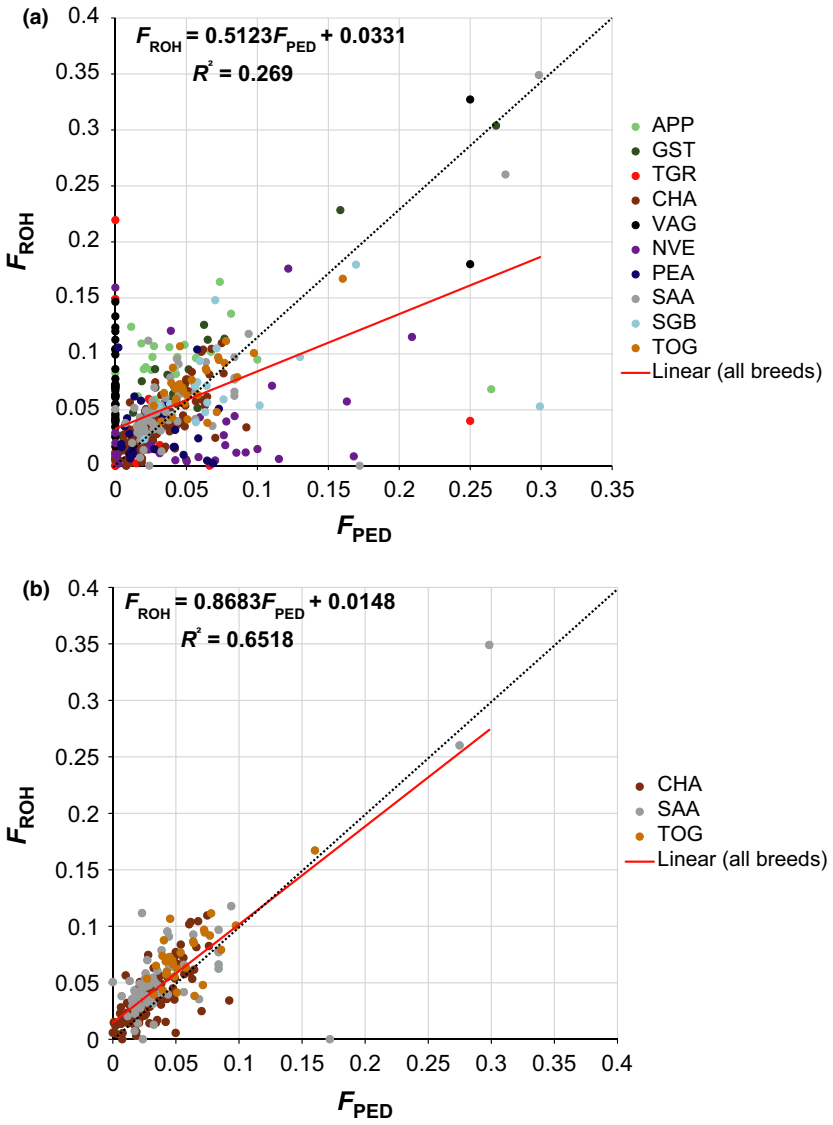

Figure 5 Plot of the regression of F_{ROH} on F_{PED} for (a) all breeds and (b) the three main breeds CHA, SAA and TOG. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

Table 4 Genomic coordinates (chromosome and start- and stop-position) of the 25 windows with highest d_i values, the breed where the d_i value was found for (in brackets, breeds for which the window became significant) and potential candidate genes.

Chr	Position of the window (start–stop, in bp)	d_i value	Candidate gene	Breed(s)
1	80 748 957–82 748 957	10.23		SGB
2	22 639 694–24 639 694	16.60		SAA (VAG)
3	43 218 512–45 218 512	14.53	*DPYD*	SGB (APP, VAG)
3	72 299 239–74 299 239	10.94		GST
4	77 744 721–79 744 721	10.88		GST (TGR, NVE, PEA)
4	82 274 001–84 274 001	10.42	*ING3*, *WNT16*	GST
5	81 288 700–83 288 700	11.45		APP
6	10 291 595–12 291 595	10.17		VAG
6	11 308 917–13 308 917	20.29	*PITX2*	APP (VAG)
6	15 242 801–17 242 801	12.05		APP
6	66 787 300–68 787 300	9.94	*KIT*	APP
6	111 755 706–113 755 706	10.83	*HTT*, *FGFR3*	VAG
6	112 315 248–114 315 248	16.18	*HTT*, *FGFR3*	VAG
8	37 728 077–39 728 077	10.16		SGB
8	83 778 542–85 778 542	10.69		SGB (TGR, NVE, PEA)
9	28 278 843–30 278 843	14.48		VAG
9	71 808 329–73 808 329	10.18	*LRP11*	TOG (APP, GST, SAA, TGR, NVE, PEA)
13	60 180 466–62 180 466	10.92	*ASIP*	CHA (APP)
14	91 134 824–93 134 824	11.46		TOG (APP)
21	0–1 791 309	10.37		APP
21	23 223 533–25 223 533	10.86		TOG (VAG, TGR, NVE, PEA)
22	30 259 191–32 259 191	12.36	*MITF*	VAG (SGB, CHA, TGR, NVE, PEA)
25	18 781 977–20 781 977	11.20		SAA
29	21 228 957–23 228 957	10.36		VAG
29	47 192 954–49 192 954	10.15	*IGF2*, *HRAS*	VAG

APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois Colored; VAG, Valais Goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

value was observed for VAG. Based on the literature, the *HTT* and *FGFR3* genes were selected as candidate genes for this window (Beever *et al.* 2006; Kemper *et al.* 2012; Benjelloun *et al.* 2015). Three genes – *DPYD*, *LOC102175079* and *PTBP2* – were positioned in the region of the window with a d_i value of 14.53 in SGB (Table S4). This window was also significant in APP and VAG. Santana *et al.* (2015) proposed *DPYD* as a candidate gene for the trait rib eye area in Nellore cattle. Rib eye area is related to muscle quantity and, consequently, to body size (Table 4). Seven out of the top 25 d_i values were found in VAG, followed by five in APP and four in SGB. A total of 15 windows were observed in one breed only and the remaining 10 in at least two breeds. The window that was significant in five breeds was in the region of the *LRP11* gene on CHI9 (from 71 808 329 to 73 808 329 bp). Kim *et al.* (2013) mentioned the *LRP11* gene in a haplotype-based association study on protein yield in Holstein cattle. However, 18 other genes are assigned to the region of this window (Table S4).

The *MITF* gene was located within the window on CHI22 (from 30 259 191 to 32 259 191 bp) and was significant in four breeds. The window harboring the *KIT* gene (CHI6) is ranked below the top 25 due to its signal in the APP breed (Table 4). Both genes have been associated with white markings in a range of species (e.g. Pausch *et al.* 2012; Haase *et al.* 2013). *ASIP*, another well-known candidate gene for coat color variation (Cieslak *et al.* 2011), resulted in high d_i values in CHA and APP. *ING3* and *WNT16* have been reported to influence teat color in Holstein cattle (Fan *et al.* 2014) and were also located in one out of the top 25 windows (Table 4). Apart from coat color variation, windows containing genes related to growth (*IGF2*, *HRAS*, *FGFR3*; De Simoni Gouveia *et al.* 2014; Kemper *et al.* 2012) were located in the top 25 signals (Table 4).

Discussion

With 10 local breeds, Switzerland has a remarkable pool of goat genetic resources. However, this pool consists of seven rare breeds (Table S1) with a low economic impact on the national scale. Access to samples from unrelated individuals of rare breeds is hampered by small actual population sizes. During data preparation, the constraint was set at 0.30 for genomic relationships among individuals to be considered for genetic diversity analyses (Data1). This level is still high: if we had set a more rigorous threshold, we would have lost far more of the available genotypes, especially for the highly related rare breeds.

In total, 48 019 SNPs (96%) of the initially available autosomal SNPs from the Illumina caprine 50k BeadChip passed quality control. The proportion of polymorphic SNPs

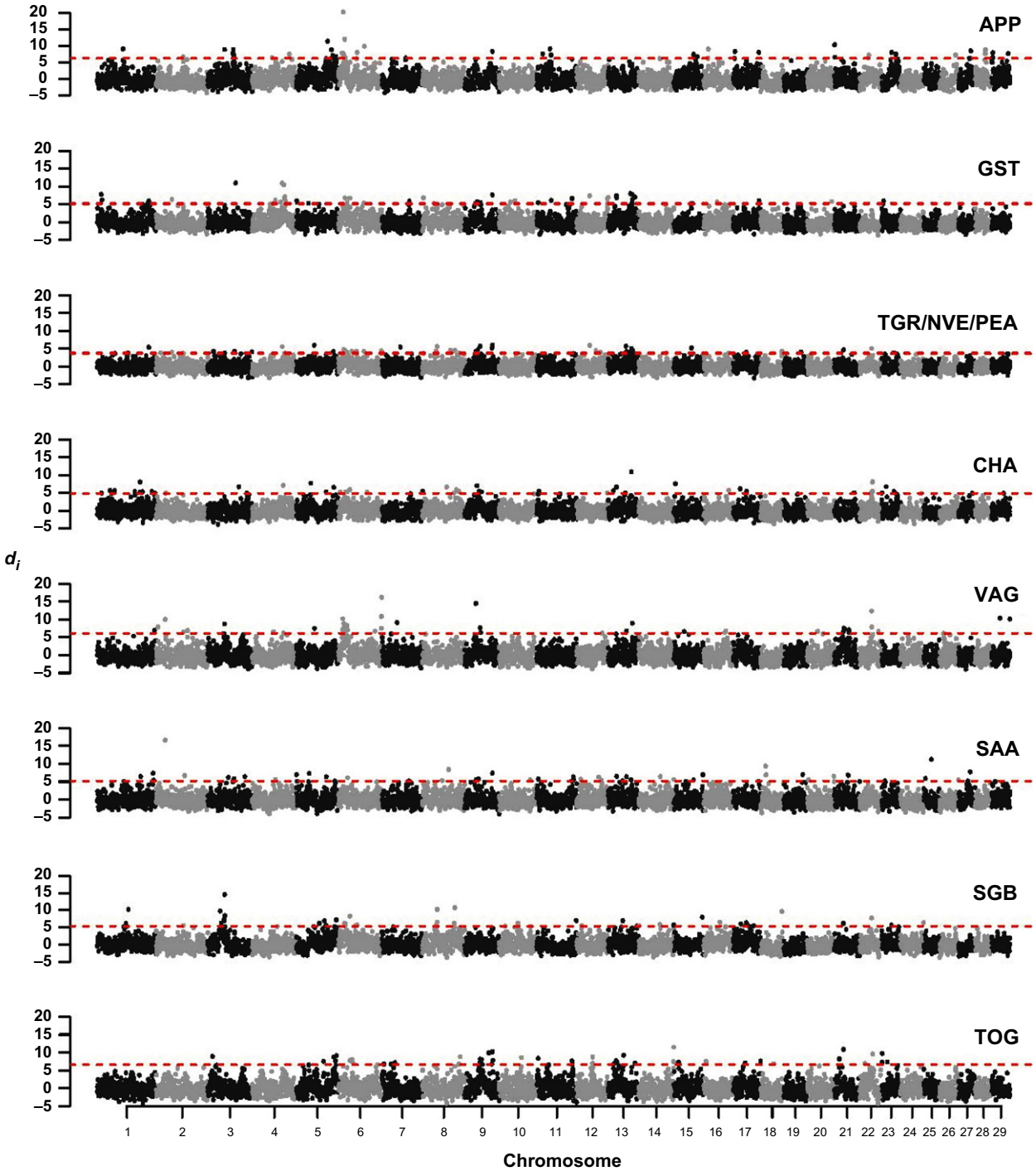

Figure 6 Selection signature detection between breed groups: Genomic distribution of the d_i statistic for all 500-kb windows across all autosomes and the eight breed groups. The dashed red line denotes the 99th percentile for each breed group. APP, Appenzell goat; GST, Grisons striped goat; TGR, Tessin grey goat; CHA, Chamois colored; VAG, Valais goat; NVE, Nera Verzasca goat; PEA, Peacock goat; SAA, Saanen goat; SGB, Booted goat; TOG, Toggenburg goat.

(P_N) was comparatively high among the 10 Swiss breeds (Table 2). These results underline the previously described high levels of genetic polymorphism in goats (Kijas *et al.* 2013; Nicoloso *et al.* 2015) and support the fact that the impact of ascertainment bias can be neglected in the interpretation of further results.

The APP, TOG, VAG and SGB breeds are of note with reference to the different genetic diversity parameters

(Table 2). These breeds show low levels of heterozygosity, mean genomic relationships >0.11 and small recent effective population sizes (Table 1). Low levels of heterozygosity for APP, TOG, VAG and SGB had already been reported by Glowatzki-Mullis *et al.* (2008). In contrast to these four breeds, we determined the highest levels of genetic diversity within TGR and NVE (Table 2).

Given the optimal number of $k = 9$ clusters, TGR and NVE could not be separated using a model-based clustering approach, whereas the other individuals were assigned to eight distinct clusters (Fig. 2). Based on 43 microsatellites, TGR, NVE and PEA were assigned to the same cluster (Glowatzki-Mullis *et al.* 2008). This difference indicates that dense genome-wide SNP data allow the analysis of short-term population structures.

The clear genetic distinction between APP, CHA, VAG, SAA and TOG is supported by PCA analysis (Fig. 1), high-resolution network analysis (Fig. S4) and the neighbor-joining tree (Fig. S5). These breeds are historically assumed to be distinct. Past influence of other breeds is assumed for GST, PEA and SGB. Combining the results from admixture analysis with the parameters of genetic diversity reveals that TGR and NVE may still be experiencing influence from other breeds. The majority of TGR and NVE are regionally restricted to Canton Ticino, and most of the animals traditionally spend the summer months on Alpine pastures in proximity to the Italian border. Against this geographical background, it seems possible that there may be an influence of goat genetic resources from Northern Italy. International collaborations like the international ADAPTmap consortium (Nicolazzi & Stella 2015) are of great value in gaining a better understanding of regional gene flow across countries.

The majority of the ROHs were in the 5 to 10-Mb length class (Fig. 4). This result is in agreement with cattle. A comparison by Purfield *et al.* (2012) of ROH results from high-density data (Illumina 50k BovineHD BeadChip revealed that 50k data could not identify ROHs < 5 Mb. However, the correlation between F_{ROH} and F_{PED} was comparable between the two marker densities, and the authors therefore concluded that 50k data is an appropriate database for the identification of ROH (Purfield *et al.* 2012).

The distribution of F_{ROH} varied between breeds, resulting in different median values. Outlier individuals with $F_{ROH} > 0.15$ were observed for almost all of the breeds (Fig. 3). The animals were sampled on farms, and blood samples were collected from families in some cases. It is known that within-herd inbreeding levels are often higher than the population mean. Consideration of data with a good coverage of different inbreeding classes is more informative for the estimation of the correlation between F_{ROH} and F_{PED}. Therefore, no constraint on genomic relationships was set for sample collection for Data2.

Mean levels of F_{ROH} ranged from 0.033 (PEA) to 0.090 (APP). Mean F_{ROH} values for APP, VAG, SGB and TOG (Table 3) support the low levels of the aforementioned genetic diversity. Kim *et al.* (2015) reported an F_{ROH} of 0.02 for the local Barki goat breed, compared to an F_{ROH} of 0.03 in the exotic breeds sample and an F_{ROH} of 0.09 for the Boer goat sample. The levels of F_{ROH} from that study are comparable with our results. The clear linear relationship between F_{ROH} and F_{PED} (McQuillan *et al.* 2008; Purfield *et al.* 2012) could initially not be confirmed based on Data2 (Fig. 5a). Further examinations provided evidence showing that some animals from rare breeds dissolved the linear relationship between F_{ROH} and F_{PED}. Limitation to animals with pedigree completeness >90% (Fig. S7b) improved the situation ($r_p = 0.599$). A correlation of $r_p = 0.808$ ($R^2 = 0.652$; Fig. 5b) was found after restricting Data2 to the three main populations (CHA, SAA and TOG) characterized by deep pedigree information and a long tradition of routine parentage testing. These results confirm that F_{ROH} based upon 50k data is a good predictor of the pedigree-based inbreeding coefficient in goats. We further conclude that, for rare goat breeds, the limited explanatory power is caused mainly by incomplete and incorrect pedigree information.

A total of 384 selection signatures for various traits were identified based on the d_i statistic proposed by Akey *et al.* (2010). Candidate genes, such as *GNAI3*, *ING3*, *WNT16*, *KIT*, *EDNRB*, *ASIP*, *MITF* for coat color variation (Cieslak *et al.* 2011; Pausch *et al.* 2012; Fan *et al.* 2014; Kim *et al.* 2015); *DPYD*, *HMGA2*, *SPP1*, *FGFR3*, *FBN1*, *BMP2*, *SFRP2*, *FGF2*, *SIGLEC5*, *FASN* and *IGF2* for growth and fatty acid composition (Cole *et al.* 2009; Kemper *et al.* 2012; Cesar *et al.* 2014; De Simoni Gouveia *et al.* 2014; Crispim *et al.* 2015; Frischknecht *et al.* 2015; Kim *et al.* 2015; Santana *et al.* 2015); *GDF9* and *PDGFD* for reproduction traits (Amills 2014; Wei *et al.* 2015); *IL2* for immune response (Pariset *et al.* 2009); and *GDF3*, *PITX2*, *ABCG2*, *SPP1*, *LRP11*, *POFUT1* and *GHRH* for milk production (Yue *et al.* 2010; Li *et al.* 2010; Kim *et al.* 2013; Zhao *et al.* 2013; Gutiérrez-Gil *et al.* 2014; Kim *et al.* 2015) are presented (Tables 4 and S4). Several genomic regions identified here appear to be under selection in cattle (e.g. Qanbari *et al.* 2014), sheep (e.g. Kjias *et al.* 2012) and other domesticated species (e.g. Haase *et al.* 2013). Our results support the suggestion of Kijas *et al.* (2013) that genes such as *KIT* and *MITF* are targets of selection across multiple populations.

PITX2, *ABCG2*, *SPP1* and *GHRH* are four genes previously described as QTL for dairy traits (Yue *et al.* 2010; Zhao *et al.* 2013; Gutiérrez-Gil *et al.* 2014; Kim *et al.* 2015). The *PITX2* gene is located on CHI6 (13.3 Mb) in the window with the highest d_i values out of all 384 windows. Apart from the high d_i value found for the APP breed, the same window harboring *PITX2* was also significant in the VAG breed. The APP breed is known for its low fat content; the mean milk fat content of 2.88% of APP is far below the mean milk fat content for other Swiss goat breeds subject to routine milk analysis (Table S1). No actual data from

routine milk analysis is available for VAG; however, a mean milk fat content of 2.35% was reported by Arbenz *et al.* (1996). Combining this phenotypic evidence with the results from selection signature analysis, it seems plausible that *PITX2* has a major negative impact on the milk fat content in the APP and VAG breed.

The broad effect of the *casein alpha s1* gene (*CSNS1S*) on different dairy phenotypes is well described (Amills 2014). No selection signature was observed in this gene here. It is plausible that the different alleles of the *CSNS1* gene are still segregating in the Swiss goat populations. Potentially high-density SNP data or whole genome sequences would allow the detection of this economically important gene cluster. Among others, *GHR* (CHI20, 31.4 Mb), *DGAT1* (CHI14, 11.3 Mb) and *LIPE* (CHI18, 49.6 Mb) are additional candidate genes related to milk yield and milk composition in goats (Amills 2014) which remained undetected in our study.

High d_i values were also observed for VAG in two windows on CHI6 harboring *FGFR3* (11.4 Mb) and *HTT* (11.3 Mb) (Tables 4 and S4). A large-effect mutation in *FGFR3* is causing limb overgrowth in Suffolk sheep (Beever *et al.* 2006) and achondroplasia in humans (Kemper *et al.* 2012). Benjelloun *et al.* (2015) observed the strongest selective sweep in the *HTT* gene in Moroccan black goat by using whole genome sequences. This gene has been comprehensively studied in humans, where it is associated with Huntington's disease. The authors assume that the gene is involved in adaptation to physiological and pathological conditions leading to endoplasmatic reticulum stress (Benjelloun *et al.* 2015). Further investigations are required to clarify whether the signal we identified is associated with growth traits or adaptation.

The allocation of candidate genes was not possible for several significant signals. For example, a total of 21 genes are annotated in the window with the second highest d_i value found in SAA and VAG on CHI2 (22.6–24.4 Mb; Table 4). However, a review of the literature did not reveal any association between these and the inheritance of phenotypically important traits in mammals. The selection signature in the region of *UGT8* (CHI6) reported by Kim *et al.* (2015) could be confirmed with our study (Table S4) but a direct link to trait of interest could not be drawn. In comparison with cattle, pigs and sheep, studies describing QTL for relevant characters in goats are still sparse. It is to be hoped that this gap will be steadily closed over the next few years. In addition, we fully support the recently proposed requirement for improving the assembly and annotation of the goat genome (Benjelloun *et al.* 2015; Kim *et al.* 2015). The search for relevant genes is further hampered in goats by a high proportion of genes that are annotated but not identified (Benjelloun *et al.* 2015). It is therefore possible that we did not recognize relevant genes due to a lack of identification.

In summary, this study reports on genetic diversity measures, ROHs and signatures of selection in 10 indigenous goat breeds in Switzerland. Based on the results, the highest conservation priority should be allocated to APP, VAG, TOG and SGB. The study highlights several candidate genes and thus contributes toward a better understanding of the genetic architecture of important traits in goats. The amount of available genotypic and phenotypic information is still limited. The successive collection of more data and subsequent validation of the results in future studies is necessary.

Acknowledgements

We thank the Swiss Goat Breeders Federation and ProSpecieRara for the distribution of pedigree information. Various goat breeders and Swiss Quality Testing Services (SQTS) are acknowledged for the provision of blood samples. We also thank the Centre National de Génotypage and NCCR Genomics Platform of the University of Geneva for SNP genotyping. The Swiss Federal Office of Agriculture (OFAG) is acknowledged for the financial support of the study.

References

Akey J.M., Ruhe A.L., Akey D.T., Wong A.K., Connelly C.F., Madeoy J., Nicholas T.J. & Neff M.W. (2010) Tracking footprints of artificial selection in the dog genome. *Proceedings of the National Academy of Sciences of the Unites States of America* **107**, 1160–5.

Alexander D.H., Novembre J. & Lange K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655–64.

Amills M. (2014) The application of genomic technologies to investigate the inheritance of economically important traits in goats. *Advances in Biology* **2014**, 1–13.

Arbenz M., Enz J., Grünenfelder H.P. *et al.* (1996) *Schweizer Ziegen*. Urs Weiss Publishing, Bisikon, Switzerland.

Barbato M., Orozco-ter Wengel P., Tapio M. & Bruford M.W. (2015) SNEP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics* **6**, 1–6.

Becker D., Otto M., Ammann P., Keller I., Drögemüller C. & Leeb T. (2014) The brown coat colour of Coppernecked goats is associated with a non-synonymous variant at the *TYRP1* locus on chromosome 8. *Animal Genetics* **46**, 50–4.

Beever J.E., Smit M.A., Meyers S.N., Hadfield T.S., Bottema C., Albretsen J. & Cockett N.E. (2006) A single-base change in the tyrosine kinase II domain of ovine *FGFR3* causes hereditary chondrodysplasia in sheep. *Animal Genetics* **37**, 66–71.

Benjelloun B., Alberto F.J., Streeter I. *et al.* (2015) Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS data. *Frontiers in Genetics* **6**, 1–14.

Burren A., Signer-Hasler H., Neuditschko M., Tetens J., Kijas J.W., Drögemüller C. & Flury C. (2014) Fine-scale population structure

analysis of seven local Swiss sheep breeds using genome-wide SNP data. *Animal Genetics Resources* **55**, 67–76.

Carillier C., Larroque H., Palhière I., Clément V., Rupp R. & Robert-Granié C. (2013) A first step toward genomic selection in the multi-breed French dairy goat population. *Journal of Dairy Science* **96**, 7294–305.

Carillier C., Larroque H. & Robert-Granié C. (2014) Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. *Genetics Selection Evolution* **46**, 1–10.

Cesar A.S.M., Regitano L.C.A., Mourao G.B. *et al.* (2014) Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genetics* **15**, 1–15.

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. *Giga Science* **4**, 1–16.

Cieslak M., Reissmann M., Hofreiter M. & Ludwig A. (2011) Colours of domestication. *Biological Reviews* **86**, 885–99.

Cole J.B., Van Raden P.M., O'Connell J.R., Van Tassel C.P., Sonstegard T.S., Schnabel R.D., Taylor J.F. & Wiggans G.R. (2009) Distribution and location of genetic effects for dairy traits. *Journal of Dairy Science* **92**, 2931–46.

Crispim A.C., Kelly M.J., Guimarães S.E.F., e Silva F.F, Fortes M.R.S., Wenceslau R.R. & Moore S. (2015) Multi-trait GWAS and new candidate genes annotation for growth curve parameters in Brahman cattle. *PLoS One* **10**, 1–19.

Curik I., Ferenčaković M. & Sölkner J. (2014) Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livestock Science* **166**, 26–34.

De Simoni Gouveia J.J., da Silva M.V.G.B., Paiva S.R. & de Oliveira S.M.P. (2014) Identification of selection signatures in livestock species. *Genetics and Molecular Biology* **37**, 330–42.

Dong Y., Xie M., Jiang Y. *et al.* (2013) Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology* **31**, 135–41.

Fan Y., Wang P., Fu W. *et al.* (2014) Genome-wide association study for pigmentation traits in Chinese Holstein population. *Animal Genetics* **10**, 740–4.

Ferenčaković M., Hamzić E., Gredler B., Curik I. & Sölkner J. (2011) Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *Agriculturae Conspectus Scientificus* **76**, 325–8.

Ferenčaković M., Hamzić E., Gredler B., Solberg T.R., Klemetsdal G., Curik I. & Sölkner J. (2013a) Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics* **130**, 286–93.

Ferenčaković M., Sölkner J. & Curik I. (2013b) Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution* **45**, 1–9.

Frischknecht M., Jagannathan V., Plattet P. *et al.* (2015) A non-synonymous *HMGA2* variant decreases height in Shetland Ponies and other small horses. *PLoS One* **10**, e0140749.

Glowatzki-Mullis M.L., Muntwyler J., Bäumle E. & Gaillard C. (2008) Genetic diversity measures of Swiss goat breeds as decision-making support for conservation policy. *Small Ruminant Research* **74**, 202–11.

Guillot G., Estoup A., Mortier F. & Cosson J.F. (2005) A spatial statistical model for landscape genetics. *Genetics* **170**, 1261–80.

Gutiérrez-Gil B., Arranz J.J., Pong-Wong R., Garcia-Gamez E., Kijas J. & Wiener P. (2014) Application of selection mapping to identify genomic regions associated with dairy production in sheep. *PLoS One* **9**, e94623.

Haase B., Signer-Hasler H., Binns M.M., Obexer-Ruff G., Hauswirth R., Bellone R.R., Burger D., Rieder S., Wade C.M. & Leeb T. (2013) Accumulating mutations in series of haplotypes at the *KIT* and *MITF* loci are major determinants of white markings in Franches Montagnes horses. *PLoS One* **8**, e75071.

Herrero-Medrano J.M., Megens H.J., Groenen M.A.M., Ramis G., Bosse M., Pérez-Enciso M. & Crooijmans R.P.M.A. (2013) Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genetics* **14**, 1–13.

Huson D.H. & Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–67.

Li H., Wang Z., Moore S.S., Schenkel F.S. & Stothard P. (2010) Genome-wide scan for positional and functional candidate genes affecting milk production traits in Canadian Holstein Cattle. *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig, Germany.

Kemper K.E., Visscher P.M. & Godder M.E. (2012) Genetic architecture of body size in mammals. *Genome Biology* **13**, 1–13.

Kijas J.W., Lenstra J.A., Hayes B., Boitard S., Porto Neto L.R., San Cristobal M., Servin B, McCulloch R, Whan V, Gietzen K., Paiva S., Barendse W., Ciani E., Raadsma H., McEwan J., Dalrymple B. & other members of the International Sheep Genomics Consortium. (2012) Genome-wide analysis of the World's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology* **10**, e1001258.

Kijas J.W., Ortiz J.S., McCulloch R., James A., Brice B., Swain B., Tosser-Klopp G. & the International Goat Genome Consortium. (2013) Genetic diversity and investigation of polledness in divergent goat populations using 52 088 SNPs. *Animal Genetics* **44**, 325–35.

Kim E.S., Cole J.B., Huson H., Wiggans G.R., Van Tassel C.P., Crooker B.A., Liu G., Da Y. & Sonstegard T.S. (2013) Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS One* **8**, e80813.

Kim E.S., Elbeltagy A.R., Aboul-Naga A.M., Rischkowsky B., Sayre B., Mwacharo J.M. & Rothschild M.F. (2015) Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity* **10**, 1–10.

MacCluer J.W., Boyce A.J., Dyke B., Weitkamp L.R., Pfennig D.W. & Parsons C.J. (1983) Inbreeding and pedigree structure in Standardbred horses. *The Journal of Heredity* **74**, 394–9.

McQuillan R., Leutenegger A.L., Abdel-Rahman R. *et al.* (2008) Runs of homozygosity in European populations. *American Journal of Human Genetics* **83**, 659–72.

Mucha S., Bunger L. & Conington J. (2015) Genome-wide association study of footrot in Texel sheep. *Genetics Selection Evolution* **47**, 1–10.

Neuditschko M., Khatkar M.S. & Raadsma H.W. (2012) NetView: a high-definition network-visualization approach to detect fine-scale population structures from genome-wide patterns of variation. *PLoS One* **7**, 1–13.

Nicolazzi E.L. & Stella A. (2015) ADAPTMAP: a first instight into goat adaptation. Plant and Animal Genome Conference XXIII, San Diego.

Nicolazzi E.L., Picciolini M., Strozzi F., Schnabel R.D., Lawley C., Pirani A., Brew F. & Stella A. (2014) SNPchiMp: A database to

disentangle the SNPchip jungle in bovine livestock. *BMC Genomics* **15**, 1–6.

Nicoloso L., Bomba L., Colli L. *et al.* (2015) Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* **47**, 1–10.

Pariset L., Joost S., Ajmone Marsan P., Valentini A. & Econogene Consortium (EC). (2009) Landscape genomics and biased FST approaches reveal single nucleotide polymorphisms under selection in goat breeds of North-East Meditrranean. *BMC Genetics* **10**, 1–8.

Pausch H., Wang X., Jung S., Krogmeier D., Edel C., Emmerling R., Götz K.U. & Fries R. (2012) Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. *PLoS One* **7**, e36346.

Purcell S., Neale B., Todd-Brown K., Thomas L. & Ferreira M.A.R (2007) PLINK: a toolset for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Purfield D., Berry D., McParland S. & Bradley D. (2012) Runs of homozygosity and population history in cattle. *BMC Genetics* **13**, 70.

Qanbari S., Pausch H., Jansen S., Somel M., Strom T.M., Fries R., Nielsen R. & Simianer H. (2014) Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genetics* **10**, e1004148.

Reber I., Keller I., Becker D., Flury C., Welle M. & Drögemüller C. (2015) Wattles in goats are associated with the *FMN1/GREM1* region on chromosome 10. *Animal Genetics* **46**, 316–20.

Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–8.

Santana M.H.A., Ventura R.V., Utsunomiya Y.T. *et al.* (2015) A genomewide association mapping study using ultrasound scanned information identifies potential genomic regions and candidate genes affecting carcass traits in Nellore cattle. *Journal of Animal Breeding and Genetics* **132**, 420–7.

Sargolzaei M., Iwaisaki H. & Colleau J.J. (2005) A fast algorithm for computing inbreeding coefficients in large populations. *Journal of Animal Breeding and Genetics* **122**, 325–31.

Sargolzaei M., Iwaisaki H. & Colleau J.J. (2006) CFC: a tool for monitoring genetic diversity. *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, Belo Horizonte, Minas Gerais, Brazil, 13–18 August, 27–8.

Silió L., Rodríguez M.C., Fernández A., Barragán C., Benítez R., Óvilo C. & Fernández A. (2013) Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics. *Journal of Animal Breeding and Genetics* **130**, 349–60.

Szpiech Z.A., Jakobsson M. & Rosenberg N.A. (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498–504.

Tosser-Klopp G., Bardou P., Bouchez O. *et al.* (2014) Design and characterization of a 52K SNP chip for goats. *PLoS One* **9**, e86227.

Wei C., Wang H., Liu G., Wu M., Cao J., Liu Z., Liu R., Zhao F., Zhang L., Lu J., Liu C. & Du L. (2015) Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. *BMC Genomics* **16**, 1–12.

Wright S. (1922) Coefficients of inbreeding and relationship. *The American Naturalist* **56**, 330–8.

Wright S. (1943) Isolation by distance. *Genetics* **28**, 114–38.

Yue W., Fang X., Zhang C., Pang Y., Xu H. & Gu C. (2010) Two novel SNPs of the *ABCG2* gene and its associations with milk traits in Chinese Holsteins. *Molecular Biology Reports* **38**, 2927–32.

Zhao H., Wu X., Cai H., Pan C., Lei C., Chen H. & Lan X. (2013) Genetic variants and effects on milk traits of the caprine *paired-like homeodomain transcription factor 2* (*PITX2*) gene in dairy goats. *Gene* **532**, 203–10.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 Development of LD (R^2) for the 10 breeds and different distances.

Figure S2 Boxplot of genomic relationships for the 10 breeds.

Figure S3 Development of the cross-validation for different number of clusters (k).

Figure S4 Dynamic network by considering the admixture results $k = 9$.

Figure S5 Neighbor-joining tree for the 10 breeds under investigation.

Figure S6 Relation number of ROH compared to sum of ROH.

Figure S7 Plot of the regression of F_{ROH} on F_{PED} with pedigree completeness >90% for (a) all breeds and (b) the three main breeds CHA, SAA and TOG.

Figure S8 Frequency of significant windows ($n = 384$) observed in one, two, three, four or five breeds and chromosome respectively.

Table S1 Phenotypic characteristics of the coat, hair, horns, body-size, performance and distinction between main and rare breeds of the 10 Swiss goat breeds.

Table S2 Overview of the number of informative SNPs, total length, average distance between adjacent SNPs, maximum distance of adjacent SNP and average r^2 (linkage disequilibrium) between adjacent markers of the 29 autosomes.

Table S3 Breed-specific development of N_e over time.

Table S4 List of genes 1 Mb up- and downstream of the middle position of the 25 top significant windows and the eight breed groups.