

RESEARCH ARTICLE

Open Access



Signatures of selection reveal candidate genes involved in economic traits and cold acclimation in five Swedish cattle breeds

Seyed Mohammad Ghoreishifar¹, Susanne Eriksson^{2*} , Anna M. Johansson², Majid Khansefid³, Sima Moghaddaszadeh-Ahrabi⁴, Nahid Parna¹, Pourya Davoudi⁵ and Arash Javanmard⁶

Abstract

Background: Thousands of years of natural and artificial selection have resulted in indigenous cattle breeds that are well-adapted to the environmental challenges of their local habitat and thereby are considered as valuable genetic resources. Understanding the genetic background of such adaptation processes can help us design effective breeding objectives to preserve local breeds and improve commercial cattle. To identify regions under putative selection, GGP HD 150 K single nucleotide polymorphism (SNP) arrays were used to genotype 106 individuals representing five Swedish breeds i.e. native to different regions and covering areas with a subarctic cold climate in the north and mountainous west, to those with a continental climate in the more densely populated south regions.

Results: Five statistics were incorporated within a framework, known as de-correlated composite of multiple signals (DCMS) to detect signatures of selection. The obtained *p*-values were adjusted for multiple testing (FDR < 5%), and significant genomic regions were identified. Annotation of genes in these regions revealed various verified and novel candidate genes that are associated with a diverse range of traits, including e.g. high altitude adaptation and response to hypoxia (*DCAF8*, *PPP1R12A*, *SLC16A3*, *UCP2*, *UCP3*, *TIGAR*), cold acclimation (*AQP3*, *AQP7*, *HSPB8*), body size and stature (*PLAG1*, *KCNA6*, *NDUFA9*, *AKAP3*, *CSH12orf4*, *RAD51AP1*, *FGF6*, *TIGAR*, *CCND2*, *CSMD3*), resistance to disease and bacterial infection (*CHI3L2*, *GBP6*, *PPFIBP1*, *REP15*, *CYP4F2*, *TIGD2*, *PYURF*, *SLC10A2*, *FCHSD2*, *ARHGEF17*, *RELT*, *PRDM2*, *KDM5B*), reproduction (*PPP1R12A*, *ZFP36L2*, *CSPP1*), milk yield and components (*NPC1L1*, *NUDCD3*, *ACSS1*, *FCHSD2*), growth and feed efficiency (*TMEM68*, *TGS1*, *LYN*, *XKR4*, *FOXA2*, *GBP2*, *GBP5*, *FGD6*), and polled phenotype (*URB1*, *EVA1C*).

Conclusions: We identified genomic regions that may provide background knowledge to understand the mechanisms that are involved in economic traits and adaptation to cold climate in cattle. Incorporating *p*-values of different statistics in a single DCMS framework may help select and prioritize candidate genes for further analyses.

Background

According to the natural selection theory, favorable mutations may increase the chance of individuals to survive and reproduce, which results in an increased frequency of fitness-related alleles in a species/breed over generations [1]. Likewise, human-mediated selection

may increase the frequency of favorable alleles associated with economic traits. Therefore, the combination of these two events shapes the genetic architecture of cattle breeds by leaving footprints in the genome that might be detectable. The identification of signatures of selection becomes more complicated when demographic events such as population bottleneck, admixture, genetic drift, and inbreeding have occurred. However, when selection footprints are successfully identified, they can contribute to a better understanding of the processes that cause

*Correspondence: susanne.eriksson@slu.se

² Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

diversity among breeds, and to pinpoint the causal variants that are involved in a phenotype under selection.

The recent development of cost-effective genotyping technologies has made it possible to scan the genome to uncover regions that are under putative selection. Similarly, different statistical tests have been conceptualized to detect signatures of selection and pinpoint the causal variants by using various models. The rationale behind selection signature theory is that the frequency of alleles that are under selection can vary in opposite directions (low or high), and thus results in stretches of consecutive homozygous genotypes [2], or in modifying the length and frequency of haplotypes around the region. Several statistical tests have been proposed to detect selection signals e.g. [3–6]. One limitation of using a single statistical test is that the variant under selection may have an effect on adjacent loci, depending on the extent of linkage disequilibrium, which can reduce the resolution of the mapping of selection signatures [7]. Different single statistical tests often do not return consistent results [8, 9]. The inconsistency among results may be attributed to factors such as different sensitivity to sampling design, details of the selective sweep, and the demographic history of populations under study [8–10]. A popular strategy to overcome these issues is to prioritize selection signals detected by several single statistical tests e.g. [11, 12]; but following this strategy, we may lose some loci that show weak signals [9]. Hence, strategies based on combining *p*-values of different test statistics (composite measures of selection) have been developed [13–15]. One of these is the de-correlated composite of multiple signals (DCMS) [15] used in this study. Ma et al. [15] reported that the resolution of selection signature mapping and the power of detecting selection signals were improved by using DCMS compared to most single statistics. Moreover, composite measures such as DCMS have been reported to identify the causal variants (i.e. the variants under selection in the detected signature regions) more precisely [7].

According to a FAO report, ~1000 distinct cattle breeds are listed and are geographically distributed in different countries around the world. Swedish cattle breeds (*Bos taurus*) include both indigenous/native and commercial breeds [16]. Whereas the latter are economically superior regarding production of meat and milk, the former are well adapted to the challenges imposed by their local environment [16], and show superior phenotypes for certain traits. One example is the favorable milk protein composition for cheese-making of the two Swedish mountain breeds: Fjäll and Fjällnära cattle [16, 17]. Until the 20th century, the old local breeds were of major importance in Sweden, representing several hundred thousand heads [18]. The number of individuals

decreased rapidly in the early to mid-20th century due to the development of more efficient breeding programs using e.g. artificial insemination (AI), and to the increased competition from breeds with a higher milk yield [19]. However, due to the growing awareness of the value of local breeds as genetic resources, several actions have been adopted for their preservation. The native Swedish cattle breeds still display considerable phenotypic diversity in terms of coat color, body weight and size, milk content, and polled phenotype, as well as substantial genetic variation, as described by Upadhyay et al. [16]. Moreover, indigenous breeds may carry specific gene variants that contribute to adaptation to their local environment. Knowledge about the genomic regions that display signatures of selection in breeds adapted to different environments is of great value to design future breeding strategies, both in local breeds and in larger commercial breeds [20, 21].

Well-documented studies on signatures of selection have shown that they can help identify polymorphisms and/or candidate genes that underlie economical traits in cattle breeds. As reviewed by Gutiérrez-Gil et al. [22], several genes have been identified, for example *ATP binding cassette subfamily G member 2 (junior blood group)* (*ABCG2*), which is associated with milk composition, *coiled-coil-helix-coiled-coil-helix domain containing 7* (*CHCHD7*) and *PLAG1 zinc finger* (*PLAG1*) with body size, *diacylglycerol O-acyltransferase 1* (*DGAT1*) with milk production, *XK related 4* (*XKR4*) with growth trait, and *melanocortin 1 receptor* (*MC1R*) and *KIT proto-oncogene, receptor tyrosine kinase* (*KIT*) with coat color and spotting. Such findings are important to understand the mechanisms that explain the phenotypic diversity among breeds. However, these studies were often based on single statistical tests which have limited power to detect selection signatures. Thus, the purpose of our study was to combine multiple statistics of signatures of selection within a single DCMS framework [15] by taking the correlation between them into account to detect signatures of selection in the genome of Swedish indigenous and commercial cattle breeds with greater statistical power and higher resolution.

Methods

Sample collection and data quality control

Genotype data of Swedish cattle breeds obtained with GeneSeek® Genomic Profiler High-Density Bovine 150 K (GGP HD150K) single nucleotide polymorphisms (SNPs) array that were previously described in a study by Upadhyay et al. [16] were downloaded from the DRYAD (<https://datadryad.org/>) public data repository. Five breeds including Fjällnära (n=16), Fjäll (n=23), Swedish Holstein–Friesian (n=24), Swedish Red (n=25),

and Swedish Red Polled ($n=18$) were retained for subsequent analyses (Table 1), whereas the remaining four breeds with a small sample size were discarded. All individuals from the five retained breeds, except one Swedish Red Polled individual born in 2016, were born between the mid-1970s and the early 2000s [16]. Inclusion of very close relatives such as full-sibs or parent–offspring–pairs was avoided as much as possible by using available information on the location of the farms and on pedigrees [16]. Detailed information about sample collection and DNA extraction is in [16]. This dataset overlapped partially with that used in the study by Johansson et al. [17].

Data quality control was performed using the PLINK v1.9 [23] software, separately for each of the two breed groups, including the northern breeds (Fjällnära and Fjäll) and middle–southern Swedish breeds (Swedish Holstein–Friesian, Swedish Red, and Swedish Red Polled) (see Table 1). SNPs with a call rate lower than 0.95 and a MAF lower than 0.05, and those for which the Hardy–Weinberg equilibrium Chi square test p -value was lower than 10^{-6} were discarded. In addition, SNPs that were duplicated in the map file, or located on a sex chromosome, and/or had an unidentified position on the UMD3.1 assembly [24] were removed using the PLINK --exclude option. After merging the two datasets (using PLINK's --merge option), a subset of 105,362 mutual SNPs and 102 individuals (with an overall call rate of 99.92%) remained for subsequent analyses. It should be noted that within-group quality control was performed to provide high-quality SNPs for haplotype phasing as requested by SHAPEIT2 [25].

Principal component (PC) analysis

We used the PLINK --ibs-matrix command to estimate the identity-by-state (IBS) matrix between individuals. The output was used to perform a principal component

(PC) analysis of genetic distances with the *prcomp* R function to visualize the distribution of samples, and the results were plotted by using R (<https://www.r-project.org/>).

De-correlated composite of multiple signals (DCMS)

In this study, five statistics including fixation index (F_{ST}) [3], haplotype homozygosity (H1) [5], modified haplotype homozygosity (H12) [5], Tajima's D index [6], and nucleotide diversity (π) [4] were combined into a single DCMS framework [15] as described in Yurchenko et al. [20]. DCMS is similar to composite selection signals (CSS) [13] by combining p -values, with the difference that it takes the respective correlation between the various statistics into account [7, 15]. The DCMS statistic can be calculated at position l as follows [15]:

$$DCMS_l = \sum_{t=1}^n \frac{\log \left[\frac{1-p_{lt}}{p_{lt}} \right]}{\sum_{i=1}^n |r_{it}|}, \quad (1)$$

where p_{lt} shows the p -value at position l for statistic t ; r_{it} refers to the correlation between the test statistic of the i^{th} and t^{th} methods, and n is the total number of test statistics (combined) in the DCMS. The expression $1/\sum_{i=1}^n |r_{it}|$ is called weight factor, which ranges from $1/n$ to 1. For example, in a given dataset with $n=3$ different uncorrelated ($r_{i \neq j} = 0$) test statistics (t), the weight factor will be 1 for each t and $DCMS_l$ will be the sum of the $\log((1-p_{lt})/p_{lt})$; if the three statistics are fully correlated ($r_{i \neq j} = 1$), the weight factor for each statistic (t) will be $1/3$, and $DCMS_l$ will be the average of $\log((1-p_{lt})/p_{lt})$. Indeed, the weight factors help avoiding excessive contribution of highly correlated statistics in the DCMS calculation.

Table 1 Descriptive statistics for the studied Swedish cattle breeds

Breed	N samples	Type (distribution)	Characteristic	Citation
FNC	16	Native (north)	Smaller body size than the Fjäll breed; lower milk yield; said to be hardy and hold strong ability to find food in natural pastures	[16]
SMC	23	Native (north)	Small body size of cows about 400–450 kg; said to be hardy and hold strong ability to find food in natural pastures	[16]
SHF	24	Commercial (south)	The old Friesian type was more of a dual-purpose; the SHF is a commercial dairy breed with live weight of cows around 700 kg	[16]
SRC	25	Commercial (south middle to south)	A commercial dairy breed; live weight of cows 550–650 kg	[16]
SRP	18	Native (middle)	The focus is on conservation; said to be hardy; live weight of cows 350–600 kg	[16]

Breed: FNC (Fjällnära Cattle), SMC (Fjäll also known as Swedish Mountain Cattle), SHF (Swedish Holstein–Friesian), SRC (Swedish Red Cattle), SRP (Swedish Red Polled)
 N samples: number of genotyped samples (i.e. before data quality control); note that according to data quality control, four samples (due to animal call-rate < 0.95) were removed, and principal component (PC) analyses, three samples (due to locating outside their expected breed cluster) were removed; resulting in 99 individuals for analyses of selection signatures

Haplotype-based H1 and H12 statistics

Effective population size (N_e) is a required parameter in haplotype phasing. It was initially estimated, separately, for each of the two breed groups using the SNeP software [26] with the parameters described in [27]. SHAPEIT2 [25] was used for haplotype phasing of the autosomal genome, also separately for each breed group. In SHAPEIT2, we set the parameters conditioning states to 400 (--states 400), and the N_e to 108 (--effective-size 108) for the group of southern–middle breeds and 60 (--effective-size 60) for the group of northern breeds. A bovine genetic map [28] was used to accompany SHAPEIT2 for haplotype phasing in order to correct for the variation in recombination rate along the cattle genome. Then, using a customized R script, the phased haplotypes were transformed to the format requested by the H12_H2H1.py script (<https://github.com/ngarud/SelectionHapStats>). This python script with a window size of 14 SNPs and step size of 1 (-window 14 -jump 1) was run to estimate the H1 and H12 statistics for each autosome and each breed, as described in [20].

Tajima's D and nucleotide diversity (π) statistics

Both Tajima's D and π statistics were estimated with the vcftools software [29]. Tajima's D statistics were calculated for each breed and chromosome, separately, using the --TajimaD function considering non-overlapping sliding windows of 300 Mb (--TajimaD 300000). The estimated D values for each 300-Mb bin were assigned to SNPs within that bin, and missing values were converted into zeros. π statistics were calculated for each breed and chromosome separately with the --site- π function; and in order to reduce noise, the outputs were smoothed for each chromosome through the R's *runmed* function with a window size of 31 SNPs ($k=31$, *endrule="constant"*) as described in [20].

Fixation index (F_{ST})

Fixation index, which is a measure of population differentiation, was calculated for each SNP and each breed against the remaining samples from the other breeds using PLINK --fst and --within functions. F_{ST} values less than 0 were converted into zeros and the statistics were then smoothed using the *runmed* function as described for the π statistics.

Calculation of the DCMS statistic

The genome-wide DCMS was calculated for each breed, separately, by combining the five statistics (H1, H12, Tajima's D , π , and F_{ST}) for each SNP. Initially, for each statistic within a given breed, genome-wide p -values based on fractional ranks (i.e. the *stat_to_pvalue* function

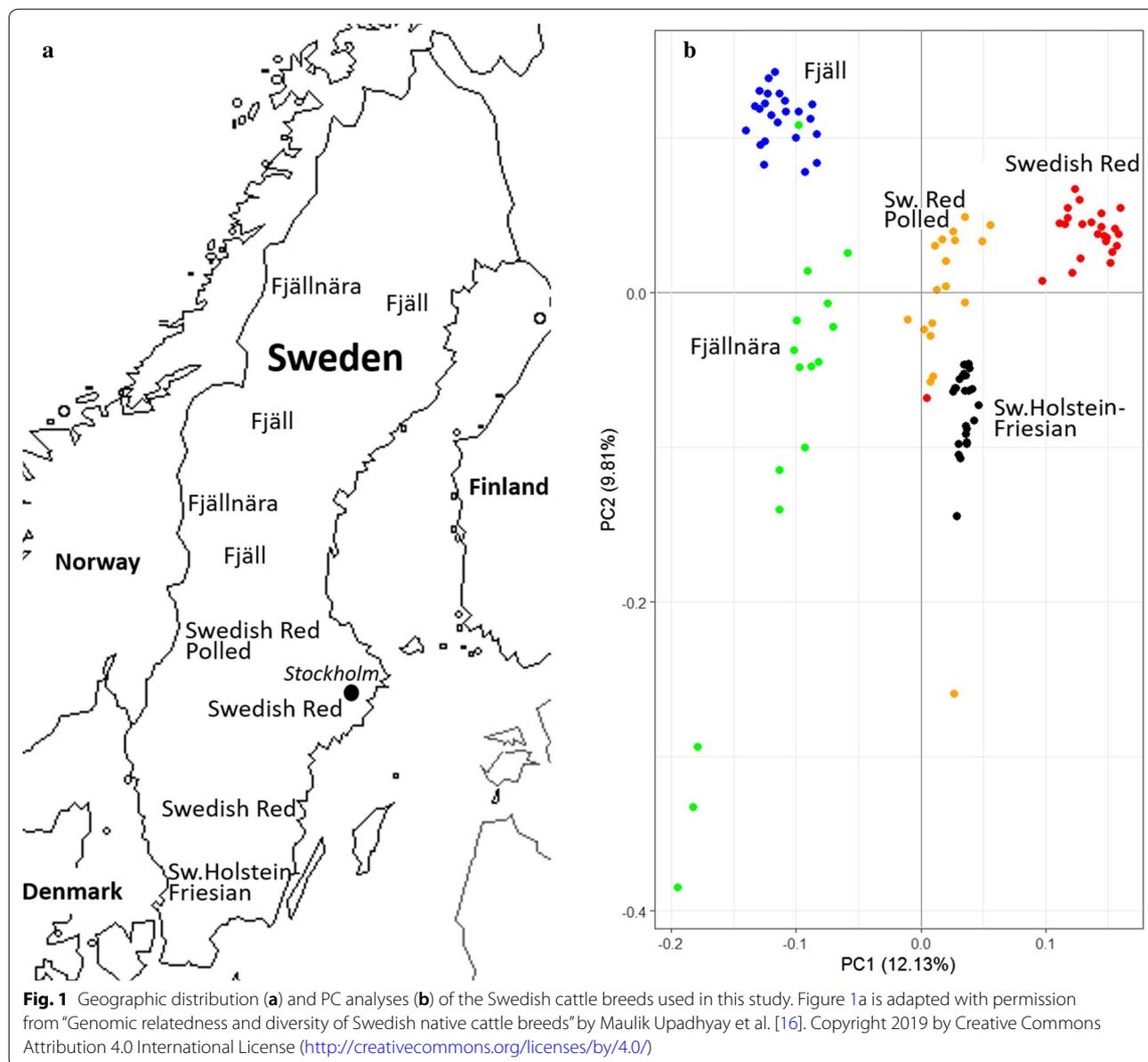
of the R MINOTAUR package [30]) were calculated for all the SNPs. To this end, a left tailed test for the π and D statistics (two.tailed=FALSE, right.tailed=FALSE) and a right tailed test for the H1, H12, and F_{ST} statistics (two.tailed=FALSE, right.tailed=TRUE) were applied. Then, we called the *covNAMcd* function ($\alpha=0.75$, $nsamp=50,000$) from the *rrcovNA* R package [31] to calculate an $n \times n$ correlation matrix (i.e. the minimum covariance determinant estimator of multivariate location and scatter) between the included statistics (where n represents the number of statistics to estimate the DCMS values). This matrix was used as input in the *DCMS* function of the MINOTAUR R package [30] to calculate genome-wide DCMS values. Once the DCMS values were generated, they were fitted to a normal distribution using the robust linear model (*rlm*) function of the MASS R package [32] as described in [20, 21]: *model=rlm* (*dcms~1*), in which the *dcms* object is a vector containing the raw DCMS values. The outputs of the fitted model (i.e. μ [mean] and SD [standard deviation]) were used as input in the *pnorm* R function to calculate the p -values of the DCMS statistics: *dcms_pvalues=pnorm* ($q=dcms$, $mean=\mu$, $sd=SD$, *lower.tail=FALSE*). In order to control multiple testing false discovery rate (FDR) among rejected null hypotheses [33], the DCMS p -values were transformed into the corresponding q -values according to the Benjamini and Hochberg method [34] using the *p.adjust* R function: *dcms_qvalues=p.adjust* ($p=dcms_pvalues$, *method="BH"*, $n=length(dcms_pvalues)$).

Gene annotation

Genes that were located in genomic regions including consecutive SNPs with a q -value lower than 0.05 were considered as statistically significant intervals and the boundary for each interval was determined by searching for the first flanking SNP showing a q -value higher than 0.1. Then, the protein coding genes were extracted from the significant regions based on the UMD3.1 bovine reference genome assembly [24]. Finally, we performed an extensive review of the literature to annotate functions of the identified genes.

Results

As illustrated in Fig. 1a, five cattle breeds that originated from the northern, middle and southern parts of Sweden were included in our study. According to the PC analyses (Fig. 1b), the first PC clearly differentiated our dataset into two clusters of breeds, i.e. the northern breeds (Fjäll and Fjällnära) and the middle–southern breeds (Swedish Holstein–Friesian, Swedish Red Cattle, and Swedish Red Polled). PC analysis of these cattle breeds (Fig. 1b) led us to remove three samples that were located outside of their expected breed cluster. Therefore, 99 individuals



(genotyped for 105,362 SNPs) remained for analyses of signatures of selection. Table 1 presents the characteristics and descriptive statistics for each breed. The average distance between the adjacent SNP pairs among the 105,362 SNP genotypes was 24.89 (± 21.48) kb, which represented 2.62 Gb of the UMD 3.1 bovine genome assembly. The average distance between adjacent SNP pairs ranged from 24.36 on *Bos taurus* chromosome (BTA)27 to 25.29 on BTA11 (data not shown).

De-correlated composite of multiple signals (DCMS)

After calculation of the within-breed DCMS statistics for 105,362 SNPs, p -values were fitted to a

normal distribution and corrected for multiple testing ($FDR < 0.05$). We identified 58, 37, 38, 39, and 51 genomic regions for the Fjällnära, Fjäll, Swedish Holstein–Friesian, Swedish Red, and Swedish Red Polled breeds, respectively (Table 2) and [see Additional file 1: Table S1]. The average length of the significant regions ranged from 142.3 (± 187.1) kb in Swedish Red Polled to 315.4 (± 400.8) kb in Swedish Red Cattle (Table 2). The total length of the genome covered with significant signals and the number of identified protein coding genes between breeds ranged from 72.6 to 134.1 Mb and from 61 to 99, respectively (Table 2). Overall, 355 unique protein coding genes were detected within the significant intervals for all

Table 2 Descriptive statistics of the genomic regions identified through the de-correlated composite of multiple signals (DCMS) in the five Swedish native cattle breeds

Breed	N regions	Average \pm SD (Kb)	N SNPs	Total size (Mb)	N genes
Fjällnära Cattle	58	231.1 \pm 298.1	719	134.1	90
Fjäll Cattle	37	229.5 \pm 226.2	421	84.9	61
Sw. Holstein–Friesian	38	249.1 \pm 289.4	543	94.7	99
Sw. Red Cattle	39	315.4 \pm 400.8	641	123.0	84
Sw. Red Polled	51	142.3 \pm 187.1	499	72.6	61

N Regions: number of segments identified using the DCMS method as significant genomic regions harboring signatures of selection

N SNPs: number of SNPs identified within the regions detected as signatures of selection using the DCMS method (q -value $<$ 0.05)

N genes: number of protein coding genes identified within the significant regions detected by the DCMS method

the breeds (Table 2) and [see Additional file 2: Table S2], among which several were detected in more than one breed [see Additional file 3: Table S3].

The distribution of the regions of signatures of selection across the genome of the five Swedish breeds is represented in Fig. 2. The most significant genomic regions were identified on BTA5 in Fjällnära (BTA5: 105,75–106,52), on BTA1 in Fjäll (BTA1: 2,25–2,52), and on BTA14 in Swedish Holstein–Friesian (BTA14: 24,42–25,11 & 14: 25,35–25,73), Swedish Red Cattle (BTA14: 24,00–24,89), and Swedish Red Polled (BTA14: 24,00–24,26) breeds (Fig. 2) and [see Additional file 1: Table S1].

Gene annotation of the identified regions detected various verified and novel candidate genes that are associated with a diverse range of traits including high altitude adaptation and response to hypoxia (*DCAF8*, *PPP1R12A*, *SLC16A3*, *UCP2*, *UCP3*, *TIGAR*), cold acclimation (*AQP3*, *AQP7*, *HSPB8*), body size and stature (*PLAG1*, *KCNA6*, *NDUFA9*, *AKAP3*, *C5H12orf4*, *RAD51API*, *FGF6*, *TIGAR*, *CCND2*, *CSMD3*), resistance to disease and bacterial infection (*CHI3L2*, *GBP6*, *PPFIBP1*, *REP15*, *CYP4F2*, *TIGD2*, *PYURF*, *SLC10A2*, *FCHSD2*, *ARHGEF17*, *RELT*, *PRDM2*, *KDM5B*), reproduction (*PPP1R12A*, *ZFP36L2*, *CSPP1*), milk yield and components (*NPC1L1*, *NUDCD3*, *ACSSI*, *FCHSD2*), growth and feed intake (*TMEM68*, *TGS1*, *LYN*, *XKR4*, *FOXA2*, *GBP2*, *GBP5*, *FGD6*), and polled phenotype (*URB1*, *EVA1C*) (see Table 3).

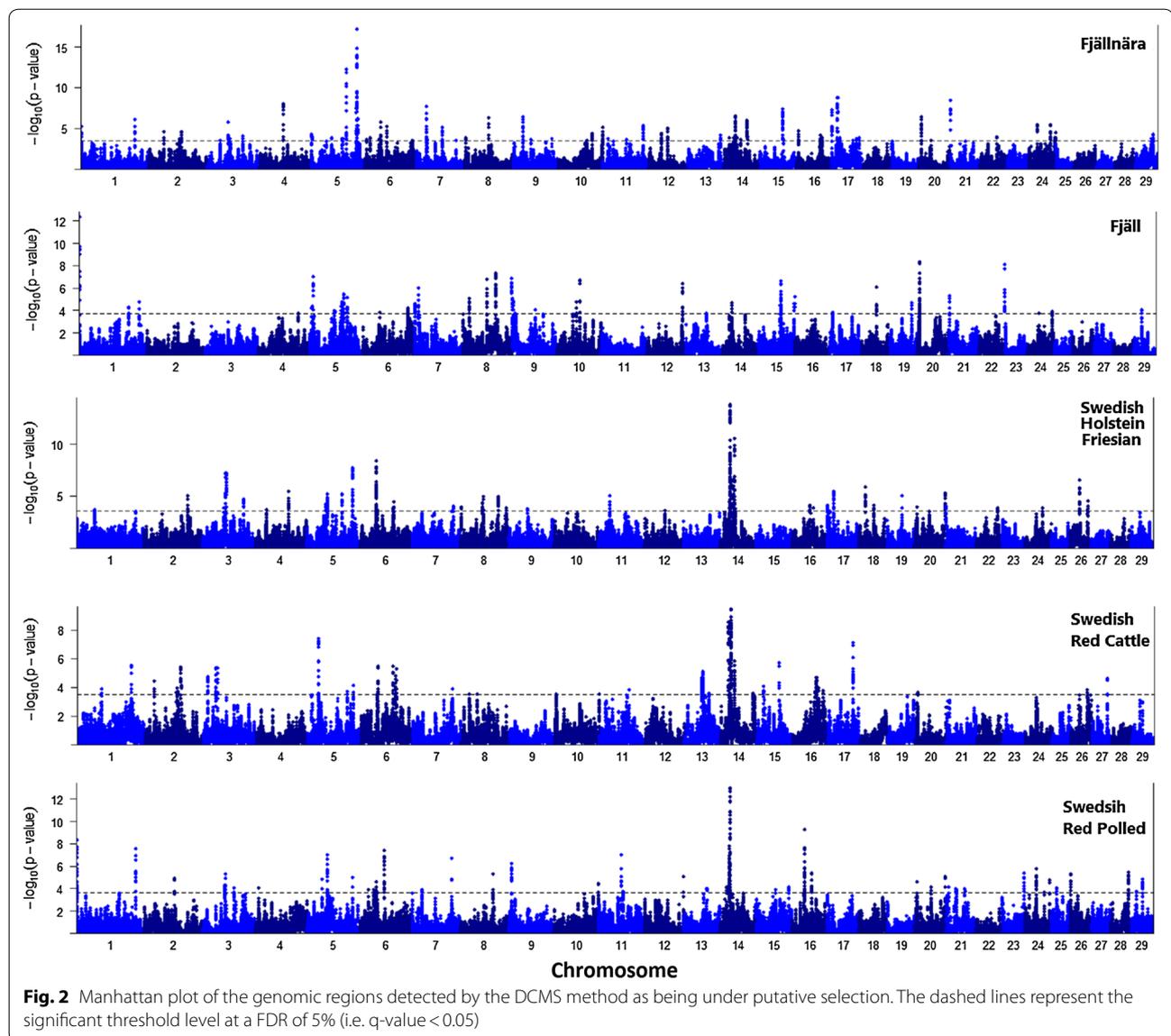
Discussion

Both the economic and fitness-related traits that were highlighted in the current study are quantitative traits, which are generally affected by many genes, most of which have small effects. Selection for complex traits may occur simultaneously across many loci (with less intensity), which would leave weak signals across the genome [35]. In spite of this, selection for polygenic traits with some alleles that have a large effect may leave detectable signals. In the current study, genomic regions that were

under putative selection in five native and commercial Swedish dairy cattle breeds from different parts of the country were studied. Selection signatures were identified by decomposition of p -values of five test statistics, instead of using single statistical tests separately or focusing on regions detected in more than one test. According to previous studies, composite measures of signatures of selection can provide an unbiased criterion to identify variants under selection more precisely [15, 36]. Therefore, by using decomposition of p -values to identify signatures of selection, candidate genes can be identified with higher power and greater precision for future research in medicine, agriculture, and livestock breeding.

The clear clustering of breeds into two groups (mountain breeds vs. middle–southern breeds) as shown in the PC plot presented here was not in complete accordance with the results of Upadhyay et al. [16] who could not clearly differentiate all the Swedish cattle breeds in separate clusters based on the first PC. This could be attributed to the fact that a larger number of Swedish breeds, i.e. representing an additional source of variation, were included in their study [16]. In addition, a recent study showed that the Swedish Fjäll cattle are closely related to the Northern, Western and Eastern Finn cattle and Icelandic cattle [37]. However, the map of signatures of selection reported here is in line with previous studies [16, 38, 39] that showed that Swedish Red Polled and Fjäll are two distinct breeds sharing no close genetic relationship with each other.

The historical background and usage differ between Swedish cattle breeds. Sweden spans about 1572 km from north to south [40] and climatic conditions for dairy production range from the subarctic cold climate in the more mountainous north, to the continental climate in the more densely populated southern part [41]. According to our results, none of the detected regions with identified genes was significant in all the five breeds analyzed. However, significant genomic regions harboring genes were associated with the polled phenotype, body size



and stature, resistance to gastrointestinal nematodes, response to hypoxia, growth and feed intake, feed efficiency, and reproduction overlapped between some of the breeds [see Additional file 3: Table S3]. The limited overlap between the detected putative regions across these Swedish cattle breeds suggest that some breed-specific selection occurred in the past. Nevertheless, in our study of signatures of selection, we found several genomic regions and genes that are involved in economic traits, such as milk production, growth, feed efficiency, reproduction and cold acclimation and that can be used for mapping causal mutations, identifying candidate genes and making better selection decisions in future cattle genetic improvement programs.

Signatures of potential human-mediated selection

In the commercial breeds, which originate from southern Sweden, such as the Swedish Red and the Swedish Holstein–Friesian, production of large quantities of milk has been economically advantageous. The old Swedish Friesian cattle that was rather a dual purpose breed was gradually transformed into the Swedish Holstein cattle from the 1970s to the mid-1990s, when imported Holstein bulls were used to improve milk yield and udder conformation [42]. Thus, the Swedish Holstein–Friesian became a more distinct dairy breed type, with a tall and lean conformation. Interestingly in our study, genes that are known to be associated with these traits were identified for this breed.

Table 3 Genomic autosomal regions detected (through the DCMS analyses) as being under putative selection in Swedish native and commercial cattle breeds

Region (Mb)	Q-value	Breed	Candidate gene (rank)	Trait	Citation
1: 1.58-2.39	0.00274	FNC	<i>EVA1C</i> (1), <i>C1H21orf59</i> (2), <i>SYNJ1</i> (3), <i>PAXBP1</i> (4), <i>C1H21orf62</i> (5), <i>OLIG1</i> (6), <i>LOC526226</i> (7), <i>IFNAR2</i> (8)	Polled phenotype	[105]
1: 2.25-2.52	4.9E-08	SMC	<i>URB1</i> (1), <i>EVA1C</i> (2), <i>MRAP</i> (3), <i>MIS18A</i> (4)	Polled phenotype	[105]
1: 1.14-1.33	0.00289	SRP	<i>CRYZL1</i> (5), <i>DONSON</i> (4), <i>SON</i> (3), <i>GART</i> (2), <i>DNAJC28</i> (10)	Polled phenotype	[105]
1: 1.52-2.51	1.2E-05	SRP	<i>IL10RB</i> (12), <i>IFNAR2</i> (11), <i>LOC526226</i> (10), <i>OLIG1</i> (9), <i>C1H21orf62</i> (8), <i>PAXBP1</i> (7), <i>SYNJ1</i> (6), <i>C1H21orf59</i> (5), <i>EVA1C</i> (2), <i>URB1</i> (1), <i>NSBTAG0000023718</i> (3), <i>MIS18A</i> (4)	Polled phenotype	[105]
2: 71.46-71.66	0.00633	SRP	<i>TMEM37</i> (1)	Disease resistance	[94]
3: 32.06-32.34	0.00325	SRC	<i>CHIA</i> (2), <i>CH13L2</i> (4)	Bacterial infection	[96, 97]
3: 9.23-9.72	0.00933	SRC	<i>DCAF8</i> (1)	Altitude adaptation	[88]
3: 9.57-9.58	0.00933	SRC	<i>PEA15</i> (2)	RFI	[77]
3: 54.20-54.22	6.8E-05	SHF	<i>GBP6</i> (4)	Bacterial infection	[103]
3: 54.34-54.36	6.87E-05	SHF	<i>GBP2</i> (3)	RFI	[77]
3: 54.64-54.68	6.8E-05	SHF	<i>GBP4</i> (11)	Stature	[106]
3: 54.25-54.26	6.8E-05	SHF	<i>GBP5</i> (2)	RFI	[77]
4: 77.43-77.70	0.00202	SHF	<i>NPC1L1</i> (5), <i>NUDCD3</i> (6)	Milk yield	[107]
5: 105.75-106.52	2.60e-13	FNC	<i>KCNA6</i> (13), <i>NDUFA9</i> (11), <i>AKAP3</i> (10), <i>C5H12orf4</i> (7), <i>RAD51AP1</i> (6), <i>FGF6</i> (2), <i>TIGAR</i> (3), <i>CCND2</i> (4)	Body weight	[76, 77]
			<i>CCND2</i> (4)	Stature	[78]
5: 82.49-82.73	2.37e-09	FNC	<i>PPF1BP1</i> (2), <i>REP15</i> (3)	Bacterial infection	[99, 108]
5: 106.15-106.32	2.4E-05	SHF	<i>FGF6</i> (2), <i>TIGAR</i> (3), <i>CCND2</i> (4)	Body weight	[76]
			<i>CCND2</i> (4)	Stature	[78]
5: 9.30-9.56	0.00039	SMC	<i>PPP1R12A</i> (1)	Reproduction	[109]
			<i>PPP1R12A</i> (1)	Altitude adaptation	[85, 86]
5: 24.92-25.02	8.17E-05	SRC	<i>FGD6</i> (3)	RFI	[77]
6: 37.06-37.27	5.7E-06	SHF	<i>TIGD2</i> (2)	Bacterial infection	[110]
6: 37.64-37.72	0.01986	SHF	<i>PYURF</i> (1)	Bacterial infection	[110]
7: 8.40-8.48	0.00159	SMC	<i>CYP4F2</i> (1)	Disease resistance	[93, 94]
8: 76.23-76.78	0.00022	SMC	<i>AQP3</i> (5), <i>AQP7</i> (5)	Acclimation	[89]
11: 25.57-25.83	0.00433	SHF	<i>ZFP36L2</i> (1)	Reproduction	[111]
12: 83.19-83.37	0.00082	SMC	<i>SLC10A2</i> (1)	Bacterial infection	[95]
13: 40.38-42.28	0.00985	SRC	<i>FOXA2</i> (6)	Growth traits	[63]
13: 43.40-43.42	0.00474	SRC	<i>ASB13</i> (9)	RFI	[77]
13: 42.73-44.04	0.00474	SRC	<i>ACSS1</i> (7)	Milk content	[112, 113]
14: 53.71-54.81	0.00068	FNC	<i>CSMD3</i> (1)	Body length	[80]
14: 24.42-25.11	2.2E-10	SHF	<i>TMEM68</i> (2), <i>TGS1</i> (1), <i>LYN</i> (3)	Growth traits	[114]
14: 24.42-25.11	2.2E-10	SHF	<i>TGS1</i> (1), <i>LYN</i> (3), <i>RPS20</i> (4), <i>MOS</i> (5), <i>PLAG1</i> (6), <i>CHCHD7</i> (7)	Pleiotropic	[44, 48, 50, 56, 114]
14: 25.35-25.73	2.2E-10	SHF	<i>IMPAD1</i> (1)	Pleiotropic	[51, 64, 66]
14: 26.01-26.24	0.00031	SHF	<i>FAM110B</i> (1)	Pleiotropic	[64]
14: 24.00-24.89	2.8E-06	SRC	<i>XKR4</i> (1), <i>TMEM68</i> (2)	Growth, feed intake	[114]
14: 25.50-25.58	0.00624	SRC	<i>IMPAD1</i> (1)	Pleiotropic	[64, 115]
14: 33.02-33.41	0.00513	SRC	<i>CSPP1</i> (1)	Reproduction	[116]
14: 33.02-33.41	0.00513	SRC	<i>COPSS</i> (2)	Feed intake	[117]
14: 32.10-32.63	0.00137	SRC	<i>CRH</i> (1)	Growth	[60]
15: 53.10-54.25	5.6E-05	FNC	<i>FCHSD2</i> (1)	Milk fat content	[70]
15: 53.10-54.25	5.6E-05	FNC	<i>FCHSD2</i> (1), <i>ARHGEF17</i> (5), <i>RELT</i> (6)	Bacterial infection	[92]
			<i>UCP2</i> (13)	Hypoxia	[118]
			<i>UCP2</i> (13), <i>UCP3</i> (14)	Milk fat content, carcass quality	[71-73, 119]
15: 53.13-53.80	0.00073	SMC	<i>FCHSD2</i> (1)	Milk fat content	[70]
15: 53.13-53.80	0.00073	SMC	<i>FCHSD2</i> (1), <i>ARHGEF17</i> (4), <i>RELT</i> (5)	Bacterial infection	[92]

Table 3 (continued)

Region (Mb)	Q-value	Breed	Candidate gene (rank)	Trait	Citation
15: 53.52-53.67	0.00178	SRC	<i>ARHGEF17</i> (1), <i>RELT</i> (2)	Bacterial infection	[92]
15: 78.33-78.51	0.02265	SRP	<i>SLC39A13</i> (1)	Stature	[120, 121]
16: 55.09-55.15	0.00963	SRC	<i>PRDM2</i> (1)	Bacterial infection	[77, 99]
			<i>PRDM2</i> (1), <i>KDM5B</i> (2)	Bacterial infection	[98, 99]
			<i>PRDM2</i> (1)	RFI	[77]
16: 44.96-45.25	0.00299	SRP	<i>SPSB1</i> (1)	Immune regulation	[100, 101]
17: 57.91-58.43	0.00011	SRC	<i>HSPB8</i> (1)	Heat and cold stress	[122, 123]
19: 51.21-51.30	0.01373	SMC	<i>SLC16A3</i> (1)	Altitude adaptation	[88]
19: 35.10-35.14	0.00442	SHF	<i>GID4</i> (1), <i>ATPAF2</i> (2)	Milk protein	[124]
20: 71.49-71.80	0.00272	SHF	<i>CEP72</i> (3), <i>SLC9A3</i> (2)	Feed efficiency	[125]
20: 71.65-71.80	0.00498	SRP	<i>SLC9A3</i> (2), <i>EXOC3</i> (3)	Feed efficiency	[125]
24: 49.90-49.94	0.00205	FNC	<i>ACAA2</i> (1)	Milk content	[126]
				RFI	[72]
26: 22.21-23.02	0.00021	SHF	<i>ELOVL3</i> (1) <i>NFKB2</i> (8)	Milk content	[127]
				Milk content	[128]

q-value: q-value of the most significant SNP within the significant genomic region

Breed: breed names are shown in abbreviated form (full names are in Table 1)

Candidate genes: candidate genes within significant genomic regions

RFI Residual feed intake

The full list of genes is in Additional file 2

The most significant peak identified by DCMS was located on BTA14 in the Swedish Holstein–Friesian breed (BTA14:24.42-25.11; Mb; q-value (most significant SNP) = $2.2E - 10$). Previous studies have reported that the regions around this location contain numerous genes and QTL that affect cattle stature and related traits [43–47]. The genes that we identified within this region, including *PLAG1*, *CHCHD7*, *MOS*, *RPS20*, and *LYN*, are known as major genes with a role in both human height and cattle stature [44, 46, 48–52]. This region (i.e. BTA14:24.4-25.4) is also known to affect a variety of reproduction-related traits such as maturity index, scrotal circumference, age at puberty in Brahman cattle [53, 54], and blood levels of IGF1 in Brahman and Tropical Composite cattle [55]. The *PLAG1* gene is significantly correlated with the lactation phenotype [56], which suggests that this region has a major pleiotropic effect in various cattle breeds. Other genes were also identified such as *GBP4*, *FGF6*, *TIGAR*, *CCND2* (involved in body weight and stature), *GID4*, *ATPAF2*, *ELOVL3*, *NFKB2* (involved in milk yield and components), *TMEM68*, *TGS*, *LYN*, *CEP72*, *SLC9A3* (involved in growth and feed efficiency) (see Table 3).

The modern Swedish Red Cattle is a dairy breed, but historically it was important for beef production in Sweden, and still shows today a better average carcass conformation classification and a slightly higher carcass fat content than Swedish Holstein [57]. The Swedish Red

Cattle breed is also known for producing higher concentrations of milk fat and protein than Swedish Holstein [58]. Our results also highlight putative signatures of selection in regions on BTA13 and 14 related to carcass quality and milk composition in Swedish Red Cattle. One of the interesting genes under putative selection is *CRH* (top ranked gene in the DCMS analysis), which is consistent with the breeding history of this breed. *CRH*, which is located at 31.49 Mb on BTA14, plays an important role in several physiological and biological pathways regarding the stimulation of ACTH (adrenocorticotropin) secretion, that up-regulates cortisol [59]. The cortisol hormone stimulates gluconeogenesis in liver and lipolysis in adipose tissue, and inhibits glucose consumption in adipose tissue and muscle [59]. Polymorphisms in the *CRH* gene are associated with marbling score [59], growth and carcass yield [60], and milk production [61, 62], which suggests a pleiotropic effect of this gene. *FOXA2* is another identified gene that is known to be one of the important transcriptional activators and plays a role in the regulation of energy homeostasis and feeding. *FOXA2* is significantly associated with chest girth, body weight, and growth traits in Chinese cattle [63]. Other interesting genes were identified in Swedish Red Cattle such as *PLAG1*, *XKR4*, and *IMPAD1*, which are positional candidate genes for pleiotropic QTL for growth traits [64]. According to the literature, *PLAG1* is associated with carcass weight, stature, body weight and milk

yield [45, 46, 65–68], *XKR4* with birth weight, growth, and feed intake [51, 67, 69], and *IMPAD1* with carcass weight, stature and body weight [56, 65].

In native breeds from remote mountain areas in northern Sweden, which are located far from commercial dairies, the production of milk suitable for storable dairy products such as cheese was most valuable in the past. A putative signature of selection was identified in breeds from the northern region (i.e. Fjäll and Fjällnära) containing the first ranked gene, i.e. *FCHSD2* and the gene *AQP7* in the Fjäll breed. A meta-analysis on French dairy cattle breeds reported *AQP7* and *FCHSD2* as candidate genes within QTL regions on BTA8 and 15, respectively, associated with milk fat percentage [70]. We also identified two other genes (*UCP2* and *UCP3*) in the Fjällnära breed, for which significant associations were reported between a polymorphism in *UCP2* and calving interval in dairy cattle, and between polymorphisms in *UCP3* and production and fat content in dairy cattle [71] and carcass phenotype in beef cattle [72, 73]. Gene expression analyses also revealed an association between *UCP2* and residual feed intake in Nellore cattle [74].

Body size and cold acclimation

Compared to the commercial southern breeds, and in addition to selection pressures for milk yield and composition, historically the native breeds had to cope with cold weather, in some high mountain regions, a limited amount of feed, and even starvation during winter [75]. These breeds generally have a smaller body size (see Table 1). The DCMS analyses identified potential evidence of adaptation to the cold climate of the mountain regions characterized by a very limited amount of feed supply (see Table 1). We identified several genes including *CCND2*, *FGF6*, *TIGAR*, *KCNA6*, *NDUFA9*, *AKAP3*, *CSH12orf4*, and *RAD51AP1* on BTA5 (BTA5: 105.75–106.52) that have been under strong selection (q -value $2.6E-13$) in Fjällnära cattle. Interestingly, a genome-wide association study (GWAS) on cattle populations from Siberia showed associations between these genes and body measurements [76]. In a gene expression study of bovine QTL, *TIGAR* was found to be significantly differentially expressed and associated with body weight [77]. Moreover, a meta-analysis of GWAS studies for cattle stature reported *CCND2* as the second most significant gene regulating stature in mammals [78]. Another gene, *CSMD3*, which encodes a transmembrane protein [79] was reported to be associated with body measurements in cattle [80].

Body size and shape vary among breeds [2, 81]. Whereas most bovine breeds from temperate regions with access to a good supply of high-quality nutrients have a larger body size, breeds that live in high altitude

areas with a limited amount of feed often have a smaller body size. Our findings also agree with the Geist theory [82], according to which food availability per animal during the growing season is a determining factor for body size evolution in mammals. We hypothesized that small body size was subject to natural selection in Swedish northern breeds since it probably allowed them to develop hardiness and endurance characteristics necessary for searching food and water over long distances. This is supported by a previous study that compared the grazing pattern of a native cattle (Fjäll) with that of a commercial breed (Swedish Holstein–Friesian) and showed that the Fjäll breed explored across extensive regions with an inclination towards various types of plants [83]. In addition, shortage of stored feed during winters in years with bad harvests or extra cold and long winters occurred before the 20th century. Such historical scarcity of feed, especially towards the end of long winters when animals were kept indoors, probably favored animals with a small body weight because this is associated with lower maintenance requirements [84]. As shown in Table 1, Fjäll cattle have a small body size, which is only a bit larger than that of the Fjällnära breed. However, we did not identify genes associated with body measurements in the Fjäll breed, possibly because they have been under stronger artificial selection for milk production and kept in areas of northern Sweden with less harsh conditions.

We also identified genes under selection in the northern breeds, including *PPP1R12A*, *AQP3*, *AQP7* and *SLC16A3* in the Fjäll breed, and *UCP2*, *UCP3*, *ACAA2* and *TIGAR* in the Fjällnära breed, which are known to be related with altitude adaptation and response to hypoxia. Hypoxia or hypoxic stress is described as a decline in oxygen levels below the normal levels of 20.8 to 20.95% and results in metabolic adaptation at both the cellular and organism levels. Phosphorylation of the *PPP1R12A* (also called *MYPT1*) gene is known to increase as the level of oxygen decreases [85, 86]. In a study comparing Ethiopian sheep breeds that are adapted to different ecological regions, Edea et al. [87] identified four genes including *PPP1R12A* that showed signatures of selection likely related to high altitude adaptation. Gene expression analyses revealed *SLC16A3* as one of the significant up-regulated genes (p -value = $1.77E-02$) in response to high altitude adaptation of sheep fetal carotid arteries [88]. For the other northern breed, Fjällnära, we found that the four genes under putative selection, *UCP2*, *UCP3*, *ACAA2* and *TIGAR*, are involved in the response to hypoxia (GO:0001666) and other related Gene Ontologies. The genes identified by the DCMS analysis suggest that the Fjällnära and Fjäll breeds are probably adapted to the high altitude mountains. However, it should be noted

that the highest mountain reaches ~2000 m in Sweden, and less than 2500 m in Scandinavia (Norway), and that these breeds have been maintained at lower altitudes. For example, the remaining Fjällnära cattle were sampled on farms located at 485 to 682 m above average sea level. Thus, their adaptation to high altitudes may have occurred before their recorded history in Scandinavia. A functional gene network analysis, *AQP3* and *AQP7* (and *AQP5*, another gene of the aquaporin family, which we did not identify here but was reported in a study on Russian cattle breeds by Yurchenko et al. [20]) were also identified as candidate genes for the regulation of thermal adaptation via the transport of water, glycerol, and small solutes across cell membranes [89].

Signatures of selection for disease resistance genes

Until the end of the 19th century, the northern cattle grazed on woodlands in the summertime and had to be guarded from predators, which is why they are assumed to have developed hardiness and robustness to cope with their local environment [90, 91]. In concordance with the historical background of these breeds, our results suggest selection for stress response and immunity-related traits. Our results revealed an overlapping putative region under selection including the *FCHSD2*, *P2RY2*, *P2RY6*, *ARHGEF17* and *RELT* genes in the Fjällnära and Fjäll breeds. Moreover, the *P2RY6*, *ARHGEF17* and *RELT* genes were also identified in Swedish Red Cattle [see Additional file 3 Table S3]. *FCHSD2*, *RELT* and *ARHGEF17* have already been identified in sheep as candidate genes that play a role in the serum levels of immunoglobulin A (IgA), which is an indicator of resistance to gastrointestinal nematodes [92]. In addition, two other genes, *CYP4F2* and *SLC10A2*, were identified on BTA7 and 12 in Fjäll Cattle. In cattle, *CYP4F2* was reported to be involved in the host resistance to the intestinal worm *Cooperia oncophora* [93] and to gastrointestinal nematodes [94]. Hempel et al. [95] showed that *SLC10A2* was among the top five most significantly up-regulated genes (q-value = 0.000897) in a comparison between infected vs. uninfected cows with *Mycobacterium avium* subsp. *paratuberculosis* (an intracellular bacterium). Paratuberculosis infections are very rare in Sweden, but this gene could be involved in the response to infections by other bacteria or selected for other reasons related to its role in the uptake of intestinal bile acids. Our findings suggest a possible selection for adaptive immunity genes in these Swedish indigenous breeds. These genes explain a small proportion of the variance [92], which suggests that resistance/susceptibility to gastrointestinal nematodes is a complex trait that is determined by multiple groups

of gene networks rather than the effect of an individual gene. However, since the *P2RY6*, *ARHGEF27*, *RELT*, *FCHSD2*, *CYP4F2* and *SLC10A2* genes appear to be under selection in the Swedish indigenous breeds, it would be interesting to analyze them in more detail. Moreover, future studies should investigate whether phenotypic variation in susceptibility to gastrointestinal nematodes or other important traits exists between these breeds.

The four *CHIA*, *CHI3L2*, *PRDM2* and *KDM5B* genes that we identified on BTA4 and 16 in the Swedish Red Cattle were previously reported as disease resistance genes (i.e. to gastrointestinal nematodes or bacterial infection) in independent studies on sheep and cattle breeds [96–99]. In Swedish Red Polled, we found that the *TMEM37* gene (BTA2: 71.46-71.66), known to be involved in the resistance to gastrointestinal nematodes, ranked as the closest gene to the most significant SNP showing a signature of selection (q-value = 0.0063). Li et al. [94] reported a set of 64 candidate genes that displayed significant overexpression at a high 5% FDR threshold level for resistance to gastrointestinal nematodes and included *TMEM37*. Similarly, we detected *SPSB1* (BTA16: 44.96-45.25), which was previously reported to be under selection in a meta-assembly of studies on signatures of selection in various cattle breeds [100], and in another analysis of bovine signatures of selection [101]. *SPSB1* (along with *SPSB2*) regulates the amount of nitric oxide (NO) produced via the induction of NO synthase [102]. Nitric oxide is one of the main contributors of reactive nitrogen known to have a major role in host defense by killing intracellular pathogens [102]. In Swedish Holstein–Friesian, the *GBP6* gene identified on BTA3 (BTA3: 53.73-55.25), is part of the six top genes that are known as key players in immune response to *Mycobacterium avium* subsp. *Paratuberculosis* [103].

A relatively small number of genes (n = 61) was identified in significant genomic regions of the Swedish Red Polled breed (Table 1). More importantly, by reviewing the literature, we found only a small number of genes in the detected selection signature region in this breed that were associated with acclimation or economic traits (Table 3). This could be attributed to the rapid reduction of N_e of this breed (genetic bottleneck) with only 20 cows still present in 1979 [104], which suggests that strong genetic drift may have occurred in this population, although the genetic variation in this breed increased in the last decades by the importation of red polled cattle from Norway and Finland. Schlamp et al. [8] reported that the identification of signatures of selection in small populations can be problematic, since they can be

weakened by the strong genetic drift that occurs in such small populations.

Conclusions

In total, 108 genomic regions (including 355 unique protein-coding genes) that display signatures of selection in Swedish indigenous (northern breeds) and commercial (southern origin) cattle breeds were identified by incorporating *p*-values of different statistics in a single DCMS framework. These signatures of selection are in line with the history of these breeds. Some of these signatures of selection are located in regions carrying genes associated with economic traits such as milk yield and components, and growth and carcass traits, thus describing response to human-mediated selection in commercial breeds. Other signatures of selection that were detected in regions that harbor genes for body size may be involved in the adaptation of these native breeds to the specific climate of the northern part of the country, which is characterized by cold weather and, historically, by a limited amount of feed during the winter. For example, four genes that are involved in body size traits were subject to selection in both Fjällnära (a native mountain breed with a small body size) and Swedish Holstein–Friesian (a commercial breed with a larger body size). Indeed, our results may provide background knowledge to better understand the genetic mechanisms that are involved in economic traits and adaptation to local climate. Additional high-resolution studies focusing on the putative regions of signatures of selection and using larger sample sizes and recorded phenotypes may help pinpoint the candidate genes under selection.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12711-020-00571-5>.

Additional file 1: Table S1. List of genomic regions under putative selection (*q*-value < 0.05) identified using the DCMS (De-correlated composite of multiple signals) method for each breed. The first column represents the abbreviated name for each breed, and the second column shows the number (ID) of each genomic region identified as under putative selection within each breed.

Additional file 2: Table S2. Full list of protein coding genes identified (using UMD3.1 genome assembly) from the genomic regions under putative selection. The 1st column represents the abbreviated name for each breed, and the 6th column represents the number (ID) of significant genomic region under putative selection (within each breed) and the rank of each gene within that region (i.e. the value 1 represents the closest gene to the most significant SNP within each region).

Additional file 3: Table S3. Mutual genes identified among breeds.

Acknowledgements

We gratefully acknowledge contribution of samples and information about breeds and animals from the different breed organizations, individual animal owners, VikingGenetics, and the sample collections at the Biobank and Dept. of Animal Breeding and Genetics at the Swedish University of Agricultural Sciences.

Authors' contributions

SMG, SE and AMJ conceived and designed the study, and SMG conducted the analyses and wrote the first draft. NP and PD helped with data analyses. NP and SMA helped with writing the draft. SMG, SE, AMJ, MK, AJ, and SMA revised the manuscript. All authors read and approved the final manuscript.

Funding

Open access funding provided by the Swedish University of Agricultural Sciences. Financial support to SE and AMJ from the Swedish Research Foundation (FORMAS, DNR 942-2015-1015) is acknowledged.

Availability of data and materials

The genotyping data of Swedish cattle breeds used for conducting this research were already deposited into the DRYAD public data repository: <https://datadryad.org/stash/dataset/doi:10.5061/dryad.wdbrv15j4>.

Ethics approval and consent to participate

The samples had already been used for analyses in previous studies, and therefore, no ethical approval was needed.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj 31587-11167, Iran. ² Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden. ³ AgriBio Centre for AgriBioscience, Agriculture Victoria, Bundoora, VIC 3083, Australia. ⁴ Department of Animal Science, Faculty of Agriculture and Natural Resources, Islamic Azad University, Tabriz Branch, Tabriz, Iran. ⁵ Department of Animal Science and Aquaculture, Dalhousie University, Truro, NS B2N5E3, Canada. ⁶ Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.

Received: 23 April 2020 Accepted: 21 August 2020

Published online: 04 September 2020

References

- Darwin C, Beer G. The origin of species. 1st ed. Oxford: Oxford University Press; 1859.
- Ghoreishifar SM, Moradi-Shahrababak H, Fallahi MH, Jalil Sarghale A, Moradi-Shahrababak M, Abdollahi-Arpanahi R, et al. Genomic measures of inbreeding coefficients and genome-wide scan for runs of homozygosity islands in Iranian river buffalo *Bubalus Bubalis*. BMC Genet. 2020;21:16.
- Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38:1358–70.
- Nei M, Li W-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA. 1979;76:5269–73.
- Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. PLoS Genet. 2015;11:e1005004.
- Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–95.
- Lotterhos KE, Card DC, Schaal SM, Wang L, Collins C, Verity B. Composite measures of selection can improve the signal-to-noise ratio in genome scans. Methods Ecol Evol. 2017;8:717–27.
- Schlamp F, van der Made J, Stambler R, Chesebrough L, Boyko AR, Messer PW. Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. Mol Ecol. 2016;25:342–56.
- Lotterhos KE, Whitlock MC. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol Ecol. 2015;24:1031–46.

10. Vatsiou AI, Bazin E, Gaggiotti OE. Detection of selective sweeps in structured populations: a comparison of recent methods. *Mol Ecol*. 2016;25:89–103.
11. Salek Ardestani S, Aminafshar M, Zandi Baghche Maryam MB, Banabazi MH, Sargolzaei M, Miar Y. Signatures of selection analysis using whole-genome sequence data reveals novel candidate genes for pony and light horse types. *Genome*. 2020;63:387–96.
12. Salek Ardestani S, Aminafshar M, Zandi Baghche Maryam MB, Banabazi MH, Sargolzaei M, Miar Y. Whole-genome signatures of selection in sport horses revealed selection footprints related to musculoskeletal system development processes. *Animals*. 2020;10:53.
13. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genet*. 2014;15:34.
14. Utsunomiya YT, O'Brien AMP, Sonstegard TS, Van Tassell CP, do Carmo AS, Mészáros G, et al. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. *PLoS One*. 2013;8:e64280.
15. Ma Y, Ding X, Qanbari S, Weigend S, Zhang Q, Simianer H. Properties of different selection signature statistics and a new strategy for combining them. *Heredity*. 2015;115:426–36.
16. Upadhyay M, Eriksson S, Mikko S, Strandberg E, Stålhammar H, Groenen MA, et al. Genomic relatedness and diversity of Swedish native cattle breeds. *Genet Sel Evol*. 2019;51:56.
17. Johansson AM, Upadhyay M, Strandberg E, Eriksson S. Genetic differentiation between subpopulations of Swedish mountain (Fjäll and Fjällnära) cattle. *Acta Agr Scand A-An*. 2019;69:39–46.
18. Korkman N. Svensk nötkreatursavel under ett sekel. *K Skogs- o Lantrbr. akad Tidskr*. 1988;20:225–38.
19. Sveriges Rödkulleförening. <https://rodkullan.se/R%3b6dkullan/Historik.html>. Accessed 20 March 2020.
20. Yurchenko AA, Daetwyler HD, Yudin N, Schnabel RD, Vander Jagt CJ, Soloshenko V, et al. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci Rep*. 2018;8:12984.
21. Yurchenko AA, Deniskova TE, Yudin NS, Dotsev AV, Khamiruev TN, Selionova MI, et al. High-density genotyping reveals signatures of selection related to acclimation and economically important traits in 15 local sheep breeds from Russia. *BMC Genomics*. 2019;20:294.
22. Gutiérrez-Gil B, Arranz JJ, Wiener P. An interpretive review of selective sweep studies in *Bos taurus* cattle populations: identification of unique and shared selection signals across breeds. *Front Genet*. 2015;6:167.
23. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.
24. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow *Bos taurus*. *Genome Biol*. 2009;10:R42.
25. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10:5–6.
26. Barbato M, Orozco-terWengel P, Tapio M, Bruford MW. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front Genet*. 2015;6:109.
27. Ghoreishifar SM, Moradi-Shahrbabak H, Parna N, Davoudi P, Khansefid M. Linkage disequilibrium and within-breed genetic diversity in Iranian Zandi sheep. *Arch Anim Breed*. 2019;62:143–51.
28. Ma L, O'Connell JR, VanRaden PM, Shen B, Padhi A, Sun C, et al. Cattle sex-specific recombination and genetic control from a large pedigree analysis. *PLoS Genet*. 2015;11:e1005387.
29. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156–8.
30. Verity R, Collins C, Card DC, Schaal SM, Wang L, Lotterhos KE. Minotaur: a platform for the analysis and visualization of multivariate results from genome scans with R Shiny. *Mol Ecol Resour*. 2017;17:33–43.
31. Todorov V, Tempel M, Filzmoser P. Detection of multivariate outliers in business survey data with incomplete information. *Adv Data Anal Classif*. 2011;5:37–56.
32. Venables WN, Ripley BD. *Modern applied statistics with S*. 4th ed. New York: Springer; 2002.
33. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci USA*. 2003;100:9440–5.
34. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57:289–300.
35. Kemper KE, Saxton SJ, Bolormaa S, Hayes BJ, Goddard ME. Selection for complex traits leaves little or no classic signatures of selection. *BMC Genomics*. 2014;15:246.
36. Grossman SR, Shylakhter I, Karlsson EK, Byrne EH, Morales S, Frieden G, et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*. 2010;327:883–6.
37. Gautason E, Schönherz AA, Sahana G, Guldbbrandtsen B. Relationship of Icelandic cattle with Northern and Western European cattle breeds, admixture and population structure. *Acta Agr Scand A-An*. 2019;69:25–38.
38. Kantanen J, Olsaker I, Holm L-E, Lien S, Vilkki J, Brusgaard K, et al. Genetic diversity and population structure of 20 North European cattle breeds. *J Hered*. 2000;91:446–57.
39. Tapio I, Värvi S, Bennewitz J, Maleviciute J, Finland E, Grislis Z, et al. Prioritization for conservation of northern European cattle breeds based on analysis of microsatellite data. *Conserv Biol*. 2006;20:1768–79.
40. Agriculture SBo. Facts about Swedish Agriculture. 18 pp. 2009. https://www2.jordbruksverket.se/webdav/files/SJV/trycksaker/Pdf_ovrigt/ovr2gb.pdf. Accessed 18 August 2020.
41. European climate zones and bioclimatic design requirements. 2016. <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ac7b5027&appId=PPGMS>. Accessed 20 March 2020.
42. Dahlberg C. Bevarande av svensk låglandsboskap. Second cycle, AZE. Uppsala: SLU, Dept. of Animal Breeding and Genetics. 2010. <https://stud.epsilon.slu.se/1865/>. Accessed 18 August 2020.
43. McClure MC, Morsci NS, Schnabel RD, Kim JW, Yao P, Rolf MM, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim Genet*. 2010;41:597–607.
44. Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, et al. Genome-wide association study identified three major QTL for carcass weight including the *PLAG1-CHCHD7* QTN for stature in Japanese Black cattle. *BMC Genet*. 2012;13:40.
45. Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Götz K-U, et al. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. *Genetics*. 2011;187:289–97.
46. Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polymorphic regions affecting human height also control stature in cattle. *Genetics*. 2011;187:981–4.
47. Spelman RJ, Huisman AE, Singireddy SR, Coppieters W, Arranz J, Georges M, et al. Quantitative trait loci analysis on 17 nonproduction traits in the New Zealand dairy population. *J Dairy Sci*. 1999;82:2514–6.
48. Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, et al. Many sequence variants affecting diversity of adult human height. *Nat Genet*. 2008;40:609–15.
49. Karim L, Takeda H, Lin L, Druet T, Arias JA, Baurain D, et al. Variants modulating the expression of a chromosome domain encompassing *PLAG1* influence bovine stature. *Nat Genet*. 2011;43:405–13.
50. Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet*. 2008;40:584–91.
51. Utsunomiya YT, Do Carmo AS, Carvalheiro R, Neves HH, Matos MC, Zavarez LB, et al. Genome-wide association study for birth weight in Nelore cattle points to previously described orthologous genes affecting human and bovine height. *BMC Genet*. 2013;14:52.
52. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet*. 2008;40:575–83.
53. Soares ACC, Guimarães SEF, Kelly MJ, Fortes MRS, Silva FFE, Verardo LL, et al. Multiple-trait genomewide mapping and gene network analysis for scrotal circumference growth curves in Brahman cattle. *J Anim Sci*. 2017;95:3331–45.

54. Fortes MRS, Reverter A, Hawken RJ, Bolormaa S, Lehnert SA. Candidate genes associated with testicular development, sperm quality, and hormone levels of inhibin, luteinizing hormone, and insulin-like growth factor 1 in Brahman bulls. *Biol Reprod*. 2012;87:58.
55. Fortes MRS, Li Y, Collis E, Zhang Y, Hawken RJ. The IGF 1 pathway genes and their association with age of puberty in cattle. *Anim Genet*. 2013;44:91–5.
56. Fink T, Tiplady K, Lopdell T, Johnson T, Snell RG, Spelman RJ, et al. Functional confirmation of *PLAG1* as the candidate causative gene underlying major pleiotropic effects on body weight and milk characteristics. *Sci Rep*. 2017;7:44793.
57. Eriksson S, Ask-Gullstrand P, Fikse WF, Jonsson E, Eriksson J-Å, Ståhlhammar H, et al. Different beef breed sires used for crossbreeding with Swedish dairy cows—effects on calving performance and carcass traits. *Livest Sci*. 2020;232:103902.
58. Wedholm A, Hallén E, Bach Larsen L, Lindmark-Månsson H, Hans Karlsson A, Allmere T. Comparison of milk protein composition in a Swedish and a Danish dairy herd using reversed phase HPLC. *Acta Agr Scand A-An*. 2006;56:8–15.
59. Wibowo TA, Michal JJ, Jiang Z. Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. *Genome*. 2007;50:939–45.
60. Buchanan FC, Thue TD, Yu P, Winkelman-Sim DC. Single nucleotide polymorphisms in the corticotrophin-releasing hormone and pro-opiomelanocortin genes are associated with growth and carcass yield in beef cattle. *Anim Genet*. 2005;36:127–31.
61. Fontanesi L, Scotti E, Samoré AB, Bagnato A, Russo V. Association of 20 candidate gene markers with milk production and composition traits in sires of Reggiana breed, a local dairy cattle population. *Livest Sci*. 2015;176:14–21.
62. Kulig H, Kowalewska-Luczak I, Szydłowska J. *CRH* gene polymorphism in relation to milk production traits in cattle. *Acta Sci Pol Zootech*. 2011;10:49–54.
63. Liu M, Li M, Wang S, Xu Y, Lan X, Li Z, et al. Association analysis of bovine *Foxa2* gene single sequence variant and haplotype combinations with growth traits in Chinese cattle. *Gene*. 2014;536:385–92.
64. Smith JL, Wilson ML, Nilson SM, Rowan TN, Oldeschulte DL, Schnabel RD, et al. Genome-wide association and genotype by environment interactions for growth traits in US Gelbvieh cattle. *BMC Genomics*. 2019;20:926.
65. Lee SH, Choi BH, Lim D, Gondro C, Cho YM, Dang CG, et al. Genome-wide association study identifies major loci for carcass weight on BTA14 in Hanwoo (Korean cattle). *PLoS One*. 2013;8:e74677.
66. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics*. 2014;15:442.
67. Seabury CM, Oldeschulte DL, Saatchi M, Beaver JE, Decker JE, Halley YA, et al. Genome-wide association study for feed efficiency and growth traits in US beef cattle. *BMC Genomics*. 2017;18:386.
68. Utsunomiya YT, Milanese M, Utsunomiya ATH, Torrecilha RBP, Kim E-S, Costa MS, et al. A *PLAG1* mutation contributed to stature recovery in modern cattle. *Sci Rep*. 2017;7:1C–15.
69. Lindholm-Perry AK, Kuehn LA, Smith TPL, Ferrell CL, Jenkins TG, Freetly HC, et al. A region on BTA14 that includes the positional candidate genes *LYPLA1*, *XKR4* and *TMEM68* is associated with feed intake and growth phenotypes in cattle. *Anim Genet*. 2012;43:216–9.
70. Marete AG, Guldbrandtsen B, Lund MS, Fritz S, Sahana G, Boichard D. A meta-analysis including pre-selected sequence variants associated with seven traits in three French dairy cattle populations. *Front Genet*. 2018;9:522.
71. Kowalewska-Luczak I, Głosińska J, Czerniawska-Piątkowska E. Effect of *UCP2* and *UCP3* genes polymorphisms on functional traits in Dairy cattle. *Russ J Genet*. 2018;54:853–7.
72. Chung E-R, Shin S-C, Heo J-P. Association between SNP marker of uncoupling protein 3 gene and meat yield and marbling score traits in Korean cattle. *Korean J Food Sci An*. 2011;31:530–6.
73. Ryu J, Kim Y, Kim C, Kim J, Lee C. Association of bovine carcass phenotypes with genes in an adaptive thermogenesis pathway. *Mol Biol Rep*. 2012;39:1441–5.
74. Fonseca LFS, Gimenez DFJ, Mercadante MEZ, Bonilha SFM, Ferro JA, Baldi F, et al. Expression of genes related to mitochondrial function in Nellore cattle divergently ranked on residual feed intake. *Mol Biol Rep*. 2015;42:559–65.
75. Nilsson R. *Fjällkon: historik, avel och framtid*. 1st ed. Övre Svartlå: Robert Nilsson; 2007.
76. Igoshin AV, Yudin NS, Belonogova NM, Larkin DM. Genome-wide association study for body weight in cattle populations from Siberia. *Anim Genet*. 2019;50:250–3.
77. Khansefid M, Pryce JE, Bolormaa S, Chen Y, Millen CA, Chamberlain AJ, et al. Comparing allele specific expression and local expression quantitative trait loci and the influence of gene expression on complex trait variation in cattle. *BMC Genomics*. 2018;19:793.
78. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. *Nat Genet*. 2018;50:362–7.
79. Shimizu A, Asakawa S, Sasaki T, Yamazaki S, Yamagata H, Kudoh J, et al. A novel giant gene CSMD3 encoding a protein with CUB and sushi multiple domains: a candidate gene for benign adult familial myoclonic epilepsy on human chromosome 8q23.3–q24. *Biochem Biophys Res Commun*. 2003;309:143–54.
80. An B, Xia J, Chang T, Wang X, Xu L, Zhang L, et al. Genome-wide association study reveals candidate genes associated with body measurement traits in Chinese Wagyu beef cattle. *Anim Genet*. 2019;50:386–90.
81. Kim ES, Elbeltagy AR, Aboul-Naga AM, Rischkowsky B, Sayre B, Mwacharo JM, et al. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity*. 2016;116:255–64.
82. Geist V. Bergmann's rule is invalid. *Can J Zool*. 1987;65:1035–8.
83. Hessele A, Dahlström F, Bele B, Norderhaug A, Söderström M. Effects of breed on foraging sites and diets in dairy cows on mountain pasture. *Int J Biodivers Sci Ecosyst Serv Manag*. 2014;10:334–42.
84. VandeHaar MJ, Armentano LE, Weigel K, Spurlock DM, Tempelman RJ, Veerkamp R. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. *J Dairy Sci*. 2016;99:4941–54.
85. Dada LA, Novoa E, Lecuona E, Sun H, Sznajder JI. Role of the small GTPase RhoA in the hypoxia-induced decrease of plasma membrane Na⁺-K-ATPase in A549 cells. *J Cell Sci*. 2007;120:2214–22.
86. Zeng Y, Zhang X, Kang K, Chen J, Wu Z, Huang J, et al. MicroRNA-223 attenuates hypoxia-induced vascular remodeling by targeting RhoB/MLC2 in pulmonary arterial smooth muscle cells. *Sci Rep*. 2016;6:24900.
87. Edea Z, Dadi H, Dessie T, Kim K-S. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes Genomics*. 2019;41:973–81.
88. Goyal R, Van Wickle J, Goyal D, Matei N, Longo LD. Antenatal maternal long-term hypoxia: acclimatization responses with altered gene expression in ovine fetal carotid arteries. *PLoS One*. 2013;8:e82200.
89. Valero KCW, Pathak R, Prajapati I, Bankston S, Thompson A, Usher J, et al. A candidate multimodal functional genetic network for thermal adaptation. *PeerJ*. 2014;2:e578.
90. Människan och mjölkdjuren. <https://www.isof.se/matkult/mjolk/manniskan-och-mjolkdjuren.html>. Accessed 20 March 2020.
91. Svensk Fjällrasavel. <http://www.fjallko.se/fjallrasen/fjallrasen>. Accessed 20 March 2020.
92. Atlija M, Arranz J-J, Martínez-Valladares M, Gutiérrez-Gil B. Detection and replication of QTL underlying resistance to gastrointestinal nematodes in adult sheep using the ovine 50 K SNP array. *Genet Sel Evol*. 2016;48:4.
93. Li RW, Wu S, Li C-J, Li W, Schroeder SG. Splice variants and regulatory networks associated with host resistance to the intestinal worm *Cooperia oncophora* in cattle. *Vet Parasitol*. 2015;211:241–50.
94. Li RW, Rinaldi M, Capuco AV. Characterization of the abomasal transcriptome for mechanisms of resistance to gastrointestinal nematodes in cattle. *Vet Res*. 2011;42:114.
95. Hempel RJ, Bannantine JP, Stabel JR. Transcriptional profiling of ileocecal valve of Holstein dairy cows infected with *Mycobacterium avium* subsp. *paratuberculosis*. *PLoS One*. 2016;11:e0153932.
96. McRae KM, McEwan JC, Dodds KG, Gemmill NJ. Signatures of selection in sheep bred for resistance or susceptibility to gastrointestinal nematodes. *BMC Genomics*. 2014;15:637.

97. Yaro M, Munyard KA, Morgan E, Allcock RJN, Stear MJ, Groth DM. Analysis of pooled genome sequences from Djallonke and Sahelian sheep of Ghana reveals co-localisation of regions of reduced heterozygosity with candidate genes for disease resistance and adaptation to a tropical environment. *BMC Genomics*. 2019;20:816.
98. Moiola B, D'Andrea S, De Grossi L, Sezzi E, De Sanctis B, Catillo G, et al. Genomic scan for identifying candidate genes for paratuberculosis resistance in sheep. *Anim Prod Sci*. 2016;56:1046–55.
99. Mallikarjunappa S, Sargolzaei M, Brito LF, Meade KG, Karrow N, Pant S. *J Dairy Sci*. 2018;101:7280–6.
100. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. A meta-assembly of selection signatures in vattle. *PLoS One*. 2016;11:e0153013.
101. Ramey HR, Decker JE, McKay SD, Rolf MM, Schnabel RD, Taylor JF. Detection of selective sweeps in cattle using genome-wide SNP data. *BMC Genomics*. 2013;14:382.
102. Lewis RS, Kolesnik TB, Kuang Z, D'Cruz AA, Blewitt ME, Masters SL, et al. TLR regulation of SPSB1 controls inducible nitric oxide synthase induction. *J Immunol*. 2011;187:3798–805.
103. Park H-E, Shin M-K, Park H-T, Jung M, Cho YI, Yoo HS. Gene expression profiles of putative biomarker candidates in *Mycobacterium avium* subsp. *paratuberculosis*-infected cattle. *FEMS Pathog Dis*. 2016;74:ftw022.
104. Sveriges Rödskulleförening. <http://rodskullan.se/R%3c3%b6dkullan/Historik.html>. Accessed 3 April 2020.
105. Stafuzza NB, de Oliveira Silva RM, Peripolli E, Bezerra LAF, Lôbo RB, de Ulhoa Magnabosco C, et al. Genome-wide association study provides insights into genes related with horn development in Nelore beef cattle. *PLoS One*. 2018;13:e0202978.
106. Cao X-K, Huang Y-Z, Ma Y-L, Cheng J, Qu Z-X, Ma Y, et al. Integrating CNVs into meta-QTL identified GBP4 as positional candidate for adult cattle stature. *Funct Integr Genomics*. 2018;18:559–67.
107. Bertolini F, Galimberti G, Calò D, Schiavo G, Matassino D, Fontanesi L. Combined use of principal component analysis and random forests identify population-informative single nucleotide polymorphisms: application in cattle breeds. *J Anim Breed Genet*. 2015;132:346–56.
108. Pant SD, Schenkel FS, Verschoor CP, You Q, Kelton DF, Moore SS, et al. A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in holstein cattle. *Genomics*. 2010;95:176–82.
109. Irano N, de Camargo GMF, Costa RB, Terakado APN, Magalhães AFB, de Oliveira Silva RM, et al. Genome-wide association study for indicator traits of sexual precocity in Nelore cattle. *PLoS One*. 2016;11:e0159502.
110. Al Kalaldehy M, Gibson J, Lee SH, Gondro C, Van Der Werf JH. Detection of genomic regions underlying resistance to gastrointestinal parasites in Australian sheep. *Genet Sel Evol*. 2019;51:37.
111. Ball CB, Rodriguez KF, Stumpo DJ, Ribeiro-Neto F, Korach KS, Blackshear PJ, et al. The RNA-binding protein, ZFP36L2, influences ovulation and oocyte maturation. *PLoS One*. 2014;9:e97324.
112. Bionaz M, Looor JJ. Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genomics*. 2008;9:366.
113. Suárez-Vega A, Gutiérrez-Gil B, Arranz JJ. Transcriptome expression analysis of candidate milk genes affecting cheese-related traits in 2 sheep breeds. *J Dairy Sci*. 2016;99:6381–90.
114. Terakado APN, Costa RB, de Camargo GMF, Irano N, Bresolin T, Takada L, et al. Genome-wide association study for growth traits in Nelore cattle. *Animal*. 2018;12:1358–62.
115. Fortes MRS, Lehnert SA, Bolormaa S, Reich CM, Fordyce G, Corbet N, et al. Finding genes for economically important traits: brahman cattle puberty. *Anim Prod Sci*. 2012;52:143–50.
116. Cochran SD, Cole JB, Null DJ, Hansen PJ. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genet*. 2013;14:49.
117. Connor EE, Kahl S, Elsasser TH, Parker JS, Li RW, Van Tassell CP, et al. Enhanced mitochondrial complex gene function and reduced liver size may mediate improved feed efficiency of beef cattle during compensatory growth. *Funct Integr Genomics*. 2010;10:39–51.
118. Varela L, Schwartz ML, Horvath TL. Mitochondria controlled by UCP2 determine hypoxia-induced synaptic remodeling in the cortex and hippocampus. *Neurobiol Dis*. 2016;90:68–74.
119. Wang Y, Yang W, Gui L, Wang H, Zan L. Association and expression analyses of the *Ucp2* and *Ucp3* gene polymorphisms with body measurement and meat quality traits in Qinchuan cattle. *J Genet*. 2016;95:939–46.
120. Allen HL, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*. 2010;467:832–8.
121. Kemper KE, Visscher PM, Goddard ME. Genetic architecture of body size in mammals. *Genome Biol*. 2012;13:244.
122. Sengar GS, Deb R, Singh U, Raja TV, Kant R, Sajjanar B, et al. Differential expression of microRNAs associated with thermal stress in Frieswal (*Bos taurus* x *Bos indicus*) crossbred dairy cattle. *Cell Stress Chaperon*. 2018;23:155–70.
123. Xu Q, Wang YC, Liu R, Brito L, Kang L, Yu Y, et al. Differential gene expression in the peripheral blood of Chinese Sanhe cattle exposed to severe cold stress. *Gen Mol Res*. 2017;16:gmr16029593.
124. Raven L-A, Cocks BG, Kemper KE, Chamberlain AJ, Vander Jagt CJ, Goddard ME, et al. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mamm Genome*. 2016;27:81–97.
125. Olivieri BF, Mercadante MEZ, Cyrillo JN, Branco RH, Bonilha SFM, de Albuquerque LG, et al. Genomic regions associated with feed efficiency indicator traits in an experimental Nelore cattle population. *PLoS One*. 2016;11:e0164390.
126. Orford M, Hadjipavlou G, Tzamaloukas O, Chatziplis D, Koumas A, Mavrogenis A, et al. A single nucleotide polymorphism in the acetyl-coenzyme A acyltransferase 2 (*ACA2*) gene is associated with milk yield in Chios sheep. *J Dairy Sci*. 2012;95:3419–27.
127. Gebreyesus G, Buitenhuis A, Poulsen NA, Visker MHPW, Zhang Q, Van Valenberg HJF, et al. Multi-population GWAS and enrichment analyses reveal novel genomic regions and promising candidate genes underlying bovine milk fatty acid composition. *BMC Genomics*. 2019;20:178.
128. Jiang J, Gao Y, Hou Y, Li W, Zhang S, Zhang Q, et al. Whole-genome resequencing of Holstein bulls for indel discovery and identification of genes associated with milk composition traits in dairy cattle. *PLoS One*. 2016;11:e0168946.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.