Genetic diversity of United States Rambouillet, Katahdin and Dorper sheep

Gabrielle M. Becker, Jacob W. Thorne, Joan M. Burke, Ronald M. Lewis, David R. Notter, James L. M. Morgan, Christopher S. Schauer, Whit C. Stewart, R. R. Redden, Brenda M. Murdoch

Genetics Selection Evolution

RESEARCH ARTICLE Open Access

Genetic diversity of United States Rambouillet, Katahdin and Dorper sheep

Gabrielle M. Becker[1], Jacob W. Thorne[1,2], Joan M. Burke[3], Ronald M. Lewis[4], David R. Notter[5], James L. M. Morgan[6], Christopher S. Schauer[7], Whit C. Stewart[8], R. R. Redden[2] and Brenda M. Murdoch[1*]

Abstract

Background Managing genetic diversity is critically important for maintaining species fitness. Excessive homozygosity caused by the loss of genetic diversity can have detrimental effects on the reproduction and production performance of a breed. Analysis of genetic diversity can facilitate the identification of signatures of selection which may contribute to the specific characteristics regarding the health, production and physical appearance of a breed or population. In this study, breeds with well-characterized traits such as fine wool production (Rambouillet, N = 745), parasite resistance (Katahdin, N = 581) and environmental hardiness (Dorper, N = 265) were evaluated for inbreeding, effective population size (N_e), runs of homozygosity (ROH) and Wright's fixation index (F_{ST}) outlier approach to identify differential signatures of selection at 36,113 autosomal single nucleotide polymorphisms (SNPs).

Results Katahdin sheep had the largest current N_e at the most recent generation estimated with both the GONe and NeEstimator software. The most highly conserved ROH Island was identified in Rambouillet with a signature of selection on chromosome 6 containing 202 SNPs called in an ROH in 50 to 94% of the individuals. This region contained the *DCAF16*, *LCORL* and *NCAPG* genes that have been previously reported to be under selection and have biological roles related to milk production and growth traits. The outlier regions identified through the F_{ST} comparisons of Katahdin with Rambouillet and Dorper contained genes with known roles in milk production and mastitis resistance or susceptibility, and the F_{ST} comparisons of Rambouillet with Katahdin and Dorper identified genes related to wool growth, suggesting these traits have been under natural or artificial selection pressure in these populations. Genes involved in the cytokine-cytokine receptor interaction pathways were identified in all F_{ST} breed comparisons, which indicates the presence of allelic diversity between these breeds in genomic regions controlling cytokine signaling mechanisms.

Conclusions In this paper, we describe signatures of selection within diverse and economically important U.S. sheep breeds. The genes contained within these signatures are proposed for further study to understand their relevance to biological traits and improve understanding of breed diversity.

*Correspondence:
Brenda M. Murdoch
bmurdoch@uidaho.edu
Full list of author information is available at the end of the article

Background

Genetic diversity is an important resource in animal production and conservation. Loss of genetic diversity can impact a species or a breed's ability to adapt to changing environmental or production pressures [1] and can lead to the accumulation of lethal or deleterious alleles with detrimental effects on health and production [2]. Adaptability of animal production requires the preservation of genetic diversity, and the identification of signatures of selection can provide essential insights that can be useful for conservation and breed improvement objectives [3, 4].

Artificial and natural selection both result in changes in allele frequencies that can lead to the fixation of certain alleles [5]. The neutral regions that surround alleles under selection tend to lose genetic variation due to the hitchhiking effect, which increases linkage disequilibrium (LD) and can result in the formation of selective sweeps [6–9]. Analyses of runs of homozygosity (ROH) can be used to detect regions that have experienced loss of heterozygosity due to the presence of selective pressures [10]. Analysis based on Wright's fixation index (F_{ST}) identifies differences in allele frequencies between populations and is one of the most commonly used methods to identify single nucleotide polymorphisms (SNPs) under selection [11–13]. In sheep, these analyses have been used to understand the genomic regions under selection for traits such as environmental adaption [14–17], parasite resistance and susceptibility [18], morphological traits [19–21], wool quality [22], and production traits [23] in many breeds.

The purpose of this study was to evaluate measures of genetic diversity and signatures of selection in three popular U.S. sheep breeds (Rambouillet, Katahdin, and Dorper) with diverse characteristics. The U.S. Rambouillet was first established in 1840 with the import of fine-wool sheep from France [24]; today, Rambouillet is a multi-purpose breed that is noted for wool and carcass quality. Compared to the other breeds of this study, Rambouillet is a larger framed, later-maturing breed [25] and is prominent in the semi-arid western states where much of the U.S. sheep production is concentrated [26]. The Katahdin is a composite hair breed that was developed in the 1950s from St. Croix hair sheep and wool breeds including the Suffolk and Wiltshire Horn [27, 28]. Katahdin is a fast-growing and prolific breed that is regarded for parasite resistance and suitability to warm, tropical environments [29, 30]. Dorper sheep are early-maturing and produce heavily muscled carcasses with many favorable palatability and tenderness characteristics [31]. To develop the Dorper breed, Dorset Horn and fat-rumped Blackhead Persian breeds were crossed to combine maternal traits with wool-shedding ability and adaptability to harsh environmental conditions [32].

Genotype data of Rambouillet, Katahdin, and Dorper animals were analyzed for inbreeding, effective population size (N_e), and signatures of selection (through F_{ST} and ROH) in order to gain understanding of the selection pressures that affect these breeds. Genes present in regions under selection were further evaluated to identify the biological pathways that are most likely affected. The results of this study provide insights into the genetic variation present in breeds commonly raised in the U.S. for their meat and wool quality (Rambouillet), parasite resistance (Katahdin), and environmental hardiness (Dorper).

Methods

Animal sampling and DNA genotyping

Sheep belonging to the Rambouillet, Katahdin, and Dorper breeds were selected for analyses of inbreeding and genetic diversity. These breeds were chosen to facilitate analyses of sheep that have been selected for diverse purposes, including fine wool (Rambouillet), parasite resistance (Katahdin), and environmental hardiness (Dorper). In total, 745 Rambouillet sheep were sampled from the Texas A&M AgriLife Research flock (TAMU; N = 403) and from central performance ram tests held at the North Dakota State University (NDSU; N = 159) and University of Wyoming (UWY; 183), representing more than 30 seedstock producers located in Colorado, Montana, North Dakota, South Dakota, and Wyoming. The Katahdin sheep (N = 581) were sampled from 20 flocks located across the U.S. (Arkansas, Georgia, Idaho, Indiana, Missouri, New York, Ohio, Oregon, Texas, Virgina, West Virginia, and Wisconsin). The Dorper sheep were sampled from the TAMU research flock (N = 265), which was founded from Dorper or White Dorper rams and ewes incorporated from over 20 producers throughout the U.S. in the early 2000s [33]. Over the last two decades, this flock has been managed as one cohort. In total, 1591 sheep were analyzed in the current study.

Sample collection and genotyping of these sheep have been described previously [34–36]. Briefly, DNA extraction and genotyping of Katahdin sheep were conducted at Neogen Corporation-GeneSeek Operations, Lincoln, NE, USA, whereas DNA from the Rambouillet and Dorper animals was extracted from whole blood samples using the phenol–chloroform method or from tissue samples by AgResearch (GenomNZ, AgResearch, New Zealand) [37]. Genotyping was conducted with the high-density (HD) Illumina 600K SNP BeadChip (Illumina Inc., San Diego, CA, USA) that includes 606,006 SNPs, the Applied Biosystems™ Axiom™ Ovine Genotyping Array (50K) that includes 51,572 SNPs (Thermo Fisher

Scientific, catalog number 550898), or the AgResearch Sheep Genomics 60K SNP chip that includes 68,848 SNPs (GenomNZ, AgResearch, New Zealand). In total, 132 Dorper and 243 Rambouillet samples were genotyped on the AgResearch chip, 133 Dorper and 502 Rambouillet samples were genotyped on the Applied Biosystems array, and the 581 Katahdin samples were genotyped on the HD Illumina chip.

Duplicate markers within a panel were filtered to retain the SNPs with the highest call rate (CR) at each unique position. Compatible SNPs across panels were merged using the Plink v1.9 software [38, 39]. Non-autosomal SNPs and SNPs with a CR lower than 90% were removed, which resulted in a dataset of 36,113 SNPs. The 1591 sheep had a genotype CR ≥ 90%. Removing SNPs with a low minor allele frequency (MAF) or that are in high LD can hinder the detection of ROH [40] and MAF thresholds can affect the comparison of structure between populations [41]. For these reasons, no further filtering was applied to the data prior to principal component analysis (PCA), ROH analyses, Wright's fixation index (F_{ST}) outlier approach, and N_e estimation.

The Rambouillet and Katahdin sheep analyzed in this study were sampled from geographically distant flocks and can be considered representative of their breed in the context of the U.S. sheep industry. Due to a limited representation of the current breed diversity in the sampled Dorper animals, the allele frequencies reported in these sheep are not anticipated to be representative of the U.S. Thus, for the Dorper breed, although analyses of these animals are still valuable and worth reporting, care should be taken to avoid over interpretation. For this reason, the Dorper ROH and N_e results are shared as supplementary data. The Dorper data were used in the pairwise F_{ST} signatures of selection analyses to facilitate interpretation of outlier SNPs that are potentially under selection in the Rambouillet and Katahdin breeds.

Population structure, inbreeding, and effective population size

The population structure between breeds was visualized through PCA that was conducted with the Plink v1.9 software. The results were visualized with the package ggplot2 in R version 4.2.3, with the first principal component (PC1) plotted on the x-axis and PC2 plotted on the y-axis [42–44]. The proportions of variance explained (PVE) by PC1 and PC2 were calculated by dividing the first and second eigenvalues, respectively, by the sum of all eigenvalues.

To understand the level of inbreeding within each breed, homozygosity was evaluated through the proportion of observed and expected homozygous SNPs and through two methods of inbreeding calculations. The number of expected and observed homozygous SNPs and the method-of-moments inbreeding coefficient (F) were calculated using Plink v1.9 (–het) for each animal [45]. The F statistic is calculated as ([observed homozygous count] – [expected count])/([total observations] – [expected count]). ROH-based inbreeding coefficients were calculated with the package detectRuns using the parameters described for the ROH analysis. The genome-wide ROH-based inbreeding coefficient, F_{ROH}, was calculated as the sum of the individual's ROH length over the length of the genome [46]. The distribution of F_{ROH} by breed was visualized with ggplot2. The Kruskal–Wallis test was used as a non-parametric alternative to the one-way analysis of variance (ANOVA) to determine whether there were statistical differences in F or F_{ROH} between breeds [47]. The Kruskal–Wallis test was followed by the Dunn's test to calculate *P*-values for pairwise breed comparisons [48].

The N_e of each breed was estimated using three methods. Historic N_e was estimated using the LD-based method in the SNeP software version 1.11 for each breed [49]. Default parameters were set for each analysis. LD was measured using r^2, the squared correlation coefficient between pairs of SNPs [50]. The rate of the decline in N_e was calculated between each consecutive reported generation and for the overall distribution. The NeEstimator v2.1 software was used to estimate current N_e for each breed using the LD method within the random mating model [51]. For comparison, current and historic N_e were also estimated using the GONe software [52], with default parameters. Results of SNeP, NeEstimator, and GONe were visualized in R with ggplot2.

Runs of homozygosity and F_{ST} analyses

Analysis of ROH was conducted to identify regions which may be under selection pressure in the Rambouillet, Katahdin, and Dorper breeds. The ROH were identified using the sliding window method with the R package detectRuns [46]. Each window comprised 15 SNPs with a maximum of two opposite SNP genotypes (heterozygous/homozygous) and one missing SNP allowed per window. A ROH was required to have a minimum length of 250,000 bp and to contain at least 30 SNPs. There was a minimum density of one SNP per every 10,000 kb and a maximum gap between SNPs of 10,000 kb. A homozygous window threshold of 0.05 was used to determine SNP inclusion in a ROH. Results were visualized using the CMPlot, ggplot2, and patchwork packages in R [53, 54]. The regions that contained a ROH in more than 50% of the individuals of a breed were selected for further analyses.

The average length of ROH called in Katahdin and Rambouillet were investigated for significant differences

between breeds. All data were evaluated for normality with the Shapiro-Wilks test prior to model selection with the 'shapiro.test' function in R [44, 55]. The Wilcoxon unpaired two-sample test was performed using the 'compare_means' function in the ggpubr package of R [43, 56].

To better understand the genetic differences between breeds, Wright's fixation index (F_{ST}) was calculated in Plink v2.0 using the method proposed by Weir and Cockerham [39, 57]. Analysis with Plink v2.0 was preferred over Plink v1.9 because of its ability to generate pairwise F_{ST} output files simultaneously for all comparisons. Per-variant F_{ST} estimates were generated for each SNP and genome-wide F_{ST} estimates were reported as ratio-of-averages between each pair of breeds, which were calculated from the ratio of the average variance components [58]. Fisher's exact test in R was used to estimate p-values from contingency tables constructed with genotype counts for each SNP and each pair of breeds. Outlier SNPs with F_{ST} estimates greater than three standard deviations above the mean and with significant Fisher's test *P*-values (< 0.05) were considered to differ greatly between breeds. For each pairwise breed analysis, F_{ST} regions of interest were defined by significant outlier SNPs located within 200,000 bp of each other, ± 100,000 bp before and after the flanking markers. For the purpose of comparison, F_{ST} was also estimated through the BayeScan 2.1 software with default parameters [59]. Genotype files were prepared for import into BayesScan using the 'genomic_converter' tool in the R package radiator [60].

The identified ROH and F_{ST} regions were queried through the National Center for Biotechnology Information (NCBI) genome browser tool [61, 62]. The ARS UI_Ramb_v2.0 genome assembly [63] was used for all SNP positions and gene region analyses. Characterized genes falling within a ROH or F_{ST} region were recorded for pathway enrichment analyses. Where possible, the predicted human ortholog of *Ovis aries* genes containing the LOC symbol (pseudo or uncharacterized genes) were recorded.

Pathway analyses

Genes of interest from ROH and F_{ST} analyses were investigated to identify associated biological pathways. Pathway enrichment analyses were conducted through the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Mapper search tools [64–67]. Due to reference database availability, queries through GO for biological process, molecular function, and cellular component were made against the *Bos taurus* reference database and queries of KEGG Mapper were made against the *Homo sapiens* reference. Pathway analyses were conducted for genes located within Rambouillet, Katahdin, or Dorper ROH islands and for genes located within Katahdin-Rambouillet, Katahdin-Dorper, or Rambouillet-Dorper F_{ST} regions. In addition, pathway analyses were conducted for genes which were in common between both F_{ST} analyses of Rambouillet (Katahdin-Rambouillet and Rambouillet-Dorper) or both F_{ST} analyses of Katahdin (Katahdin-Rambouillet and Katahdin-Dorper). To better understand the biological implications of genes within ROH islands and F_{ST} outlier regions, the STRING database was queried for protein–protein interaction networks and functional enrichment analysis of these genes [68].

Results

Population structure, inbreeding, and effective population size

The population structure of the studied animals was first evaluated through PCA. The analysis revealed three clearly distinct clusters, with Rambouillet, Katahdin, and Dorper sheep clustering more closely with individuals of the same breed than individuals of other breeds (see Additional file 1: Fig. S1). The model's PC1 had an eigenvalue of 179.82 and was estimated to explain 35.99% of all the variation while PC2 had an eigenvalue of 116.82 and was estimated to explain 23.38% of the variation. The largest principal component separated the Katahdin and Rambouillet breeds, while PC2 separated Dorper from both Katahdin and Rambouillet. The placement of these breeds is similar to the results of previous PCA, which have shown the separation of breeds originating from West Africa, South Africa, and the Iberian Peninsula [69, 70].

The average proportion of expected and observed homozygous SNPs for Rambouillet, Katahdin, and Dorper are reported in Table 1. The observed homozygosity

Table 1 Homozygosity and inbreeding coefficients for the Dorper, Katahdin, and Rambouillet breeds

Breed	Expected Hom. AVG ± SD	Observed Hom. AVG ± SD	F AVG ± SD	Range F	F_{ROH} AVG ± SD	Range F_{ROH}
Rambouillet	0.6216 ± 0.0002	0.6620 ± 0.0218	0.1069 ± 0.06	−0.057 to 0.389	0.1690 ± 0.06	0.044–0.442
Katahdin	0.6216 ± 0.0002	0.6628 ± 0.0169	0.1090 ± 0.04	−0.057 to 0.364	0.1865 ± 0.04	0.027–0.428
Dorper	0.6215 ± 0.0001	0.6619 ± 0.0182	0.1067 ± 0.05	0.016 to 0.374	0.1875 ± 0.05	0.099–0.437

F is the method-of-moments inbreeding coefficient, F_{ROH} is the ROH-based inbreeding coefficient

Hom.: homozygosity; SD: standard deviation

was similar to previous estimates for Rambouillet [71] and was slightly higher than previous estimates for Katahdin and Dorper [72, 73]. For each breed, the average F_{ROH} estimate was higher than the average F estimate. The Kruskal–Wallis *P*-value for 'F ~ Breed' was 4.90e−04, and for 'F_{ROH} ~ Breed' was 2.89e−25. Based on the Dunn's test, Katahdin had significantly higher F inbreeding coefficients than Rambouillet; for F_{ROH}, the Rambouillet breed had significantly lower inbreeding coefficients than both Dorper and Katahdin (Fig. 1). Overall, F_{ROH} estimates were higher than previously reported estimates for worldwide sheep populations [74, 75]. Previous studies have sampled fewer individuals per breed, which may contribute to these observations.

Estimates of N_e were calculated using the SNeP, NeEstimator, and GONe software. The N_e of the Rambouillet and Katahdin breeds were estimated from 759 generations ago to 13 generations ago using SNeP, from 741 generations ago to one generation ago using GONe, and for the current generation (given the notation 0) with NeEstimator (Fig. 2). Estimations with GONe were non-linear, while estimations with SNeP showed a consistent decline in N_e from the furthest generation to the most recent. These general observations of SNeP and GONe trajectories are similar to the N_e results reported in the Master thesis work of D Adepoju with cattle data [76]. The GONe and SNeP estimates began to show more agreement approximately 30 generations ago (Fig. 2a). In the most recent generation, Katahdin were estimated to have an N_e of 161.4 by NeEstimator or 436.1 by GONe, while Rambouillet had N_e estimates of 56.9 by NeEstimator or 111.8 by GONe (Fig. 2b).

The rate of change (m) of N_e size over generational time can give insight into the rate of diversity loss. These calculations were made with the results of SNeP, as these estimates followed a linear pattern. The overall rate of

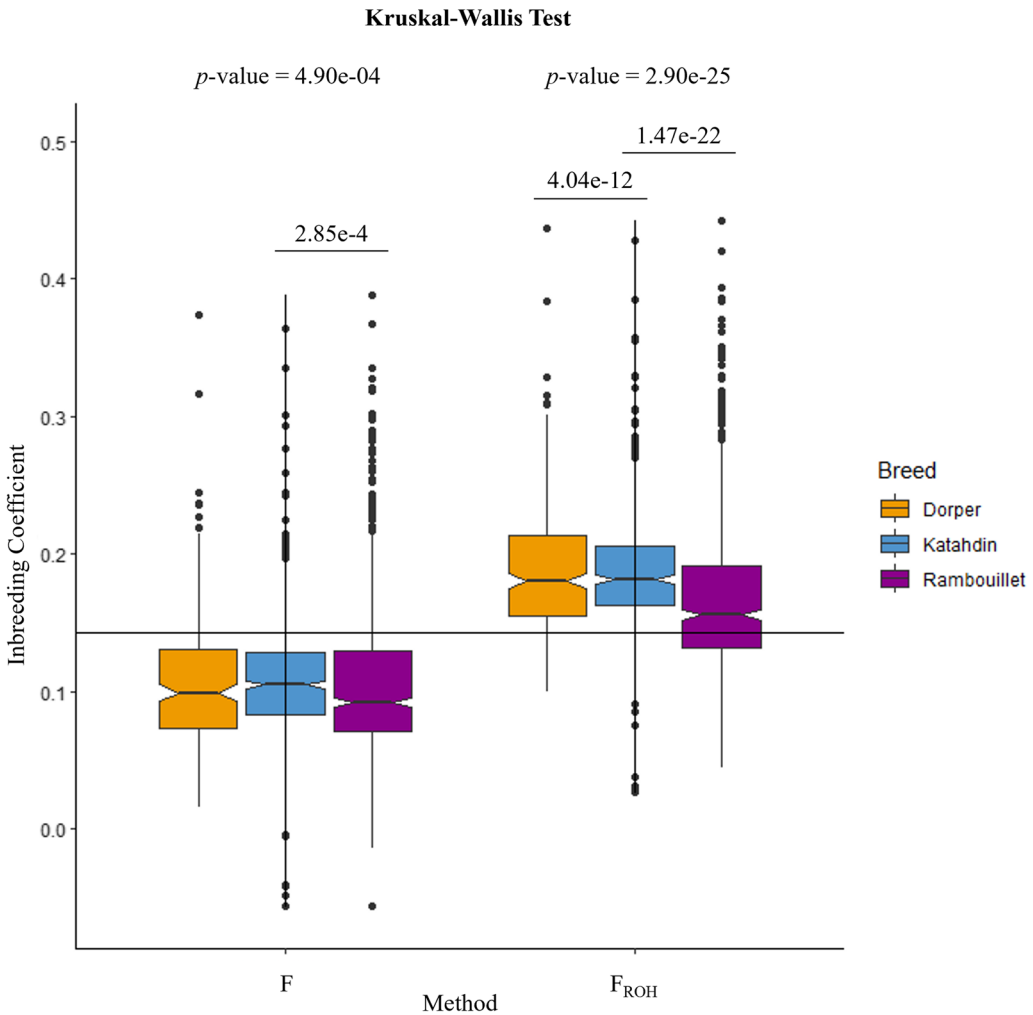

Fig. 1 Distribution of F and F_{ROH} inbreeding coefficients by breed. Boxplots of the distribution of F and F_{ROH} for Dorper, Katahdin, and Rambouillet breeds. The horizontal black line represents the overall mean inbreeding coefficient calculated from both F and F_{ROH} estimates

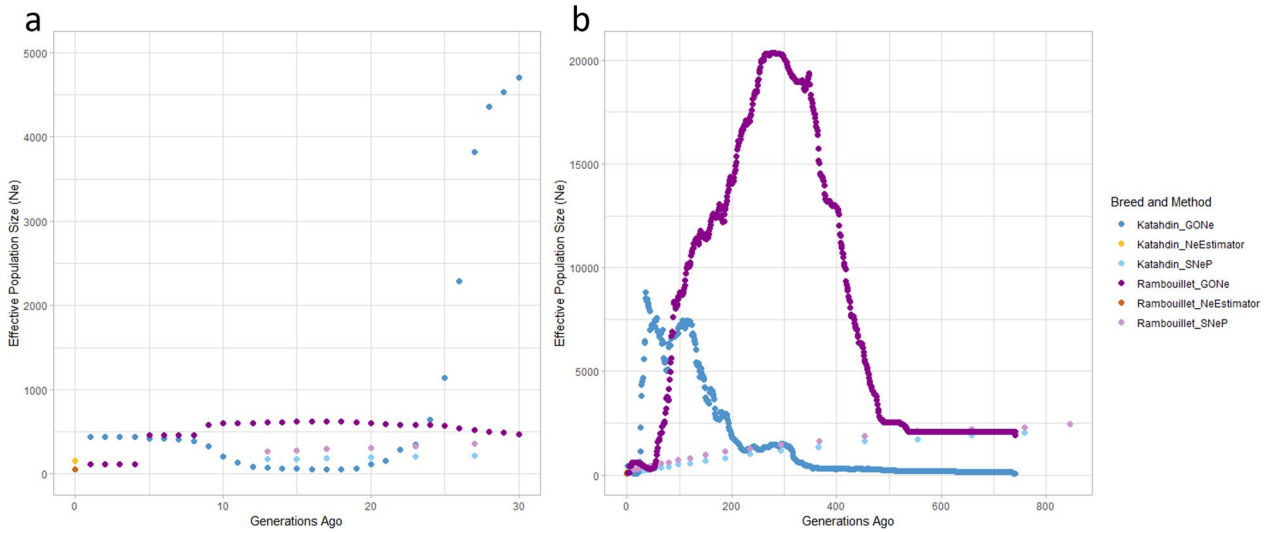

Fig. 2 Recent and historic N_e estimated through three methods. **a** N_e estimates from 30 generations ago to the current one for Katahdin and Rambouillet, estimated with GONe, NeEstimator, and SNeP; **b** Historic N_e estimates for Katahdin and Rambouillet estimated with GONe and SNeP

change was m = 3.02 for Rambouillet and m = 2.76 for Katahdin, suggesting that the Rambouillet breed has lost genetic diversity at a slightly more rapid rate than the Katahdin breed. The most rapid decrease in N_e occurred between 80 and 98 generations ago for the Katahdin (m = 5.11; Table 2) and between 15 and 17 generations ago for the Rambouillet breed (m = 9.00; Table 3). Katahdin also had a high rate of change 15 to 17 generations ago (m = 5, Table 3). In addition, both Rambouillet and Katahdin showed higher rates of change between 20 and 23 generations ago, with Rambouillet having an m of 7.33, and Katahdin having an m of 5.00 between these intervals. Each subsequent generation had a lower N_e estimate than the preceding generation, and in each generation the estimate for Rambouillet was greater than that for Katahdin. The N_e estimates for these breeds became more similar as the generations became closer to the current one. Analysis results for the Dorper sheep are reported as supplementary data (see Additional file 2: Table S1). The greatest rate of change for Dorper was estimated from 17 to 20 generations ago, with m = 4.00, and the overall rate of change was m = 2.05.

Runs of homozygosity and Wright's F_{ST} analyses

There were 51,992 and 72,946 ROH called for Katahdin and Rambouillet, respectively. The ROH were categorized into five classes by length, with classes defined by 0–6 Mb, 6–12 Mb, 12–24 Mb, 24–48 Mb, and > 48 Mb. The majority of the ROH identified were shorter than 6 Mb in length; 86% of the Rambouillet total ROH and 75% of the Katahdin total ROH were within this class (see Additional file 3: Table S2). The percentages of 6–12 Mb and 12–24 Mb long ROH were greater in Katahdin than in Rambouillet (see Additional file 4: Fig. S2). The average lengths of ROH called in Katahdin and Rambouillet were significantly different (Wilcoxon *P*-value = 1.29e−82), with Rambouillet having significantly shorter mean ROH than Katahdin (Fig. 3). The differences observed in mean ROH length and in percentage of ROH by class may reflect differences in LD decay at signatures of selection between these breeds. As LD breaks down rapidly over distance, short ROH can indicate more ancient inbreeding or selection events while long ROH are indicative of more recent selection [77, 78].

A ROH island was defined by the presence of SNPs called within a ROH in 50% or more of the animals of a breed (Fig. 4). The ROH analysis identified three ROH islands in Rambouillet and two ROH islands in Katahdin (Table 4). The ROH island identified in Rambouillet on chromosome 6 had both the largest number of SNPs called in an ROH island (202 SNPs) and the highest percentage of animals called in an ROH at the same SNP (94.23%). ROH islands in Rambouillet were identified on chromosome 3 (Fig. 5a), chromosome 6 (Fig. 5b), and chromosome 7 (Fig. 5c), and those in Katahdin on chromosome 23 (Fig. 6a) and chromosome 25 (Fig. 6b). There was some overlap between ROH islands called in Rambouillet and Dorper sheep on chromosome 6, although the ROH called in Rambouillet were both longer and more highly conserved within the breed (see Additional file 5: Table S3).

Table 2 Effective population size (N_e) for Katahdin sheep

Generations ago	N_e	Average distance (bp)	Average LD (r^2) ± SD	Number of pairwise comparisons	Rate of change (m)
13	178	3,749,388	0.0361 ± 0.05	38,212	0.50
15	179	3,272,881	0.0410 ± 0.06	34,640	5.00*
17	189	2,843,420	0.0445 ± 0.07	31,503	1.67
20	194	2,460,514	0.0497 ± 0.07	28,092	5.00*
23	209	2,116,735	0.0535 ± 0.08	25,271	2.00
27	217	1,811,459	0.0597 ± 0.08	22,366	4.40
32	239	1,541,337	0.0637 ± 0.09	19,760	2.67
38	255	1,303,408	0.0700 ± 0.10	17,620	4.14
45	284	1,095,495	0.0745 ± 0.10	15,484	3.67
54	317	914,182	0.0795 ± 0.11	13,424	4.55
65	367	757,894	0.0825 ± 0.11	11,640	3.40
80	418	623,867	0.0876 ± 0.12	10,128	5.11*
98	510	509,861	0.0878 ± 0.12	8529	3.27
120	582	413,447	0.0941 ± 0.13	7034	4.07
150	704	333,299	0.0962 ± 0.13	5838	2.59
187	800	266,817	0.1049 ± 0.14	4792	4.43
234	1008	213,215	0.1042 ± 0.14	3832	3.20
293	1197	170,203	0.1093 ± 0.14	3189	2.03
366	1345	136,424	0.1199 ± 0.16	2460	3.24
454	1630	110,075	0.1223 ± 0.16	1872	1.00
554	1730	90,236	0.1380 ± 0.18	1428	1.82
657	1917	76,085	0.1464 ± 0.19	939	1.45
759	2065	65,804	0.1554 ± 0.20	698	*Overall 2.76*

The rate of change in N_e (m) was calculated as the change in N_e over the change in generations ago ($\Delta N_e/\Delta$generation). The greatest rates of change are indicated with (*) and the overall rate of change is given in italics

The genomic regions of ROH islands were evaluated for the presence of known and predicted genes. Seventy-six genes were identified in the Katahdin ROH islands, including 20 uncharacterized LOC genes without predicted *Homo sapiens* orthologs, and two copy number variants (CNV) of the *TRNAC-GCA* gene. Fifty-five unique genes were used for pathway enrichment queries (see Additional file 6: Table S4). For Rambouillet, 117 unique genes, multiple copies of tRNA genes *TRNAS-GGA*, *TRNAH-AUG*, *TRNAW-CCA*, and *TRNAC-GCA*, and 37 uncharacterized LOC genes were identified between the ROH islands on chromosomes 3 and 6 (see Additional file 7: Table S5). The ROH island identified on chromosome 7 was intergenic. Known or predicted protein–protein interactions of genes located within ROH islands were identified through query of the STRING database. For genes within the Rambouillet ROH island on chromosome 6, the largest number of interactions was found for FAM184B (9 proteins), FAM13A (7 proteins), and for CCSER1, LAP3, LCORL, MED28, and NCAPG, all of which had interactions with six proteins. Within the ROH island on chromosome 3, the largest number of interactions was found for TMEM117 (7 proteins), ZCRB1 (6 proteins), and HNRNPA1 (5 proteins). Analysis of protein–protein interactions for the Katahdin ROH island on chromosome 23 identified four interactions with NDUFV2, and analysis of the ROH island on chromosome 25 contained only interactions of ARID4B with TOMM20 and RBM34 (see Additional file 8: Table S6).

Pairwise F_{ST} of 0.140, 0.156, and 0.161 were estimated between Katahdin and Rambouillet, between Rambouillet and Dorper, and between Katahdin and Dorper, respectively. Thresholds of + 3 standard deviations above the average were calculated for each breed comparison to identify outlier SNPs that showed the greatest amount of differentiation. As expected from the pairwise estimates, the Katahdin-Rambouillet SNP comparisons showed the lowest average F_{ST} and smallest standard deviation. The F_{ST} threshold for Katahdin-Rambouillet outlier SNPs was 0.52, and for both Rambouillet-Dorper and Katahdin-Dorper comparisons, the threshold was 0.58 (Table 5). The pairwise F_{ST} statistics for each SNP in the Katahdin-Rambouillet analysis were compared against ROH island results (Fig. 4).

The nearest gene of the ten SNPs with the highest F_{ST} estimates from each pairwise F_{ST} analysis was

Table 3 Effective population size (N_e) for Rambouillet sheep

Generations ago	N_e	Average distance (bp)	Average LD (r^2) ± SD	Number of pairwise comparisons	Rate of change (m)
13	268	3,748,453	0.0243 ± 0.04	39,003	4.00
15	276	3,272,995	0.0269 ± 0.04	35,589	9.00*
17	294	2,843,334	0.0291 ± 0.04	31,860	5.67
20	311	2,459,704	0.0316 ± 0.05	28,845	7.33*
23	333	2,116,548	0.0343 ± 0.05	25,565	6.50*
27	359	1,810,664	0.0370 ± 0.05	23,218	5.00
32	384	1,541,490	0.0406 ± 0.06	20,609	5.83
38	419	1,303,677	0.0437 ± 0.06	18,002	4.86
45	453	1,095,627	0.0479 ± 0.07	15,746	6.00
54	507	914,251	0.0512 ± 0.07	13,836	5.45
65	567	757,660	0.0550 ± 0.08	11,965	3.20
80	615	623,577	0.0612 ± 0.09	10,246	5.89
98	721	509,735	0.0637 ± 0.09	8553	3.39
121	799	413,170	0.0704 ± 0.10	7279	6.68*
149	986	333,486	0.0707 ± 0.10	5903	4.08
187	1141	267,022	0.0758 ± 0.10	4976	2.62
234	1264	213,570	0.0848 ± 0.12	4033	3.80
293	1488	170,488	0.0897 ± 0.12	3228	2.27
367	1656	136,133	0.0998 ± 0.14	2507	2.63
454	1885	110,105	0.1075 ± 0.14	1972	2.65
554	2150	90,160	0.1143 ± 0.16	1415	0.63
657	2215	75,994	0.1293 ± 0.17	959	0.69
759	2285	65,817	0.1425 ± 0.18	702	*Overall 3.02*

The rate of change in N_e (m) was calculated as the change in N_e over the change in generations ago ($\Delta N_e/\Delta$generation). The greatest rates of change are indicated with (*) and the overall rate of change is given in italics

determined (Table 6). The SNP OAR2_231739122.1 had the highest F_{ST} (0.9069) in the Katahdin-Rambouillet analysis and was positioned 935 bp downstream of the *CXCR2* gene. For the Katahdin-Dorper analysis, the SNP OAR2_88734520.1 had an F_{ST} value of 0.9735 and was within the *CCDC171* gene, and for Rambouillet-Dorper, the SNP OAR1_292866363.1 had an F_{ST} value of 0.8829 and was positioned within the *METTL6* gene. Among the top ten SNPs identified in the Rambouillet-Dorper analysis, four were located within or near the *LCORL* gene. In addition, all pairwise F_{ST} comparisons identified outlier SNPs within the region of the *FRY* and *RXFP2* genes (Table 7 and Fig. 7).

In total, 554 outlier SNPs were identified in the Katahdin-Rambouillet F_{ST} analysis. Of these 554 SNPs, 157 were positioned within 200,000 bp of at least one other outlier SNP, which together defined 63 F_{ST} regions containing 233 unique and characterized genes (see Additional file 9: Table S7). These genes included the tRNA genes *TRNAC-GCA* (four CNV), *TRNAW-CCA* (three CNV), *TRNAE-CUC* (two CNV), and *TRNAS-GGA* (one CNV). The F_{ST} regions were in part consistent with the ROH results, with seven SNPs being identified through both the Katahdin ROH and Katahdin-Rambouillet outlier F_{ST} analyses, and 18 SNPs being identified through both the Rambouillet ROH and Katahdin-Rambouillet outlier F_{ST} analyses (see Additional file 10: Table S8). The GO pathway enrichment analysis of these genes identified significant enrichment for biological processes including regulation of cellular glucuronidation and glucuronosyltransferase activity, and significant enrichment for the molecular function pathway UDP-glycosyltransferase activity (see Additional file 11: Table S9). KEGG Mapper pathway analysis revealed genes that are involved in many pathways, including, among others, B cell receptor signaling, taste transduction, circadian rhythm, cytokine-cytokine receptor interaction, endocytosis, growth hormone synthesis/secretion/action, hematopoietic cell lineage, HIF-1 signaling, IL-17 signaling, and longevity regulation pathways. In addition, there were a number of pathways related to viral infection, such as herpes simplex virus 1, human immunodeficiency virus 1, and human cytomegalovirus (see Additional file 12: Table S10).

Analyses with the Dorper breed were used to better understand allelic differentiation in the Katahdin

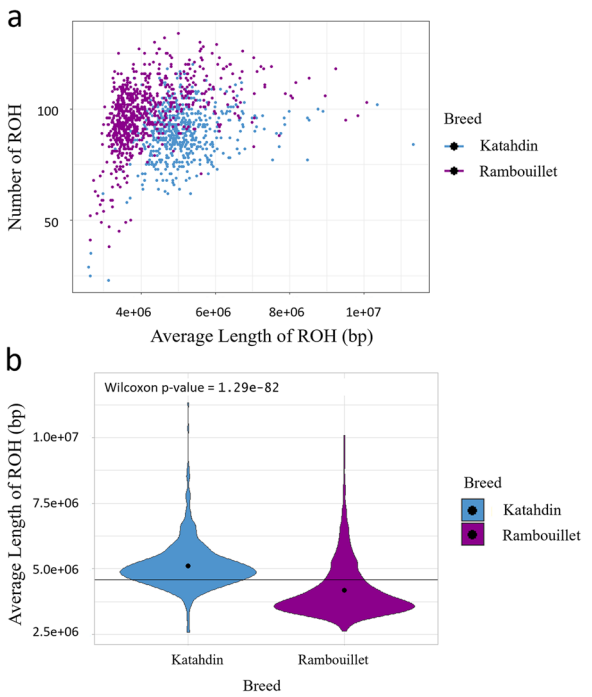

Fig. 3 Length of ROH called by breed. **a** Plot of the average length of ROH (x-axis) by the total number of ROH called per individual (y-axis), with the Katahdin sheep plotted in blue and the Rambouillet sheep in purple; **b** Average length of ROH by breed, with the overall mean indicated by the horizontal line and breed means indicated by black points. The Wilcoxon *P*-value of 'average length ~ breed' is 1.29e−82

and Rambouillet breeds. All F_{ST} regions were analysed to identify the genes that were present in both pairwise comparisons of a breed: for example, genes identified in both the Rambouillet-Dorper and Katahdin-Rambouillet F_{ST} analyses were further explored to give context to the F_{ST} signals associated with the Rambouillet. Fifty-one genes were in common for the Rambouillet, 31 for the Katahdin, and 21 for the Dorper comparisons (Table 8). Many of these genes were previously identified in studies on signatures of selection or are candidate genes for production and/or health traits in sheep. The genes reported in Table 8 for each breed were queried through KEGG. These analyses identified pathways for metabolic pathways, olfactory transduction, and cytokine-cytokine receptor interaction (Table 9). In addition, genes in common between ROH islands and F_{ST} regions were evaluated for Katahdin and Rambouillet breeds (Table 10).

The Katahdin-Dorper analysis identified 513 outlier SNPs, including 109 SNPs located near to each other, which defined 49 regions. The 144 genes contained within these regions were used for pathway analysis. Genes in these regions were part of KEGG pathway terms, including chemokine signaling and cytokine-cytokine receptor interaction (see Additional file 13: Table S11). In the Rambouillet-Dorper analysis, 569 outlier SNPs were identified, with 167 SNPs defining 73 F_{ST} regions. In total, 264 genes were located within these regions. Some of the pathways identified by KEGG Mapper analysis included biosynthesis of unsaturated fatty acids, neutrophil extracellular trap formation, olfactory transduction (with nine related genes), and prion disease (see Additional file 14: Table S12).

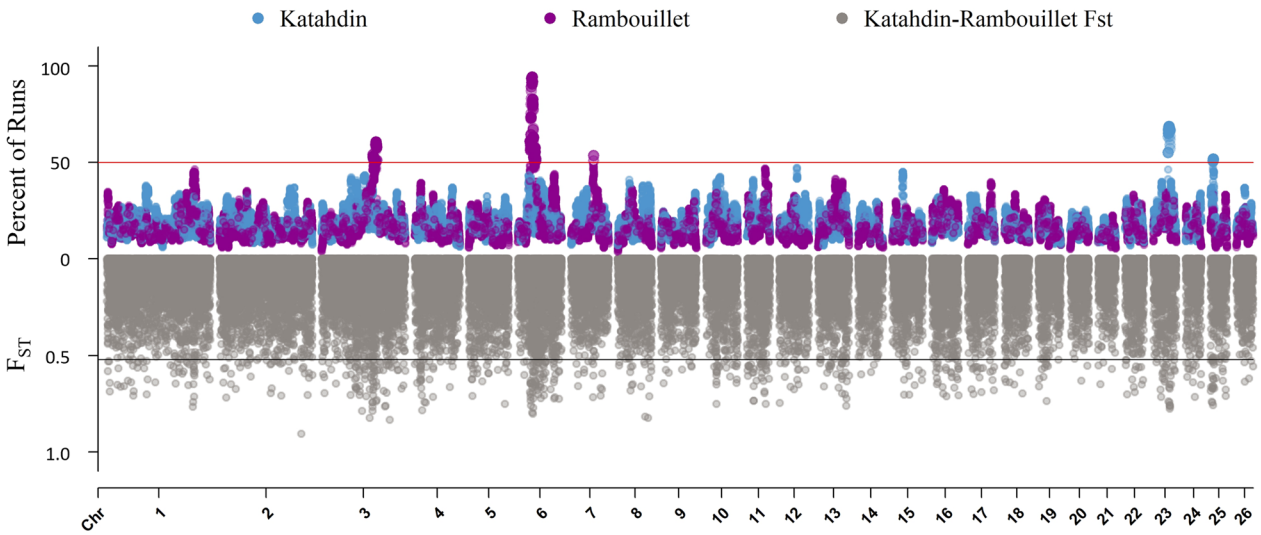

Fig. 4 Manhattan plot of ROH and F_{ST} results for Katahdin and Rambouillet sheep. The top of the plot (scale 0 to 100) is the percentage of SNPs called in an ROH in Katahdin (blue) and Rambouillet (purple) breeds. The bottom of the plot (scale 0 to 1.0) contains the results of the Katahdin-Rambouillet pairwise F_{ST} of each SNP (gray)

Table 4 Regions comprised of SNPs called within a ROH in 50% or more of the animals of a breed

Breed	Chr	Number of SNPs	AVG % SNPs in run	Max % SNPs in run	Range (bp)
Katahdin	23	50	65.82	68.67	40,163,730–45,746,963
	25	18	51.54	51.81	6,929,232–7,836,577
Rambouillet	3	71	55.70	60.67	138,563,593–148,939,496
	6	202	68.38	94.23	32,795,860–47,985,710
	7	3	52.62	53.56	60,566,239–60,647,707

Chr: chromosome number; AVG: average; Max: maximum

In Katahdin sheep, 50 SNPs were called within an ROH in 50% or more of Katahdin sheep in a region on chromosome 23 and 18 SNPs were called in a region on chromosome 25. In Rambouillet, a region on chromosome 3 contained 71 SNPs, chromosome 6 contained 202 SNPs, and chromosome 7 contained three SNPs

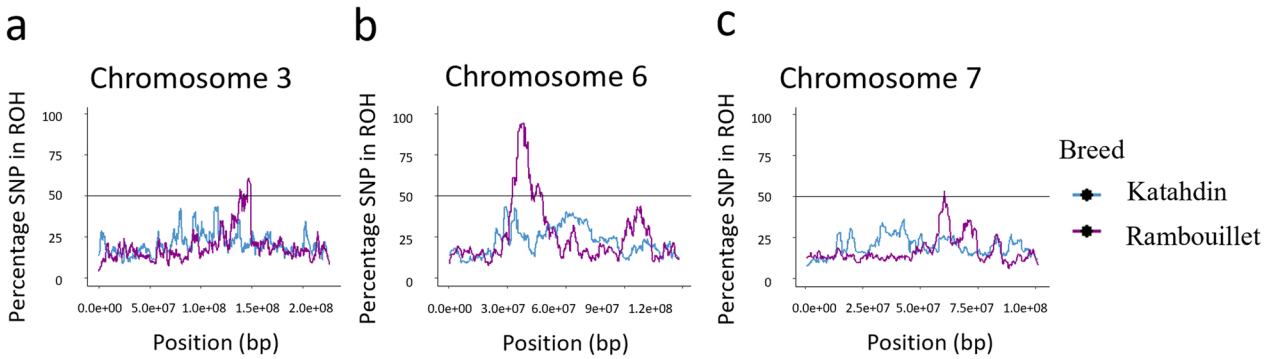

Fig. 5 Chromosomes with ROH called in 50% or more of the Rambouillet sheep. **a** ROH called on chromosome 3; **b** ROH called on chromosome 6; **c** ROH called on chromosome 7. The ROH called in Katahdin are plotted in blue and the ROH in Rambouillet are plotted in purple

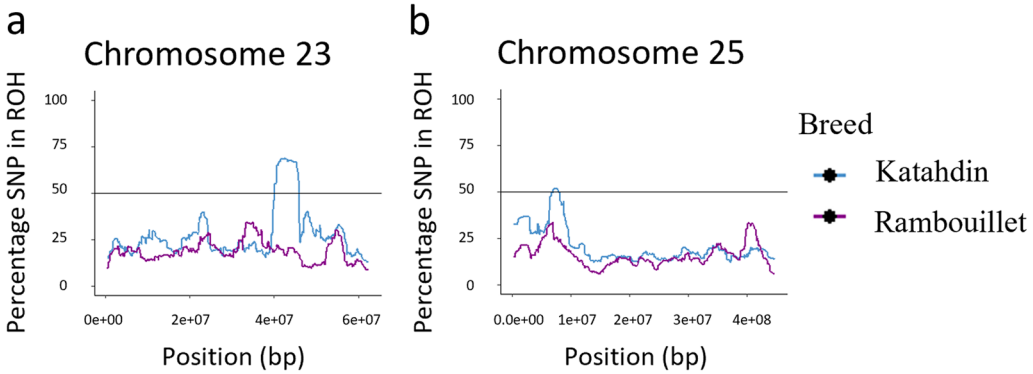

Fig. 6 Chromosomes with ROH called in 50% or more of the Katahdin sheep. **a** ROH called on chromosome 23; **b** ROH called on chromosome 25. The ROH called in Katahdin are plotted in blue and the ROH in Rambouillet are plotted in purple

Table 5 Average F_{ST} estimates between Katahdin, Dorper, and Rambouillet breeds

	Katahdin-Rambouillet	Katahdin- Dorper	Rambouillet-Dorper
AVG	0.1149	0.1319	0.1294
SD	0.1345	0.1493	0.1506
Threshold	0.5184	0.5799	0.5813

The threshold for calling outlier values was the average (AVG) + 3 standard deviation (SD)

The pathway analyses of the Dorper ROH results are provided as supplementary material (see Additional file 15: Table S13; Additional file 16: Table S14; Additional file 17: Table S15; Additional file 18: Table S16). In addition, the outlier F_{ST} results described here from the Weir and Cockerham model implemented in Plink were compared against the results of a BayeScan F_{ST} analysis (see Additional file 19: Table S17).

Table 6 Genomic information for 10 SNPs with the highest F_{ST} in each pairwise breed comparison

Pairwise breed comparison	Chr:bp	F_{ST}	Gene
Katahdin-Rambouillet			
OAR2_231739122.1	2:220,204,274	0.9069	935 bp downstream of *C-X-C motif chemokine receptor 2* (*CXCR2*)
OAR3_138331159.1	3:129,837,912	0.8236	24,522 bp from *mitochondrial ribosomal protein L42* (*MRPL42*)
OAR3_146778162.1	3:137,510,701	0.7882	*KAT8 regulatory NSL complex subunit 2* (*KANSL2*)
s39351.1	3:183,889,776	0.8339	8508 bp from *DENN domain containing 5B* (*DENND5B*)
OAR4_19418235.1	4:19,933,519	0.7842	1464 bp upstream of *PHD finger protein 14* (*PHF14*)
OAR6_42557643.1	6:39,000,766	0.796	207,018 bp from *protein SET-like* (LOC10110458)
OAR6_44450940.1	6:40,475,050	0.8022	*Slit guidance ligand 2* (*SLIT2*)
OAR8_78922714.1	8:73,722,766	0.8178	*TGF-beta activated kinase 1 (MAP3K7) binding protein 2 (TAB2)*
s01826.1	8:81,097,505	0.8236	65,227 bp from *AT-rich interaction domain 1B* (*ARID1B*)
OAR23_46900744.1	23:44,502,153	0.7756	3397 bp from heterogeneous nuclear ribonucleoproteins A2/B1-like (LOC101108009)
Katahdin-Dorper			
OAR2_88734520.1	2:84,038,487	0.9735	*Coiled-coil domain containing 171* (*CCDC171*)
OAR3_92705824.1	3:87,616,185	0.8486	*Cysteine rich transmembrane BMP regulator 1* (*CRIM1*)
s70203.1	4:35,066,909	0.9462	*Glutamate metabotropic receptor 3* (*GRM3*)
OAR6_64114132.1	6:58,970,273	0.8836	*KELCH like family member 5* (*KLHL5*)
OAR8_29480020.1	8:27,186,942	0.852	*methyltransferase like 24* (*METTL24*)
OAR14_25081675.1	14:24,402,748	0.8856	*Nucleoporin 93* (*NUP93*)
OAR14_53703182.1	14:51,456,105	0.9254	742 bp upstream of *zinc finger protein 575* (*ZNF575*)
s30024.1	25:6,684,555	0.8544	31,987 bp from *TAR (HIV-1) RNA binding protein 1* (*TARBP1*)
s44881.1	25:6,938,538	0.8551	27,307 bp from *small nucleolar RNA SNORA40* (LOC114110918)
OAR25_17689768.1	25:16,484,790	0.8508	49,538 bp from *transmembrane protein 26* (*TMEM26*)
Rambouillet-Dorper			
s41709.1	2:16,050,607	0.8904	48,461 bp from *zinc finger protein 462* (*ZNF462*)
OAR6_41003295.1	6:37,533,595	0.8818	1873 bp upstream of *matrix extracellular phosphoglycoprotein* (*MEPE*)
OAR6_41709987.1	6:38,197,739	0.896	*Ligand dependent nuclear receptor corepressor like* (*LCORL*)
OAR6_41768532.1	6:38,248,979	0.9007	26,028 bp from *LCORL*
OAR6_41850329.1	6:38,330,994	0.9009	108,043 bp from *LCORL*
OAR6_41925630.1	6:38,397,839	0.8971	174,888 bp from *LCORL*
s01370.1	9:27,800,948	0.8827	9194 bp from *tribbles pseudokinase 1* (*TRIB1*)
OAR14_25081675.1	14:24,402,748	0.8733	*Nucleoporin 93* (*NUP93*)
OAR16_42184944.1	16:39,117,011	0.9006	131,081 bp from *prolactin receptor* (*PR*)
OAR24_27348134_X.1	24:25,380,060	0.8806	12,165 bp from *interleukin 4 receptor* (*IL4R*)

Protein–protein interactions were identified for genes implicated by pairwise F_{ST} analyses with STRING. Query with the genes from the Katahdin-Rambouillet F_{ST} outlier regions revealed the largest number of interactions with ribosomal proteins RPL23A (19 proteins), RPL5 (18 proteins), and RPL10A, RPL6, and RPL7, each with 15 protein–protein interactions. Query with the genes from the Rambouillet-Dorper F_{ST} outlier regions identified 19 protein–protein interactions with RPL8, 18 interactions with RPL6, and 17 interactions each with RECQL4 and RPL10A. The largest number of interactions in the Katahdin-Dorper F_{ST} outlier regions were with EIF4A3, KCNQ1, NPM1, and TSSC4, all of which had seven protein–protein interactions (see Additional file 8: Table S6).

Discussion

In this study, we analyzed the genetic diversity and signatures of selection of Rambouillet, Katahdin, and Dorper sheep through within-breed (ROH approach) and pairwise breed (F_{ST} outlier approach) comparisons. The results reported here concern genomic regions that are under selection in the considered U.S. sheep. In addition, this study identified commonalities with previously identified signatures of selection from a diverse range of breeds, which allows our findings to contribute to a

Table 7 Fixation index (F_{ST}) scores for SNPs in the region harboring the *FRY* and *RXFP2* genes

SNP	Gene	Chr:bp	F_{ST}		
			Katahdin-Dorper	Rambouillet-Dorper	Katahdin-Rambouillet
s24045.1	*FRY, ZAR1L*	10:29,041,335	0.0531	0.6995*	0.4769
OAR10_29065568.1	*FRY*	10:29,071,493	0.0034	0.6139*	0.4614
OAR10_29223007.1	*FRY*	10:29,230,973	0.1191	0.3418	0.5809*
s02289.1	*FRY*	10:29,244,858	0.0007	0.5767	0.6366*
OAR10_29274817.1	*FRY*	10:29,281,423	−0.0012	0.5676	0.6031*
OAR10_29341212.1	*FRY*	10:29,347,138	0.0921	0.8234*	0.6517*
OAR10_29448537.1	Downstream of *RXFP2*	10:29,458,417	0.6443*	0.7264*	0.0164
OAR10_29469450.1	Downstream of *RXFP2*	10:29,479,765	−0.0001	0.6992*	0.7510*
OAR10_29654158.1	Upstream of *RXFP2*	10:29,654,912	0.8321*	0.7709	0.0094
OAR10_29737372.1	Upstream of *RXFP2*	10:29,753,249	0.5008	0.0115	0.5895*
s18834.1	Upstream of *RXFP2*	10:29,847,303	0.7320	0.0479	0.5521*

*Denotes that the F_{ST} score is above the breed comparison threshold of either 0.58 (Katahdin-Dorper and Rambouillet-Dorper) or 0.52 (Katahdin-Rambouillet) and is significant by Fisher's exact test

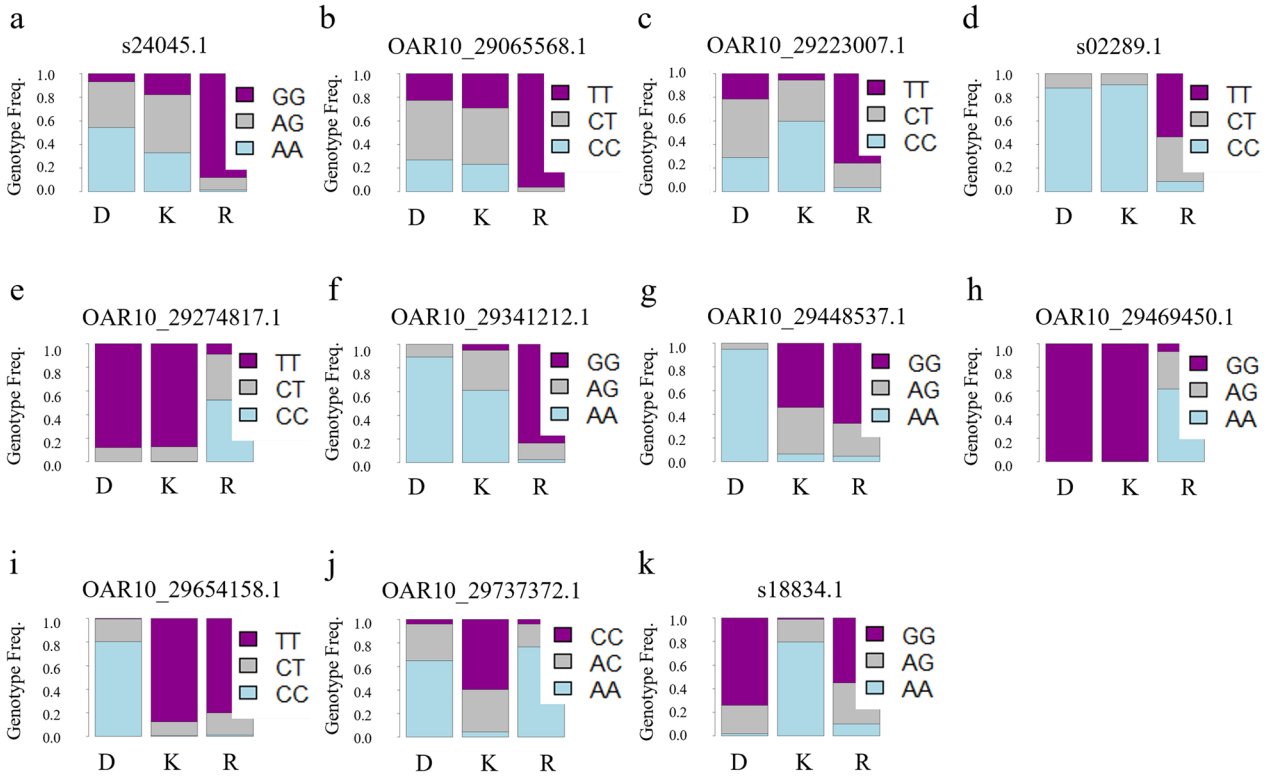

Fig. 7 SNPs with high F_{ST} in the region of the *FRY* and *RXFP2* genes. **a** Stacked bar chart representing genotype frequencies at SNP s24045.1; **b** Stacked bar chart representing genotype frequencies at SNP OAR10_29065568.1; **c** Stacked bar chart representing genotype frequencies at SNP OAR10_29223007.1; **d** Stacked bar chart representing genotype frequencies at SNP s02289.1; **e** Stacked bar chart representing genotype frequencies at SNP OAR10_29274817.1; **f** Stacked bar chart representing genotype frequencies at SNP OAR10_29341212.1; **g** Stacked bar chart representing genotype frequencies at SNP OAR10_29448537.1; **h** Stacked bar chart representing genotype frequencies at SNP OAR10_29469450.1; **i** Stacked bar chart representing genotype frequencies at SNP OAR10_29654158.1; **j** Stacked bar chart representing genotype frequencies at SNP OAR10_29737372.1; **k** Stacked bar chart representing genotype frequencies at SNP s18834.1. D: Dorper; K: Katahdin; R: Rambouillet

Table 8 Genes that were in common between F_{ST} comparisons

Breed (number of genes)	Genes within all F_{ST} analyses of a breed [reference]
Rambouillet (51)	*ADIRF* [79], *BMPR1A* [80], *CCSER1* [81], *CSF3* [82], *EEF1A1** [83], *EIF2S2** [84], *EVI2A* [85], *EVI2B* [85], *FAM25A*, *FRY** [86], *GSDMA* [87], *GSDMB*, *LDB3* [88], *LRRC3C*, *MED24*, *MMRN2*, *MRPS18C*, *MSL1*, *MTIF2*, *NF1* [89], *NR1D1* [90], *OMG* [91], *OPN4*, *OR10A2*, *OR10A5*, *OR2AG1* [92], *OR2D3*, *OR8S1*, *ORMDL3*, *PCED1B*, *PSMD3* [93], *RAB11FIP4* [94], *RASGEF1B*, *RPL10A* [95], *RPL6* [96], *RXFP2** [83], *SCAMP2*, *SLC38A1*, *SLC38A2*, *SNCG* [97], *SNORD124*, *THRA* [93], *TMA7*, *TNFSF18* [98], *TRNAC-ACA*, *TRNAC-GCA** [99], *TRNAG-UCC**, *TRNAS-GGA** [100], *TRNAW-CCA** [101], *UAP1*, *ZPBP2*
Katahdin (31)	*ADCY6* [102], *B3GALNT2*, *B3GLCT*, *CACNB3* [103], *CCNT1* [91], *DDX23*, *EEF1A1** [83], *EIF2S2** [84], *FBXL14*, *FRY** [86], *INPP4A*, *IRF2BP2*, *KANSL2* [104], *LALBA* [97], *MGAT4A*, *MRPL42* [105], *NUDT4* [105], *RND1*, *RXFP2** [83], *SNORA40*, *SOCS2* [106], *ST13*, *TARBP1*, *TEX49*, *TRNAC-GCA** [99], *TRNAE-CUC*, *TRNAG-UCC**, *TRNAS-GGA** [100], *TRNAW-CCA** [101], *UBE2N* [107], *UNC50*
Dorper (21)	*ANO6* [108], *ATXN7L3B*, *DIP2C* [81], *EEF1A1** [83], *EIF2S2** [84], *EPB41L4B* [109], *FRY** [86], *GRIN2B* [110], *MFF* [111], *PTPN3* [112], *RXFP2** [83], *SLIT2* [113], *SNORD31*, *TRNAC-GCA** [99], *TRNAG-CCC*, *TRNAG-UCC**, *TRNAH-AUG*, *TRNAS-GCU*, *TRNAS-GGA** [100], *TRNAW-CCA** [101], *ZMYND11* [114]

Genes that were present in all F_{ST} comparisons of a breed (e.g., for Rambouillet, genes that were present in both the Katahdin-Rambouillet and Rambouillet-Dorper comparisons). The list numbers of the relevant literature references for genes identified as candidate genes or within signatures of selection for production and/or health traits in sheep are in brackets. *indicates that the gene is present in more than one breed list. tRNA genes may be members of the same family located at multiple loci between breeds or analyses

Table 9 KEGG results for genes which were in common between both comparisons of Rambouillet or Katahdin

Breed	KEGG Mapper pathway (number of genes)	Enriched genes
Rambouillet	Olfactory transduction (5)	*OR2D3*, *OR10A5*, *OR2AG1*, *OR10A2*, *OR8S1*
	Ribosome (3)	*RPL10A*, *MRPS18C*, *RPL6*
	Cytokine-cytokine receptor interaction (3)	*CSF3*, *BMPR1A*, *TNFSF18*
	Coronavirus disease—COVID-19 (3)	*CSF3*, *RPL10A*, *RPL6*
Katahdin	Metabolic pathways (3)	*ADCY6*, *B3GALNT2*, *LALBA*

Genes present within the Rambouillet F_{ST} analyses (i.e. the Katahdin-Rambouillet and Rambouillet-Dorper F_{ST} comparisons) or the Katahdin F_{ST} analyses (i.e. the Katahdin-Rambouillet and Katahdin-Dorper F_{ST} comparisons) were queried through KEGG Mapper to identify pathways of interest. The Dorper sheep did not have KEGG mapper pathways enriched with three or more gene search terms

Table 10 Genes that were called within the ROH islands and pairwise F_{ST} comparison regions

Breed	Genes in common between ROH and F_{ST} regions
Katahdin	*ARID4B*, *ATMIN*, *B3GALNT2*, *GGPS1*, *GNG4*, *LYST*, *RBM34*, *TBCE*, *TOMM20*, *TRNAC-ACA*, *TRNAC-GCA*, *TRNAR-CCU*, *TRNAW-CCA*
Rambouillet	*ADGRA3*, *AMIGO2*, *ANO6*, *CCSER1*, *COL2A1*, *DCAF16*, *EIF2S3*, *FAM184B*, *GBA3*, *GRID2*, *HDAC7*, *KCNIP4*, *LAP3*, *LCORL*, *LRRK2*, *MED28*, *MFF*, *NCAPG*, *PCED1B*, *RAN*, *RPL10A*, *RPL37*, *SLC2A13*, *SLC38A1*, *SLC38A2*, *SLC38A4*, *SLIT2*, *SNCA*, *STIM2*, *TBC1D19*, *TMA7*, *TMEM106C*, *TRNAA-CGC*, *TRNAC-GCA*, *TRNAG-CCC*, *TRNAH-AUG*, *TRNAS-GGA*, *TRNAW-CCA*, *VDR*

The Katahdin genes are present in the Katahdin ROH islands as well as the F_{ST} region(s) between Katahdin-Rambouillet and/or Katahdin-Dorper, the Rambouillet genes are present in Rambouillet ROH islands and the F_{ST} region(s) between Katahdin-Rambouillet and/or Rambouillet-Dorper

broad conversation about the signatures of selection and breed diversity of worldwide sheep.

A landmark paper in sheep genetic diversity research was published in 2012 by Kijas and colleagues [115]. Their study analyzed population structure and signatures of selection of 2819 sheep sampled from 74 breeds, including 102 Rambouillet sheep. The inbreeding coefficient (F) and N_e reported for these Rambouillet sheep were 0.14 and 709, respectively. These statistics are quite different from those of the Rambouillet analyzed in the current study. While a larger number of Rambouillet (N = 745) were sampled in this paper, the average inbreeding coefficient (F_{ROH}) was found to be higher (0.169) and the current N_e was smaller (NeEstimator: 56.9, GONe: 111.8) than those reported previously [115]. It is possible that these differences result from differences between the F statistic and N_e calculation methods. A previous study found that two genotype-based inbreeding coefficients (F_{SNP} and F_{ROH}) had correlation coefficients ranging from 0.78 to 0.88, indicating that while related, the degree of inbreeding is not entirely comparable between subcategories of F statistics [116]. The results of the current study found that F and F_{ROH} showed different levels of inbreeding, with pairwise breed differences being

most significant in F_{ROH} estimates. The N_e reported in our study are estimated for current and historic generations while the N_e estimates reported previously by Kijas and colleagues [115] are not described in terms of generational distance and are therefore difficult to compare directly.

A previous study was conducted with genotype data from 4884 Katahdin sheep [72]. The 581 Katahdin sheep used in our study were also analyzed as part of this larger population, although the genotype data used between these publications were collected from separate platforms. In spite of the difference in sample sizes between these studies, the SNeP N_e estimations 13 generations ago were very similar (N_e of 172 or 178 estimated from 4884 or 581 sheep, respectively). This agreement between N_e estimations supports the validity of these estimates.

The N_e estimates reported between the SNeP, GONe, and NeEstimator approaches varied. Based on SNeP, Rambouillet had consistently larger N_e than Katahdin, and both breeds showed a decline in N_e between each subsequent generation. This ranking of the Rambouillet and Katahdin breeds was also observed in the historic N_e reported through GONe, although in more recent generations there were multiple reranking events in which the N_e of Katahdin exceeded that of Rambouillet, and in the current generation, both NeEstimator and GONe calculated a larger N_e for Katahdin than for Rambouillet. However, Katahdin had significantly higher average F and F_{ROH} compared to Rambouillet, which suggests comparatively less current breed diversity. While the largest and most conserved ROH was identified in Rambouillet, Katahdin had a greater proportion of ROH within the larger 6–12 Mb class. The presence of more ROH of greater length suggests recent inbreeding events in the history of the Katahdin breed. Taken together, these results support the importance of evaluating evidence from multiple inbreeding and N_e models, since relying on a single estimate might bias interpretation of genetic diversity.

The largest ROH Island was positioned between 32.8 and 48.0 Mb on chromosome 6 and encompassed 64 unique genes. This region was previously reported to be under selection in sheep and cattle [23, 117–120] and many of the associated genes were suggested as candidates for production traits. In cattle, the *ABCG2, IBSP*, *PIGY*, *PKD2,* and *MEPE* genes have been associated with yearling weight [121]; *NCAPG*, *LCORL,* and *LAP3* with body weights and calving ease [122]; and *ABCG2*, *LAP3*, *NCAPG*, *DCAF16,* and *LCORL* with milk total solid percentage [123]. Similarly in sheep, quantitative trait loci (QTL) associated with the *DCAF16*, *LAP3*, *LCORL*, *NCAPG, NAP1L5,* and *PPARGC1A* genes have been reported for somatic cell score or milk yield in dairy sheep [124]; *DCAF16, LCORL,* and *NCAPG* genes associated with body weight [125]; *SPP1* and *LAP3* genes associated with weight [126]; *LCORL* associated with meat productivity [127]; *SLIT2* and *ABCG2* genes associated with fat deposition and milk production, respectively [115, 128]; and genes *HERC3, HERC5, HERC6, IBSP,* and *SPP1* associated with parasite resistance [129]. In the current study, the ROH island identified in this region was highly conserved, with more than 94% of the Rambouillet animals having a ROH called. Rambouillet sheep are raised for their carcass and wool characteristics and, as a breed, are generally considered to be susceptible to parasite infection [130]. The ROH island in this region may be the result of past selection for weight or carcass traits and the homozygosity at genes relating to milk yield or parasite resistance may have occurred through hitchhiking [131].

Similarly, the ROH island identified on chromosome 3 contains genes that were previously described within signatures of selection in goats and cattle, including *ADAMTS20*, *GXYLT1*, *IRAK4*, *PRICKLE1*, *PUS7L*, *TWF1*, *YAF2*, and *ZCRB1* [132–134]. Expression of *ADAMTS20* has been reported to be upregulated in ewes with endometritis [135], and members of the *ADAMTS* family have been associated with divergent prolificacy of sheep and goats [136]. In goats, the *ADAMTS20* gene has also been associated with coat color [137]. The *IRAK4* gene encodes interleukin-1 receptor-associated kinase 4, a key regulator of innate immune signaling responses [138]. In addition, the *GXYLT1*, *PRICKLE1*, *YAF2*, and *ZCRB1* genes have been associated with resistance/susceptibility to *Mycobacterium avium* spp. *paratuberculosis* infection (responsible for Johne's disease) in Holstein cattle [139]. An association study in dogs identified *TWF1* as a candidate gene for deafness, and this gene has potential roles in hair-bundle development and melanocyte dendricity [140]. Both *IRAK4* and *TWF1* were implicated as genes of interest by Tajima's D analysis with two strains of Qinchuan cattle [133]. The occurrence of these genes within signatures of selection in other species and the associations with reproduction, coat color, and animal health suggest potential traits which may be under selection in the Rambouillet breed.

Genes that were previously reported in signatures of selection in worldwide sheep breeds were identified through F_{ST} analysis of the U.S. Rambouillet, Katahdin, and Dorper sheep in the current study. These genes include *NF1*, a negative regulator of the ras pathway (identified in the Katahdin-Rambouillet and Rambouillet-Dorper F_{ST} comparison), *OR2AG1* (identified in the Katahdin-Rambouillet and Rambouillet-Dorper F_{ST} comparison, and other olfactory receptor genes identified in all F_{ST} comparisons); and *RXFP2*, associated with horn

phenotypes in sheep (present in all F_{ST} comparisons), which were reported by Kijas et al. [115]. A number of the genes identified in both F_{ST} comparisons with Katahdin (Katahdin-Rambouillet and Katahdin-Dorper) have been associated with milk production or resistance and susceptibility for mastitis. These genes included *CCNT1*, previously associated with milk and cheese-making traits [91]; *LALBA*, associated with milk production traits [141]; *MRPL42*, a candidate gene for mastitis resistance [105]; and the gene suppressor of cytokine signaling 2 (*SOCS2*) associated with mastitis susceptibility [106, 142]. Genes present in both Rambouillet F_{ST} comparisons (Katahdin-Rambouillet and Rambouillet-Dorper) have been shown to be connected with wool growth, including: *FRY*, involved in wool development through a previous F_{ST} analysis [78]; *NR1D1*, found to be differentially expressed and differentially methylated at different stages of hair follicle development [90]; *TNFSF18*, a candidate gene for staple length and fiber diameter traits [98]; and *RXFP2*, which may be linked to hair follicle growth through the cAMP synthesis pathway [143]. The *NF1*, *EVI2A*, *EVI2B,* and *OMG* genes were also identified in both of the Rambouillet F_{ST} analyses and were previously associated with adaptive response to physical exhaustion and fat deposition [85].

Copies of the *EIF2S2* gene were identified in F_{ST} regions on chromosomes 13 and 25 in the Rambouillet-Dorper and Katahdin-Rambouillet F_{ST} analyses, respectively. The copy-number variant of *EIF2S2* on chromosome 25 is caused by retrotransposition (retroCNV), which results in an insertion of the *EIF2S2* retrogene into the 3′ UTR of *IRF2BP2*. This insertion has been found to be responsible for the "wooly" fleece phenotype of modern sheep [84, 144]. As a hair breed, the coats of Katahdin sheep are more similar to those of ancestral sheep breeds, in which an inner coat of fine wooly fibers lays below an outer coat comprised of hair fibers. Modern wooly sheep, such as the Rambouillet, lack the double coat of hair and ancestral sheep breeds, and instead possess a single coat comprised of wooly fibers of mostly uniform dimension [84]. It is unclear what is the biological mechanism that is responsible for Rambouillet having differentiation in the regions of both the retroCNV and the *EIF2S2* gene, but the high F_{ST} in the region on chromosome 25 is likely related to differences between hair and wool coat types.

The ROH island identified on chromosome 25 in the Katahdin sheep overlapped partly with the Katahdin-Rambouillet F_{ST} region containing the *IRF2BP2* gene. The F_{ST} region ranged from 6.6 to 7.0 Mb while the ROH island ranged from 6.9 to 7.8 Mb, creating an overall signature of selection encompassing a region from 6.6 to 7.8 Mb on chromosome 25. Three SNPs positioned at, respectively, 6.94 Mbp, 7.67 Mbp, and 7.73 Mbp, were identified by both the ROH and F_{ST} outlier analyses. This region contains the *TOMM20*, *RBM34*, *ARID4B*, *GGPS1*, *TBCE*, *B3GALNT2*, *ATMIN*, *GNG4*, and *LYST* genes, which were previously implicated in tail fat deposition, the ancestral-like coarse wool phenotype, coat color regulation, and response to *Brucella ovis* infection [145–149]. The identification of this signature of selection through both F_{ST} and ROH analyses, which encompasses genes that are important for both immune pathways and hair type, suggests a potential relationship between the genetic control of these traits in Katahdin sheep.

The ROH island identified on chromosome 23 in Katahdin sheep harbors genes compiled in KEGG immune pathways, including the *LAMA1* gene in the ECM-receptor interaction, toxoplasmosis, and viral myocarditis pathways; *TUBB6* in the phagosome, *Salmonella* infection and pathogenic *E. coli* infection pathways; and *PTPRM* in the adherens junction and cell adhesion molecules pathways. In a study with Scottish Blackface lambs, the *LAMA1* gene was found to be divergently expressed between lambs with low versus high fecal egg count phenotypes [150]. The *PTPRM* gene has been described to have a role in the telogen phase of hair follicle growth in Dorper sheep [151], and in studies with humans and biomedical models, it has been shown to be expressed by T cells and to have a dysregulated expression in patients with immune-mediated skin disease [152, 153]. While there is no current evidence for a specific immune role for *TUBB6* in sheep, transcription of *TUBB6* and other phagosome-associated genes were found to be down-regulated in the pituitary gland of nutrient-restricted ewes during late gestation [154], and enhanced transcription was identified during the late phase of *Eimeria bovis* infection in culture with host bovine cells [155].

The SNP with the highest F_{ST} score for the Katahdin-Rambouillet analysis was OAR2_231739122.1, which is positioned 935 bp downstream of the *CXCR2* gene. This gene encodes the principal membrane-bound chemokine receptor that is responsible for mediating neutrophil recruitment [156]. In sheep, *CXCR2* has been implicated in clinical mastitis and resistance to gastrointestinal nematode infection [157, 158]. A number of immune-related pathways were revealed through KEGG Mapper pathway analysis of the genes present within the Katahdin-Rambouillet F_{ST} regions. These pathways included leukocyte transendothelial migration (*AFDN* gene), B cell receptor signaling (*CD79B* gene), cytokine-cytokine receptor interactions (*CSF3, GH1, BMPR1A, GDF5, TNFSF18* genes), IL-17 signaling (*CSF3* gene), inflammatory mediator regulation of TRP channels (*ADCY6, NTRK1, TRPM8* genes), and NF-kappa B signaling (*BCL2L1, MAP3K14* genes), among others. The high level of F_{ST} associated with these genes and pathways may reflect differences

in allele frequencies contributing to the more robust immune response attributed to Katahdin sheep [33, 157].

Through KEGG Mapper pathway analysis for Katahdin-Dorper F_{ST} regions, pathways including chemokine signaling (*ADCY6*) and cytokine-cytokine receptor interaction (*TNFRSF13B*), NOD-like receptor signaling (*NLRP3* and *PANX1*), B cell receptor signaling (*CD81*), and phagosome pathways (*STX7*) were identified. These results are of particular interest in light of the divergent immune abilities of Dorper and Katahdin sheep [159, 160]. In addition, KEGG cytokine-cytokine receptor interaction pathways were identified in all F_{ST} comparisons, although the genes involved differed between analyses (see Additional file 20: Fig. S3). In the Katahdin-Rambouillet comparison, the *CSF3, GH1, BMPR1A, GDF5,* and *TNFSF18* genes were identified; in the Rambouillet-Dorper comparison, the *IL31RA, CSF3, IL17RA, GHR, IL6ST, BMPR1A,* and *TNFSF18* were identified; and in the Katahdin-Dorper comparison, only the *TNFRSF13B* was identified, as previously stated. About half of the pathway genes overlapped between the Rambouillet comparisons (*CSF3, BMPR1A,* and *TNFSF18*), suggesting that these F_{ST} regions were most strongly associated with differentiation of allele frequencies in Rambouillet sheep. There were no overlapping genes between the two Katahdin comparisons, which suggests that the greatest amount of differentiation between Katahdin and Rambouillet and between Katahdin and Dorper occur in different regions of the cytokine-cytokine receptor interaction pathways. These F_{ST} results may have a role in the overall immune response of Katahdin sheep compared to the Rambouillet and Dorper breeds that are susceptible to parasites.

The sheep used in this study were not phenotyped for coat color or the presence or absence of horns or scurs. However, Rambouillet are well-known for their light-colored coat [161], with variation being observed in different propensities for yellowing rather than pigment [162]. Dorper sheep have a white body and a black head and neck, while White Dorper sheep have an entirely white coat [163]. The Dorper analyzed in this study were from a flock founded by both Dorper and White Dorper sheep and showed a variety of color patterns, including some animals with black heads and/or black spotting and others with solid white coats. Katahdin sheep do not have a definite coat color and may have a variety of colors and patterns [164]. These breed-specific differences in coat color may have influenced the identification of signatures of selection that contain genes associated with coat pigment in the present study. Horn phenotypes are also expected to differ between these breeds. Many Katahdin producers prefer their sheep to be polled, although scurred and horned animals are permitted by breed standards [164]. The horn/polled phenotype in Rambouillet has been described as strictly sex-linked [165]. Rams in the African Dorper breed are variable and can have horns, scurs, or be polled, while females are scurred or polled [165]. These known differences may explain why multiple F_{ST} outlier SNPs were detected near the *RXFP2* gene in all pairwise comparisons of this study.

Conclusions

This study provides insights into the signatures of selection and genetic diversity of three popular U.S. sheep breeds. The results described here support previous reports on genes that underlie signatures of selection in sheep and provide additional insights into the biological differences between Rambouillet, Katahdin, and Dorper sheep. The results of the F_{ST} analyses indicated strong population differentiation associated with genes relevant to milk production in Katahdin sheep and wool growth in Rambouillet sheep. A large and highly conserved ROH island was identified in Rambouillet sheep that contained genes of known importance for growth and carcass traits, including *LCORL* and *NCAPG*. Signatures of selection were identified in genes relevant to ancestral versus wooly coat types (*IRF2BP2*, retroCNV *EIF2S2*) and horn/polled phenotypes (*RXFP2*) in addition to many genes involved in immune-related pathways. These findings likely have relevance to the variations in physical appearance and parasite resistance ability of these breeds. Further analysis of these F_{ST} and ROH signatures of selection may provide greater insight into the selection pressures being exerted on these important breeds in the U.S. sheep industry.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12711-024-00905-7.

Additional file 1: Figure S1. Principal component analysis with Rambouillet, Katahdin, and Dorper sheep.

Additional file 2: Table S1. Effective population size (N_e) for Dorper sheep. The rate of change (m) between each pair of generations is reported, with the overall rate of change reported in italics.

Additional file 3: Table S2. Percentage of ROH and mean size of ROH by breed and by Mbp class.

Additional file 4: Figure S2. Stacked bar plot for the proportion of ROH called within each ROH size class by breed.

Additional file 5: Table S3. ROH islands identified in Dorper sheep. ROH islands are defined by SNPs called within a ROH for more than 50% of Dorper sheep. *indicates these regions have overlap between Dorper and Rambouillet ROH calls.

Additional file 6: Table S4. Results of KEGG Mapper Pathway analysis of genes located within Katahdin ROH islands.

Additional file 7: Table S5. Results of KEGG Mapper Pathway analysis of genes located within Rambouillet ROH islands.

Additional file 8: Table S6. Results of STRING protein–protein interactions for genes contained within F_{ST} outlier regions and within-breed ROH islands. This excel file Table S6 contains seven sheets i.e. Sheet 1 entitled "Katahdin-Rambouillet F_{ST} STRING interactions", Sheet 2 entitled "Rambouillet-Dorper F_{ST} STRING interactions", Sheet 3 entitled "Katahdin-Dorper F_{ST} STRING interactions", Sheet 4 entitled "Katahdin ROH STRING interactions, chromosome 23", Sheet 5 entitled "Katahdin ROH STRING interactions, chromosome 25", Sheet 6 entitled "Rambouillet ROH STRING interactions, chromosome 3" and Sheet 7 entitled "Rambouillet ROH STRING interactions, chromosome 6".

Additional file 9: Table S7. Summary of F_{ST} regions by chromosome and breed comparison.

Additional file 10: Table S8. List of SNPs identified from both Katahdin-Rambouillet F_{ST} analysis and Katahdin or Rambouillet ROH islands.

Additional file 11: Table S9. Results of GO enrichment analysis of genes located within Katahdin-Rambouillet F_{ST} regions.

Additional file 12: Table S10. KEGG Mapper pathway results for query of Katahdin-Rambouillet genes against the *Homo sapiens* reference database.

Additional file 13: Table S11. KEGG Mapper pathway results for query of Katahdin-Dorper genes against the *Homo sapiens* reference database.

Additional file 14: Table S12. KEGG Mapper pathway results for query of Rambouillet-Dorper genes against the *Homo sapiens* reference database.

Additional file 15: Table S13. Most significantly enriched GO biological process function terms from Dorper ROH islands. Gene names were searched against the *Bos taurus* reference database. Each term in italics represents the most specific subclass with related parent terms directly below.

Additional file 16: Table S14. Significantly enriched GO cellular component terms from Dorper ROH islands. Gene names were searched against the *Bos taurus* reference database. Each term in italics represents the most specific subclass with related parent terms directly below.

Additional file 17: Table S15. Significantly enriched GO cellular molecular function terms from Dorper ROH islands. Gene names were searched against the *Bos taurus* reference database. Each term in italics represents the most specific subclass with related parent terms directly below.

Additional file 18: Table S16. KEGG Mapper pathway enrichment for genes located within Dorper ROH islands. KEGG query was made against the *Homo sapiens* reference.

Additional file 19: Table S17. Comparison of BayeScan and Plink F_{ST} results for each pairwise analysis. This excel Table S17 contains three sheets i.e. Sheet 1 entitled "Comparison of BayeScan and Plink F_{ST} results for the Katahdin-Rambouillet pairwise analysis", Sheet 2 entitled "Comparison of BayeScan and Plink F_{ST} results for the Rambouillet-Dorper pairwise analysis", and Sheet 3 entitled "Comparison of BayeScan and Plink F_{ST} results for the Katahdin-Dorper pairwise analysis".

Additional file 20: Figure S3. Genes belonging to cytokine-cytokine receptor interaction pathways identified from breed F_{ST} analyses. Genes identified in Katahdin-Dorper F_{ST} analysis are in blue, genes identified from the Katahdin-Rambouillet F_{ST} analysis are in orange, and genes identified from the Rambouillet-Dorper F_{ST} analysis are in gray. Pathway figure is modified from Kanehisa Laboratories (KEGG).

Acknowledgements
The authors would like to thank the sheep producers who contributed to the success of this study and acknowledge AgResearch and the Animal Genomics team for access to the AgResearch Sheep Genomics 60K SNP chip. The findings and conclusion in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Author contributions
BMM, JMB, DRN, and RML conceived the study; GMB, JWT, JMB, RML, DRN, JLMM, CSS, WCS, RRR, and BMM contributed to animal sampling and/or data collection efforts; GMB and BMM performed data analyses and manuscript writing; all authors contributed to manuscript editing. All authors read and approved the final manuscript.

Funding
This research was funded by the Organic Agriculture Research and Extension Initiative grant no. 2016-51300-25723, USDA-NIFA-IDA1566, Texas A&M AgriLife Research, the Idaho Department of Commerce and Agriculture via Idaho Global Entrepreneurial Mission grant #004727, and Hatch-Multistate "Increased Efficiency of Sheep Production" project accession no. 1025808 from the USDA National Institute of Food and Agriculture.

Availability of data and materials
All relevant data are included in the manuscript and its additional files. The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
The animal study was reviewed and approved by the Agricultural Research Service (ARS) Institutional Animal Care and Use Committee (IACUC) (17103), the Virginia Tech IACUC (15–133) the Ohio Agricultural Research and Development Center IACUC (2017A00000042), Texas A & M University IACUC (2020-19A), and NDSU IACUC (20210012). Written informed consent was obtained from the owners for the participation of their animals in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Animal, Veterinary and Food Science, University of Idaho, Moscow, ID, USA. [2]Texas A&M AgriLife Extension, Texas A&M University, San Angelo, TX, USA. [3]USDA, ARS, Dale Bumpers Small Farms Research Center, Booneville, AR, USA. [4]Department of Animal Science, University of Nebraska–Lincoln, Lincoln, NE, USA. [5]School of Animal Sciences, Virginia Tech, Blacksburg, VA, USA. [6]Round Mountain Consulting, Fayetteville, AR, USA. [7]Hettinger Research Extension Center, North Dakota State University, Hettinger, ND, USA. [8]Department of Animal Science, University of Wyoming, Laramie, WY, USA.

Received: 19 July 2023 Accepted: 23 April 2024
Published online: 30 July 2024

References

1. Toro MA, Caballero A. Characterization and conservation of genetic diversity in subdivided populations. Philos Trans R Soc Lond B Biol Sci. 2005;360:1367–78.
2. Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, et al. The crucial role of genome-wide genetic variation in conservation. Proc Natl Acad Sci USA. 2021;118: e2104642118.
3. Eusebi PG, Martinez A, Cortes O. Genomic tools for effective conservation of livestock breed diversity. Diversity (Basel). 2019;12:8.
4. Saravanan KA, Panigrahi M, Kumar H, Bhushan B, Dutt T, Mishra BP. Genome-wide analysis of genetic diversity and selection signatures in three Indian sheep breeds. Livest Sci. 2021;243:104367.
5. Prezeworski M, Coop G, Wall JD. The signature of positive selection on standing genetic variation. Evolution. 2005;59:2312–23.
6. Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. Genetics. 2000;155:1405–13.
7. Eydivandi S, Roudbar MA, Karimi MO, Sahana G. Genomic scans for selective sweeps through haplotype homozygosity and allelic fixation

in 14 indigenous sheep breeds from Middle East and South Asia. Sci Rep. 2021;11:2834.
8. Andolfatto P. Adaptive hitchhiking effects on genome variability. Curr Opin Genet Dev. 2001;11:635–41.
9. Zhang Z, Sui Z, Zhang J, Li Q, Zhang Y, Wang C, et al. Identification of signatures of selection for litter size and pubertal initiation in two sheep populations. Animals (Basel). 2022;12:2520.
10. Peripolli E, Munari DP, Silva MV, Lima AL, Irgang R, Baldi F. Runs of homozygosity: current knowledge and applications in livestock. Anim Genet. 2017;48:255–71.
11. Nei M. Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA. 1973;70:3321–3.
12. Almeida OA, Moreira GC, Rezende FM, Boschiero C, de Oliveira PJ, Ibelli AM, et al. Identification of selection signatures involved in performance traits in a paternal broiler line. BMC Genomics. 2019;20:449.
13. Smaragdov MG, Kudinov AA. Assessing the power of principal components and wright's fixation index analyzes applied to reveal the genome-wide genetic differences between herds of Holstein cows. BMC Genet. 2020;21:47.
14. 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 (Edinb). 2016;116:255–64.
15. Edea Z, Dadi H, Dessie T, Kim KS. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. Genes Genomics. 2019;41:973–81.
16. Ahmad SF, Mehrotra A, Charles S, Ganai NA. Analysis of selection signatures reveals important insights into the adaptability of high-altitude Indian sheep breed Changthangi. Gene. 2021;799: 145809.
17. Kominakis A, Tarsani E, Hager-Theodorides AL, Mastranestasis I, Gkelia D, Hadjigeorgiou I. Genetic differentiation of mainland-island sheep of Greece: implications for identifying candidate genes for long-term local adaptation. PLoS One. 2021;16:e0257461.
18. McRae KM, McEwan JC, Dodds KG, Gemmell NJ. Signatures of selection in sheep bred for resistance or susceptibility to gastrointestinal nematodes. BMC Genomics. 2014;15:637.
19. Moioli B, Pilla F, Ciani E. Signatures of selection identify loci associated with fat tail in sheep. J Anim Sci. 2015;93:4660–9.
20. Cesarani A, Sechi T, Gaspa G, Usai MG, Sorbolini S, Macciotta NP, et al. Investigation of genetic diversity and selection signatures between Sarda and Sardinian Ancestral black, two related sheep breeds with evident morphological differences. Small Rumin Res. 2019;177:68–75.
21. Li X, Yang JI, Shen M, Xie XL, Liu GJ, Xu YX, et al. Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. Nat Commun. 2020;11:2815.
22. Gutiérrez-Gil B, Esteban-Blanco C, Wiener P, Chitneedi PK, Suarez-Vega A, Arranz JJ. High-resolution analysis of selection sweeps identified between fine-wool Merino and coarse-wool Churra sheep breeds. Genet Sel Evol. 2017;49:81.
23. Ruiz-Larrañaga O, Langa J, Rendo F, Manzano C, Iriondo M, Estonba A. Genomic selection signatures in sheep from the Western Pyrenees. Genet Sel Evol. 2018;50:9.
24. Dickinson WF, Lush JL. Inbreeding and the genetic history of the Rambouillet sheep in America. J Hered. 1933;24:19–33.
25. Murphy TW, Stewart WC, Notter DR, Mousel MR, Lewis GS, Taylor JB. Evaluation of Rambouillet, Polypay, and Romanov-White Dorper× Rambouillet ewes mated to terminal sires in an extensive rangeland production system: body weight and wool characteristics. J Anim Sci. 2019;97:1568–77.
26. National Agricultural Statistics Service (NASS), Agricultural Statistics Board, United States Department of Agriculture (USDA). Sheep and Goats. 2023. https://usda.library.cornell.edu/concern/publications/000000018?locale=en. Accessed 7 June 2023.
27. Wildeus S. Hair sheep genetic resources and their contribution to diversified small ruminant production in the United States. J Anim Sci. 1997;75:630–40.
28. Burke JM, Miller JE. Relative resistance of Dorper crossbred ewes to gastrointestinal nematode infection compared with St Croix and Katahdin ewes in the southeastern United States. Vet Parasitol. 2002;109:265.
29. Burke JM, Miller JE. Relative resistance to gastrointestinal nematode parasites in Dorper, Katahdin, and St. Croix lambs under conditions encountered in the southeastern region of the United States. Small Rumin Res. 2004;54:43–51.
30. Vanimisetti HB, Greiner SP, Zajac AM, Notter DR. Performance of hair sheep composite breeds: resistance of lambs to *Haemonchus contortus*. J Anim Sci. 2004;82:595–604.
31. Snowder GD, Duckett SK. Evaluation of the South African Dorper as a terminal sire breed for growth, carcass, and palatability characteristics. J Anim Sci. 2003;81:368–75.
32. de Waal HO, Combrinck WJ. The development of the Dorper, its nutrition and a perspective of the grazing ruminant on veld. Small Rumin Res. 2000;36:103–17.
33. Riley DG, Thorne JW, Redden RR, Herring AD, Waldron DF. Comparative performance of Dorper and Rambouillet ewes on West Texas range. Small Rumin Res. 2020;185:106065.
34. Becker GM, Burke JM, Lewis RM, Miller JE, Morgan JL, Rosen BD, et al. Variants within genes EDIL3 and ADGRB3 are associated with divergent fecal egg counts in Katahdin sheep at weaning. Front Genet. 2022;13:817319.
35. Becker GM, Shira KA, Woods JL, Khilji SF, Schauer CS, Webb BT, et al. Angular limb deformity associated with TSPAN18, NRG3 and NOVA2 in Rambouillet rams. Sci Rep. 2023;13:16059.
36. Thorne JW, Redden R, Bowdridge SA, Becker GM, Stegemiller MR, Murdoch BM. Genome-wide analysis of sheep artificially or naturally infected with gastrointestinal nematodes. Genes (Basel). 2023;14:1342.
37. Sambrook J, Fritsch ER, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press; 1989.
38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
39. 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.
40. Meyermans R, Gorssen W, Buys N, Janssens S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. BMC Genomics. 2020;21:94.
41. Linck E, Battey CJ. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Mol Ecol Resour. 2019;19:639–47.
42. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
43. Kassambara A, Kassambara MA. Package 'ggpubr'. R package version 0.1; 2020. https://rpkgs.datanovia.com/ggpubr. Accessed 9 June 2023.
44. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2021.
45. Bannasch D, Famula T, Donner J, Anderson H, Honkanen L, Batcher K, et al. The effect of inbreeding, body size and morphology on health in dog breeds. Canine Med Genet. 2021;8:12.
46. Biscarini F, Cozzi P, Gaspa G, Marras G. detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes. CRAN (The Comprehensive R Archive Network); 2018. https://orca.cardiff.ac.uk/id/eprint/108906/. Accessed 9 June 2023.
47. Van Hecke T. Power study of anova versus Kruskal-Wallis test. J Stat Manag Syst. 2012;15:241–7.
48. Dunn OJ. Multiple comparisons using rank sums. Technometrics. 1964;6:241–52.
49. 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.
50. Hill WG, Robertson A. Linkage disequilibrium in finite populations. Theor Appl Genet. 1968;38:226–31.
51. Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour. 2014;14:209–14.
52. Coombs JA, Letcher BH, Nislow KH. GONe: software for estimating effective population size in species with generational overlap. Mol Ecol Resour. 2012;12:160–3.
53. Pedersen T. patchwork: The composer of plots. 2020. https://cran.r-project.org/web/packages/patchwork/index.html. Accessed 9 June 2023.
54. Yin L. CMplot: Circle Manhattan plot. R package version 4.1.0 2022. https://CRAN.R-project.org/package=CMplot. Accessed 9 June 2023.

55. Royston P. Approximating the Shapiro-Wilk W-test for non-normality. Stat Comput. 1992;2:117–9.
56. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc. 1952;47:583–621.
57. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38:1358–70.
58. Bhatia G, Patterson N, Sankararaman S, Price AL. Estimating and interpreting F_{ST}: the impact of rare variants. Genome Res. 2013;23:1514–21.
59. Foll M, Gaggiotti O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics. 2008;180:977–93.
60. Gosselin T, Lamothe M, Devloo-Delva F, Grewe P. radiator: RADseq data exploration, manipulation and visualization using R. R package version. 2020. https://thierrygosselin.github.io/radiator/ Accessed 7 Mar 2024.
61. Rangwala SH, Kuznetsov A, Ananiev V, Asztalos A, Borodin E, Evgeniev V, et al. Accessing NCBI data using the NCBI sequence viewer and genome data viewer (GDV). Genome Res. 2021;31:159–69.
62. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the National Center for Biotechnology Information in 2023. Nucleic Acids Res. 2023;51:D29-38.
63. Davenport KM, Bickhart DM, Worley K, Murali SC, Salavati M, Clark EL, et al. An improved ovine reference genome assembly to facilitate in-depth functional annotation of the sheep genome. Gigascience. 2022;11:giab096.
64. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000;25:25–9.
65. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51:D587–92.
66. Kanehisa M, Sato Y, Kawashima M. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci. 2022;31:47–53.
67. Gene Ontology Consortium, Carbon S, Douglass E, Good BM, Unni DR, Harris NL, et al. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021;49:D325–34.
68. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447–52.
69. Blackburn HD, Paiva SR, Wildeus S, Getz W, Waldron D, Stobart R, et al. Genetic structure and diversity among sheep breeds in the United States: identification of the major gene pools. J Anim Sci. 2011;89:2336–48.
70. Lei Z, Sun W, Guo T, Li J, Zhu S, Lu Z, et al. Genome-wide selective signatures reveal candidate genes associated with hair follicle development and wool shedding in sheep. Genes (Basel). 2021;12:1924.
71. Davenport KM, Hiemke C, McKay SD, Thorne JW, Lewis RM, Taylor T, et al. Genetic structure and admixture in sheep from terminal breeds in the United States. Anim Genet. 2020;51:284–91.
72. Becker GM, Burke JM, Lewis RM, Miller JE, Morgan JL, Rosen BD, et al. Inbreeding and effective population size of United States Katahdin sheep. In: Proceedings of the 12th World Congress on Genetics Applied to Livestock Production (WCGALP): 3–8 July 2022; Rotterdam. 2022.
73. Molotsi AH, Taylor JF, Cloete SW, Muchadeyi F, Decker JE, Whitacre LK, et al. Genetic diversity and population structure of South African smallholder farmer sheep breeds determined using the OvineSNP50 beadchip. Trop Anim Health Prod. 2017;49:1771–7.
74. Dzomba EF, Chimonyo M, Pierneef R, Muchadeyi FC. Runs of homozygosity analysis of South African sheep breeds from various production systems investigated using OvineSNP50k data. BMC Genomics. 2021;22:7.
75. Nosrati M, Nanaei HA, Javanmard A, Esmailizadeh A. The pattern of runs of homozygosity and genomic inbreeding in world-wide sheep populations. Genomics. 2021;113:1407–15.
76. Adepoju D. Estimating the effective population size of Swedish native cattle: understanding the demographic trajectories of indigenous Swedish cattle breeds. Master thesis. Uppsala: Swedish University of Agricultural Sciences. 2022.
77. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. PLoS One. 2010;5:e13996.
78. Pérez O'Brien AM, Utsunomiya YT, Mészáros G, Bickhart DM, Liu GE, Van Tassell CP, et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. Genet Sel Evol. 2014;46:19.
79. Bakhtiarizadeh MR, Salehi A, Alamouti AA, Abdollahi-Arpanahi R, Salami SA. Deep transcriptome analysis using RNA-Seq suggests novel insights into molecular aspects of fat-tail metabolism in sheep. Sci Rep. 2019;9:9203.
80. Bahire SV, Rajput PK, Kumar V, Kumar D, Kataria M, Kumar S. Quantitative expression of mRNA encoding BMP/SMAD signalling genes in the ovaries of Booroola carrier and non-carrier GMM sheep. Reprod Domest Anim. 2019;54:1375–83.
81. Serrano M, Ramón M, Calvo JH, Jiménez MÁ, Freire F, Vázquez JM, et al. Genome-wide association studies for sperm traits in Assaf sheep breed. Animal. 2021;15:100065.
82. Estrada-Reyes ZM, Tsukahara Y, Amadeu RR, Goetsch AL, Gipson TA, Sahlu T, et al. Signatures of selection for resistance to *Haemonchus contortus* in sheep and goats. BMC Genomics. 2019;20:735.
83. Wiedemar N, Drögemüller C. A 1.8-kb insertion in the 3′-UTR of RXFP2 is associated with polledness in sheep. Anim Genet. 2015;46:457–61.
84. Demars J, Cano M, Drouilhet L, Plisson-Petit F, Bardou P, Fabre S, et al. Genome-wide identification of the mutation underlying fleece variation and discriminating ancestral hairy species from modern wooly sheep. Mol Biol Evol. 2017;34:1722–9.
85. Ahbara AM, Musa HH, Robert C, Abebe A, Al-Jumaili AS, Kebede A, et al. Natural adaptation and human selection of northeast African sheep genomes. Genomics. 2022;114:110448.
86. Zhang L, Mousel MR, Wu X, Michal JJ, Zhou X, Ding B, et al. Genome-wide genetic diversity and differentially selected regions among Suffolk, Rambouillet, Columbia, Polypay, and Targhee sheep. PLoS One. 2013;8:e65942.
87. Li C, Ren L, Wang Y, Zhong J, Huang L, Lin Y, et al. Construction and functional analysis of expression vector and miRNA interference vectors of Gsdma of Tibetan sheep. J Anim Sci. 2016;94:165–6.
88. Ma X, Liu A, Tian S. A meta-analysis of mRNA expression profiling studies in sheep with different FecB genotypes. Anim Genet. 2023;54:225–38.
89. Wei C, Wang H, Liu G, Zhao F, Kijas JW, Ma Y, et al. Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. Sci Rep. 2016;6:26770.
90. Tian Y, Yang X, Du J, Zeng W, Wu W, Di J, et al. Differential methylation and transcriptome integration analysis identified differential methylation annotation genes and functional research related to hair follicle development in sheep. Front Genet. 2021;12:735827.
91. Marina H, Pelayo R, Gutiérrez-Gil B, Suárez-Vega A, Esteban-Blanco C, Reverter A, et al. Low-density SNP panel for efficient imputation and genomic selection of milk production and technological traits in dairy sheep. J Dairy Sci. 2022;105:8199–217.
92. Yaro M, Munyard KA, Morgan E, Allcock RJ, 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.
93. Wang H, Zhang L, Cao J, Wu M, Ma X, Liu Z, et al. Genome-wide specific selection in three domestic sheep breeds. PLoS One. 2015;10:e0128688.
94. Estrada-Reyes ZM, Ogunade IM, Pech-Cervantes AA, Terrill TH. Copy number variant-based genome wide association study reveals immune-related genes associated with parasite resistance in a heritage sheep breed from the United States. Parasite Immunol. 2022;44:e12943.
95. Pasandideh M, Rahimi-Mianji G, Gholizadeh M. A genome scan for quantitative trait loci affecting average daily gain and Kleiber ratio in Baluchi Sheep. J Genet. 2018;97:493–503.
96. Chen L, Liu K, Zhao Z, Blair HT, Zhang P, Li D, et al. Identification of sheep ovary genes potentially associated with off-season reproduction. J Genet Genomics. 2012;39:181–90.
97. Hao Z, Zhou H, Hickford JG, Gong H, Wang J, Hu J, et al. Transcriptome profile analysis of mammary gland tissue from two breeds of lactating sheep. Genes (Basel). 2019;10:781.

98. Zhao H, Guo T, Lu Z, Liu J, Zhu S, Qiao G, et al. Genome-wide association studies detects candidate genes for wool traits by re-sequencing in Chinese fine-wool sheep. BMC Genomics. 2021;22:127.
99. Sánchez-Ramos R, Trujano-Chavez MZ, Gallegos-Sánchez J, Becerril-Pérez CM, Cadena-Villegas S, Cortez-Romero C. Detection of candidate genes associated with fecundity through genome-wide selection signatures of Katahdin ewes. Animals (Basel). 2023;13:272.
100. Krivoruchko AY, Yatsyk OA, Safaryan EY. Candidate genes for productivity identified by genome-wide association study with indicators of class in the Russian meat merino sheep breed. Vavilovskii Zhurnal Genet Selektsii. 2020;24:836–43.
101. de Souza TC, de Souza TC, Rovadoscki GA, Coutinho LL, Mourao GB, de Camargo GM, et al. Genome-wide association for plasma urea concentration in sheep. Livest Sci. 2021;248:104483.
102. Bakhtiarizadeh MR, Alamouti AA. RNA-Seq based genetic variant discovery provides new insights into controlling fat deposition in the tail of sheep. Sci Rep. 2020;10:13525.
103. Zhang DY, Zhang XX, Li GZ, Li XL, Zhang YK, Zhao Y, et al. Transcriptome analysis of long noncoding RNAs ribonucleic acids from the livers of Hu sheep with different residual feed intake. Animal. 2021;15: 100098.
104. Zhao H, Zhu S, Guo T, Han M, Chen B, Qiao G, et al. Whole-genome re-sequencing association study on yearling wool traits in Chinese fine-wool sheep. J Anim Sci. 2021;99: skab210.
105. Banos G, Bramis G, Bush SJ, Clark EL, McCulloch ME, Smith J, et al. The genomic architecture of mastitis resistance in dairy sheep. BMC Genomics. 2017;18:624.
106. Rupp R, Senin P, Sarry J, Allain C, Tasca C, Ligat L, et al. A point mutation in suppressor of cytokine signalling 2 (Socs2) increases the susceptibility to inflammation of the mammary gland while associated with higher body weight and size and higher milk production in a sheep model. PLoS Genet. 2015;11:e1005629.
107. Liu Z, Ji Z, Wang G, Chao T, Hou L, Wang J. Genome-wide analysis reveals signatures of selection for important traits in domestic sheep from different ecoregions. BMC Genomics. 2016;17:863.
108. Chitneedi PK, Suárez-Vega A, Martínez-Valladares M, Arranz JJ, Gutiérrez-Gil B. Exploring the mechanisms of resistance to *Teladorsagia circumcincta* infection in sheep through transcriptome analysis of abomasal mucosa and abomasal lymph nodes. Vet Res. 2018;49:39.
109. Abied A, Bagadi A, Bordbar F, Pu Y, Augustino SM, Xue X, et al. Genomic diversity, population structure, and signature of selection in five Chinese native sheep breeds adapted to extreme environments. Genes (Basel). 2020;11:494.
110. Estrada-Reyes ZM, Rae DO, Mateescu RG. Genome-wide scan reveals important additive and non-additive genetic effects associated with resistance to *Haemonchus contortus* in Florida Native sheep. Int J Parasitol. 2021;51:535–43.
111. Wen Y, Li S, Zhao F, Wang J, Liu X, Hu J, et al. Changes in the mitochondrial dynamics and functions together with the mRNA/miRNA network in the heart tissue contribute to hypoxia adaptation in Tibetan sheep. Animals (Basel). 2022;12:583.
112. Mohamadipoor Saadatabadi L, Mohammadabadi M, Amiri Ghanatsaman Z, Babenko O, Stavetska R, Kalashnik O, et al. Signature selection analysis reveals candidate genes associated with production traits in Iranian sheep breeds. BMC Vet Res. 2021;17:369.
113. Mastrangelo S, Bahbahani H, Moioli B, Ahbara A, Al Abri M, Almathen F, et al. Novel and known signals of selection for fat deposition in domestic sheep breeds from Africa and Eurasia. PLoS One. 2019;14:e0209632.
114. Li LL, Ma SK, Peng W, Fang YG, Duo HR, Fu HY, et al. Genetic diversity and population structure of Tibetan sheep breeds determined by whole genome resequencing. Trop Anim Health Prod. 2021;53:174.
115. Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, San Cristobal M, et al. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 2012;10: e1001258.
116. Rodríguez-Ramilo ST, Elsen JM, Legarra A. Inbreeding and effective population size in French dairy sheep: comparison between genomic and pedigree estimates. J Dairy Sci. 2019;102:4227–37.
117. Manunza A, Cardoso TF, Noce A, Martínez A, Pons A, Bermejo LA, et al. Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. Sci Rep. 2016;6:27296.
118. Sorbolini S, Gaspa G, Steri R, Dimauro C, Cellesi M, Stella A, et al. Use of canonical discriminant analysis to study signatures of selection in cattle. Genet Sel Evol. 2016;48:58.
119. Selli A, Ventura RV, Fonseca PA, Buzanskas ME, Andrietta LT, Balieiro JC, et al. Detection and visualization of heterozygosity-rich regions and runs of homozygosity in worldwide sheep populations. Animals (Basel). 2021;11:2696.
120. Liu Z, Bai C, Shi L, He Y, Hu M, Sun H, et al. Detection of selection signatures in South African Mutton Merino sheep using whole-genome sequencing data. Anim Genet. 2022;53:224–9.
121. Naserkheil M, Mehrban H, Lee D, Park MN. Genome-wide association study for carcass primal cut yields using single-step Bayesian approach in Hanwoo cattle. Front Genet. 2021;12: 752424.
122. 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.
123. Buaban S, Lengnudum K, Boonkum W, Phakdeedindan P. Genome-wide association study on milk production and somatic cell score for Thai dairy cattle using weighted single-step approach with random regression test-day model. J Dairy Sci. 2022;105:468–94.
124. Mohammadi H, Farahani AH, Moradi MH, Mastrangelo S, Di Gerlando R, Sardina MT, et al. Weighted single-step genome-wide association study uncovers known and novel candidate genomic regions for milk production traits and somatic cell score in Valle del Belice dairy sheep. Animals (Basel). 2022;12:1155.
125. Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C. Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. Genet Sel Evol. 2015;47:66.
126. La Y, Zhang X, Li F, Zhang D, Li C, Mo F, et al. Molecular characterization and expression of SPP1, LAP3 and LCORL and their association with growth traits in sheep. Genes (Basel). 2019;10:616.
127. Zlobin AS, Nikulin PS, Volkova NA, Zinovieva NA, Iolchiev BS, Bagirov VA, et al. Multivariate analysis identifies eight novel loci associated with meat productivity traits in sheep. Genes (Basel). 2021;12:367.
128. Rochus CM, Tortereau F, Plisson-Petit F, Restoux G, Moreno-Romieux C, Tosser-Klopp G, et al. Revealing the selection history of adaptive loci using genome-wide scans for selection: an example from domestic sheep. BMC Genomics. 2018;19:71.
129. Al Kalaldeh 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.
130. Thorne JW, Bowdridge SA, Murdoch BM, Redden RR. Response of Rambouillet lambs to an artificial gastrointestinal nematode infection. Animals (Basel). 2022;12:1199.
131. Cutter AD, Payseur BA. Genomic signatures of selection at linked sites: unifying the disparity among species. Nat Rev Genet. 2013;14:262–74.
132. Bertolini F, Servin B, Talenti A, Rochat E, Kim ES, Oget C, et al. Signatures of selection and environmental adaptation across the goat genome post-domestication. Genet Sel Evol. 2018;50:57.
133. Mei C, Wang H, Liao Q, Khan R, Raza SH, Zhao C, et al. Genome-wide analysis reveals the effects of artificial selection on production and meat quality traits in Qinchuan cattle. Genomics. 2019;111:1201–8.
134. Mastrangelo S, Di Gerlando R, Sardina MT, Sutera AM, Moscarelli A, Tolone M, et al. Genome-wide patterns of homozygosity reveal the conservation status in five Italian goat populations. Animals (Basel). 2021;11:1510.
135. Safhi FA, Ateya A. New insights into polymorphisms in candidate genes associated with incidence of postparturient endometritis in Ossimi Sheep (*Ovis aries*). Agriculture. 2023;13:2273.
136. Hernández-Delgado P, Felix-Portillo M, Martínez-Quintana JA. ADAMTS proteases: Importance in animal reproduction. Genes (Basel). 2023;14:1181.
137. Dong Y, Zhang X, Xie M, Arefnezhad B, Wang Z, Wang W, et al. Reference genome of wild goat (*Capra aegagrus*) and sequencing of goat breeds provide insight into genic basis of goat domestication. BMC Genomics. 2015;16:431.
138. De S, Karim F, Kiessu E, Cushing L, Lin LL, Ghandil P, et al. Mechanism of dysfunction of human variants of the IRAK4 kinase and a role for

its kinase activity in interleukin-1 receptor signaling. J Biol Chem. 2018;293:15208–20.
139. 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.
140. Kluth S, Distl O. Congenital sensorineural deafness in Dalmatian dogs associated with quantitative trait loci. PLoS One. 2013;8:e80642.
141. Garcia-Gamez E, Gutierrez-Gil B, Sahana G, Sanchez JP, Bayon Y, Arranz JJ. GWA analysis for milk production traits in dairy sheep and genetic support for a QTN influencing milk protein percentage in the LALBA gene. PLoS One. 2012;7:e47782.
142. Li R, Ma Y, Jiang L. Research progress of dairy sheep milk genes. Agriculture. 2022;12:169.
143. Megdiche S, Mastrangelo S, Ben Hamouda M, Lenstra JA, Ciani E. A combined multi-cohort approach reveals novel and known genome-wide selection signatures for wool traits in Merino and Merino-derived sheep breeds. Front Genet. 2019;10:1025.
144. Schrider DR, Navarro FC, Galante PA, Parmigiani RB, Camargo AA, Hahn MW, et al. Gene copy-number polymorphism caused by retrotransposition in humans. PLoS Genet. 2013;9: e1003242.
145. Bakhtiarizadeh MR, Salami SA. Identification and expression analysis of long noncoding RNAs in fat-tail of sheep breeds. G3 (Bethesda). 2019;9:1263–76.
146. Lukic B, Curik I, Drzaic I, Galić V, Shihabi M, Vostry L, et al. Genomic signatures of selection, local adaptation and production type characterisation of East Adriatic sheep breeds. J Anim Sci Biotechnol. 2023;14:142.
147. Wang J, Hua G, Cai G, Ma Y, Yang X, Zhang L, et al. Genome-wide DNA methylation and transcriptome analyses reveal the key gene for wool type variation in sheep. J Anim Sci Biotechnol. 2023;14:88.
148. de Paula Antunes JM, Allendorf SD, Appolinário CM, Peres MG, Vicente AF, Cagnini DQ, et al. Microarray analysis of gene expression in rams experimentally-infected with the virulent strain of *Brucella ovis*. J Biotechnol Biomater. 2015;5:203.
149. Fan R, Xie J, Bai J, Wang H, Tian X, Bai R, et al. Skin transcriptome profiles associated with coat color in sheep. BMC Genomics. 2013;14:389.
150. McRae KM, Good B, Hanrahan JP, McCabe MS, Cormican P, Sweeney T, et al. Transcriptional profiling of the ovine abomasal lymph node reveals a role for timing of the immune response in gastrointestinal nematode resistance. Vet Parasitol. 2016;224:96–108.
151. Sun H, Meng K, Wang Y, Wang Y, Yuan X, Li X. LncRNAs regulate the cyclic growth and development of hair follicles in Dorper sheep. Front Vet Sci. 2023;10:1186294.
152. Rioux G, Turgeon F, Le-Bel G, Grenier C, Guérin SL, Pouliot R. Gene profiling of a 3D psoriatic skin model enriched in T cells: downregulation of PTPRM promotes keratinocyte proliferation through excessive ERK1/2 signaling. Cells. 2022;11:2904.
153. Arimura Y, Yagi J. Comprehensive expression profiles of genes for protein tyrosine phosphatases in immune cells. Sci Signal. 2010;3: rs1.
154. Liu S, Lu H, Mao S, Zhang Z, Zhu W, Cheng J, et al. Undernutrition-induced substance metabolism and energy production disorders affected the structure and function of the pituitary gland in a pregnant sheep model. Front Nutr. 2023;10:1251936.
155. Taubert A, Wimmers K, Ponsuksili S, Jimenez CA, Zahner H, Hermosilla C. Microarray-based transcriptional profiling of *Eimeria bovis*-infected bovine endothelial host cells. Vet Res. 2010;41:70.
156. Semple BD, Bye N, Ziebell JM, Morganti-Kossmann MC. Deficiency of the chemokine receptor CXCR2 attenuates neutrophil infiltration and cortical damage following closed head injury. Neurobiol Dis. 2010;40:394–403.
157. Rafeie F, Abdoli R, Hossein-Zadeh NG, Talebi R, Szmatoła T. Interaction networks and pathway analysis of genetic resistance to gastrointestinal nematodes in sheep. Trop Anim Health Prod. 2023;55:34.
158. Li T, Gao J, Zhao X, Ma Y. Digital gene expression analyses of mammary glands from meat ewes naturally infected with clinical mastitis. R Soc Open Sci. 2019;6:181604.
159. Silva MV, Sonstegard TS, Hanotte O, Mugambi JM, Garcia JF, Nagda S, et al. Identification of quantitative trait loci affecting resistance to gastrointestinal parasites in a double backcross population of Red Maasai and Dorper sheep. Anim Genet. 2012;43:63–71.
160. Palomo-Couoh JG, Aguilar-Caballero AJ, Torres-Acosta JF, González-Garduño R. Comparing the phenotypic susceptibility of Pelibuey and Katahdin female lambs against natural gastrointestinal nematode infections under hot humid tropical conditions. Parasitol Res. 2017;116:1627–36.
161. Monty DE Jr, Kelley LM, Rice WR. Acclimatization of St.-Croix, Karakul and Rambouillet sheep to intense and dry summer heat. Small Rumin Res. 1991;4:379–92.
162. Cameron BA, Stobart RH. The yellowing propensity of Rambouillet wool. Sheep Goat Res J. 2008;23:11–4.
163. Wanjala G, Astuti PK, Bagi Z, Kichamu N, Strausz P, Kusza S. Assessing the genomics structure of Dorper and white Dorper variants, and Dorper populations in South Africa and Hungary. Biology. 2023;12:386.
164. Barton B. Katahdin sheep an excellent meat breed. Ranch Rural Living. 2014;96:29–31.
165. Simon R, Drögemüller C, Lühken G. The complex and diverse genetic architecture of the absence of horns (polledness) in domestic ruminants, including goats and sheep. Genes (Basel). 2022;13:832.

Publisher's Note