Animal 18 (2024) 101322

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

# Genome-wide association studies for milk production traits in two autochthonous Aosta cattle breeds

F. Bernini [a,*], E. Mancin [b], C. Sartori [b], R. Mantovani [b], M. Vevey [c], V. Blanchet [c], A. Bagnato [a], M.G. Strillacci [a]

[a] Department of Veterinary Medicine and Animal Science, Università degli Studi di Milano, Via dell'Università 6, 26900 Lodi, Italy
[b] Department of Agronomy, Food, Natural Resources, Animals and Environment, Università degli Studi di Padova, Viale dell'Università 16, 35020 Legnaro, Italy
[c] Associazione Nazionale Bovini di Razza Valdostana, Frazione Favret 5, 11020 Gressan, Italy

ARTICLE INFO

*Article history:*
Received 23 April 2024
Revised 29 August 2024
Accepted 2 September 2024
Available online 12 September 2024

*Keywords:*
Cattle
Genomics
Local breeds
Productive traits
Quantitative Traits Loci

ABSTRACT

Genome-wide association studies (**GWASs**) are used to identify quantitative trait loci for phenotypic traits of interest. The use of multilocus mixed models allows to correct for population stratification and account for long-range linkage disequilibrium. In this study, GWASs were conducted to identify the genetic bases of milk production (milk yield, protein and fat composition, and yield) in two autochthonous dual-purpose cattle breeds from the Aosta Valley. Using either the breeding values or the deregressed proofs, common significative single nucleotide polymorphisms have been identified for milk yield, protein percentage, and fat percentage. Two major quantitative trait loci regions have been identified on the chromosomes 5 and 14 for the fat percentage, harbouring the *MGST1, CYHR1, VPS28*, and *CPSF1* genes. For the protein percentage, a candidate region has been identified on BTA 6; in this region, the *CSN1S1*, *CSN2, HSTN, CSN3*, and *RUFY3* genes are annotated. Most of the identified genes have already been associated with milk composition in other studies on cosmopolitan and local cattle. These results show that the genes involved in milk composition quantitative traits in the Aosta cattle are common also in other cattle breeds and they can be further investigated with the use of whole genome sequencing data.


## Implications

This genome-wide association study is a first step towards the development of genomic selection models for the autochthonous Aosta cattle breeds. Implementing genomic selection in these cattle represents an important step forward in the efficiency of selection, maintenance of their genetic variability, and most importantly, preservation of their hardiness. Despite further analyses are still needed to develop genomic selection for these breeds, this study highlighted that Aosta cattle breeds share quantitative trait loci with other cosmopolitan breeds.

## Introduction

Genome-wide association studies have been used for years to identify quantitative trait loci for phenotypic traits of interest in cosmopolitan cattle breeds (Chen et al., 2022; Raven et al., 2014). At present, the increased availability of genotypes collected directly on female cattle, opens new possibilities for genome-wide association studies to disclose quantitative trait loci for low heritability and innovative traits (Pedrosa et al., 2023; Strillacci et al., 2023) and to perform genome-wide association studies on small autochthonous cattle, as already occurred in many local breeds (Korkuć et al., 2021; Mancin et al., 2022). In cattle, genome-wide association studies have been used for decades to identify quantitative trait loci for complex traits (Glantz et al., 2012) and more recently for innovative traits, e.g., heat tolerance, methane emissions, and feed efficiency (Manzanilla-Pech et al., 2021; Nguyen et al., 2017). To date, most genome-wide association studies have been carried out within single breeds, however, multibreed genome-wide association studies may lead to increased power and precision (van den Berg et al., 2016). Multibreed genome-wide association studies can enhance the statistical power, reduce the likelihood of false-positive associations, and improve the mapping resolution of genetic variants (Bouwman et al., 2018). This approach captures a broader spectrum of genetic diversity, allowing for the identification of both universal and breed-specific genetic markers associated with economically important traits such as milk production.

However, in dairy cattle, only a fraction of individuals in a population are genotyped, and not all genotyped animals possess the phenotypes. Therefore, the most straightforward approach appeared to use Estimated Breeding Value (**EBV**) as pseudo-

* Corresponding author.
*E-mail address:* francesca.bernini@unimi.it (F. Bernini).

https://doi.org/10.1016/j.animal.2024.101322
1751-7311/

phenotypes to include animals with genotypes but no phenotypes in genome-wide association studies. In fact, EBVs are already adjusted for environmental factors and readily available from routine evaluations.

Nonetheless, due to the random nature of EBVs, they tend to shrink towards zero when their reliability declines. This can lead to spurious associations between Single Nucleotide Polymorphism (**SNP**) and EBVs, increasing the likelihood of type-I errors (Sahana et al., 2023). In a small, heterogeneous population such as the Aosta breeds, this may result in two main drawbacks: (i) the effects of SNPs would be underestimated due to lower accuracy compared to more cosmopolitan breeds, and (ii) different cohorts of animals might experience varying degrees of reliability values for EBVs, leading to different levels of shrinkage. To address this problem, the most common approach is to deregress the EBV by their accuracy and/or remove redundant information such as parent average effects, obtaining the Deregressed Proofs (**DRPs**) (Garrick et al., 2009).

Aosta cattle have been recently classified into two breeds: the Aosta Red Pied (**ARP**) and the Aosta Black Pied - Chestnut with its subgroup of Chestnuts with Heréns ascendant (**ABCH**); the latter ones are two strains of the same population with different evolutionary history and only recently grouped as a unique breed by the Herd book based on their genomic similarity (Strillacci et al., 2020). These two breeds have different selection programmes: the ARP is selected for meat and milk production, while the ABCH is also selected for combativity. Like many other local cattle, Aosta breeds are particularly important for their own region, not only for their production of milk and meat but also for the cultural value and the maintenance of the mountain landscapes and environment (Strillacci et al., 2020).

The milk of Aosta cattle is almost entirely used to produce the Protected Designation of Origin Fontina cheese. These products own a disciplinary rule that specifies that the production, processing, and preparation process must take place in a specific geographical location of the production zone. In the case of Fontina cheese, the disciplinary requires that all the milk used for the cheese production is obtained by Aosta cattle, and the production and ageing processes can only take place in the Aosta Valley.

The selection process of Aosta cattle breeds is based on performance testing for meat traits, and sires of cows (young bulls) and sires of bulls (proven bulls) are also tested for milk traits (Pagnacco et al., 1989). The same authors proposed a mating plan to maintain as much as possible genetic variability in the population while addressing the relaxed selection programme, which has been adopted in the population for decades. A routine milk data collection is performed every 4–5 weeks by the expert technicians of the Italian Farmers Associations and involves all the farms registered to the National Breeders Association for Aosta cattle breeds, representing the majority of the Aosta cattle breeders. The estimated breeding values for productive traits are then calculated twice per year for the entire population using a repeatability animal model accounting for pedigree information, similar to the one described in the pilot study of Mazza et al. (2016). The implementation of genomic selection would represent an important step forward in making the efficiency of selection for milk and meat traits higher, maintaining the genetic variability, and, even more importantly, keeping the hardiness of the Aosta cattle. Recently, many females have been genotyped with SNP arrays thanks to the funding of the European Agricultural Fund for Rural Development, and this information can now be used to identify quantitative trait loci for traits of interest and to apply the genomic estimation of EBV. The aim of this study was to provide a first insight into the genomic identification of quantitative trait loci for milk production traits, i.e. milk yield - **MY**, protein yield - **PY**, fat yield - **FY**, protein content - **PP**, and fat content – **FP**, and annotating genes in the regions identified with SNP markers.

## Material and methods

### *Sampling and genotyping*

For this study, 4 247 female genotypes were provided by the National Breeders Association for Aosta cattle breeds. Sample distribution for each Aosta breed was as follows: 1 361 - ABCH and 2 886 - ARP. The initial dataset consisted of 89 762 SNP markers, obtained with the GGP Bovine 100 K SNP chip (GeneSeek®) by Neogen. These SNPs were mapped according to the ARS-UCD1.2 bovine reference genome (GCA_002263795.2), on the bovine autosomes.

The samples used for the analysis had a call rate higher than 95%. The SNPs were subject to quality control, and only the ones with (i) call rate $\geq$ 0.95, (ii) minor allele frequency > 0.01, and (iii) Hardy-Weinberg equilibrium $P$ > 1e-6 were kept. The pruned dataset was composed of 78 194 SNPs.

### *Statistical analysis for obtaining the estimated breeding values and deregression*

Variance components for MY, PY, FY, PP, and FP were obtained using the repeatability test day model currently used in routine genetic evaluations for productive traits. Variances were estimated on datasets of 42 716 individuals with records and 65 081 in the pedigree for ARP, and 22 799 with records and 35 914 in the pedigree for ABCH provided by the National Breeders Association for Aosta cattle breeds. The model included the fixed effects of herd, lactation number, gestation class, age at parity class within lactation, and month at parity class within lactation. The last two effects were covaried by the days in milk expressed as third-order Legendre polynomials. The herd-test day, the permanent environment, and the additive genetic component were included as random effects. The analyses were run under a Bayesian framework using a Gibbs sampling algorithm implemented within the software GIBBS3f90 of the BLUPF90 software family (Misztal et al., 2014). The EBVs were thus obtained from the analysis as individual additive genetic values. For 2 228 animals (521 ABCH and 1 707 ARP) part of the initial dataset, the DRPs were calculated according to Garrick et al. (2009). Only animals with reliability greater than 0.25 and only informative animals (with a reliability higher than the sum of the reliability of the two parents divided by four) were kept.

### *Genome-wide association analysis*

The Mixed Model GWA analysis was performed with the software SNP & Variation Suite v8.9.1 by Golden Helix ®, and the data of ARP and ABCH were used together. The Efficient Mixed Model Association eXpedited – EMMAX algorithm was used considering the additive genetic model. The identity–by–state matrix was included to account for the relatedness of the subjects sampled, and the breed was considered as fixed effect (Kang et al., 2010). After the analysis, the False Discovery Rate and Bonferroni correction thresholds were set at 5% genome-wide to correct for multiple testing.

### *Gene annotation*

For all the significant SNPs, over the 5% false discovery rate threshold, the rsID was assigned based on the SNP position using the Ensembl Database (McLaren et al., 2016). Once the rsID was

obtained, the gene annotation was performed using the Variant Effect Predictor - VEP tool by Ensembl (Hubbard et al., 2002).

## Results

The additive genetic variance and heritability of target traits (Table 1) are greater in ABCH than in ARP for all the traits considered. Target traits showed a moderate heritability, ranging from 0.12 and 0.13 or FY and FP in ARP, to 0.28 for MY in ABCH. Also, the average EBV and DRP values and SD were different between the two breeds (Table 1). In particular, the variability of MY, FY, and PY was smaller in the ARP with respect to the ABCH for both EBV and DRP. The results of the genome-wide association studies are shown in the Manhattan plots of Fig. 1 for the EBV and DRP of MY, PP, and FP. Significant SNPs associated to quantitative trait loci were found in several chromosomes and are reported in Tables 2–4 with their genomic classification and gene annotation.

When considering the Bonferroni threshold, a total of (i) 22 and 14 SNPs were significantly associated with EBV_FP and DRP_FP, respectively (Table 2), (ii) 21 and 30 SNPs were significantly associated with EBV_PP and DRP_PP, respectively (Table 3), and (iii) only two SNPs were significantly associated with EBV_MY (Table 4). For FY and PY, no significant SNPs have been found, either using the EBV or DRP, probably due to a lower heritability compared to the FP and PP. The Manhattan plots for these two traits are reported in Supplementary Figure S1.

## Discussion

The genetic parameters in Table 1 show a greater genetic variability for ABCH compared to the ARP. This may be explained as this group, recently arranged but administratively considered as the same breed, includes both the Aosta Black Pied and the Aosta Chestnut strains (Strillacci et al., 2020). Heritability values in ARP were previously estimated by Mazza et al. (2016) and in the ABCH by Sartori et al. (2020). The ARP values of 0.198, 0.132, and 0.169 were respectively estimated for MY, FY, and PY, while for the same traits, they were 0.227, 0.129 and 0.167 in the ABCH.

### *Fat percentage trait*

Six quantitative trait loci regions were significantly associated with FP. The region identified on chromosome 5, at about 93 Mbp, is defined by 21 SNPs, with rs211210569 and rs210744919 being the rsID numbers with the highest significance (Fig. 1, Table 2). This region harbours the microsomal glutathione S-transferase 1 - *MGST1* gene, which is involved in glutathione transport (GO:0034635), cellular detoxification processes (GO:0098869), and cellular response to lipid hydroperoxide (GO:0071449), an oxygenated product of polyunsaturated fatty acids, suggesting a potential role in lipid metabolism (Jayawardana et al., 2023). *MGST1* is a quantitative trait loci in numerous association studies performed on different breeds (e.g., Holstein, Braunvieh, Fleckvieh, Montbéliarde, and Normande) exploring the genetic basis of milk composition, especially for fat content (Sanchez et al., 2017; Tribout et al., 2020). Research in this field has aimed to decipher the complex genetic factors influencing milk composition, considering the economic and nutritional importance of milk and dairy products. Understanding the genetic determinants, including the potential contribution of *MGST1*, could have implications for livestock breeding programmes and the dairy industry, ultimately influencing milk quality and its nutritional value.

Recently, Korkuć et al. (2023) conducted multiple genome-wide association studies on Whole Genome Sequencing data for milk production traits in German Black Pied cattle. They identified significative SNPs for FP in the *MGST1* gene, speculating that this gene might contribute to milk fat via the regulation of energy and/or fatty acids to produce milk fat in the mammary gland. Cruz et al. (2019) identified two important regions for milk fatty acids groups, one on chromosome 5 harbouring the *MGST1* gene that was confirmed also when fitting the effect of the diacylglycerol O-acyltransferase 1 - *DGAT1* gene, and one on the chromosome 14 on the cysteine and histidine–rich 1- *CYHR1* gene, which was also identified in our study. Integrating genome–wide and RNA sequencing information, Littlejohn et al. (2016), suggested a role for *MGST1* as a detoxification enzyme whose impact on milk lipid synthesis or secretion is still unknown.

The quantitative trait loci region identified on chromosome 14 (Table 2) harbours eight genes, previously reported in association with milk traits. In particular, the *CYHR1* gene has been associated with milk fat yield and content in numerous studies considering different breeds (Oliveira et al., 2019; Pedrosa et al., 2021). Nevertheless, its role is still unclear due to the possible effect of the nearby DGAT1 gene that is well known to be a locus affecting milk fat content in cattle (Kühn et al., 2004). In our study, the linkage disequilibrium squared correlation statistics suggests no linkage between each of the eight genes reported in Table 2 and the DGAT1, as shown in Supplementary Figure S2.

A meta-analysis on Holstein cattle from different countries (i.e., Australia, Canada, China, France, Germany, Ireland, and Italy) (Bakhshalizadeh et al., 2021) reported two significant SNPs for FP also found in our study, the rs17870736 and rs134432442. They are mapping within the *VPS28* subunit of ESCRT-I (*VPS28*) and the cleavage and polyadenylation specific factor 1 (*CPSF1*) genes, respectively, and in linkage with one SNP (rs109968515, about 2.5 Kb apart from our significant rs137727465) that annotated close to the *CYHR1* gene.

The rs134432442 SNP is a missense variant (ACC/ATC codon) causing a change of the Threonine (amino acid position: 403) with

**Table 1**
Average (SD) values for each trait of the additive genetic variance ($\sigma^2_a$), heritability ($h^2$), EBV and DRP, for the Aosta cattle breeds.

| Item | ABCH | | | | ARP | | | |
|---|---|---|---|---|---|---|---|---|
| | $\sigma^2_a$ | $h^2$ | EBV | DRP | $\sigma^2_a$ | $h^2$ | EBV | DRP |
| n* | 291 060 | 291 060 | 1 361 | 521 | 718 966 | 718 966 | 2 886 | 1 707 |
| MY | 9.72 (0.42) | 0.28 (0.01) | −4.81 (340.42) | 1.68 (1.75) | 8.32 (0.29) | 0.22 (0.01) | −12.23 (274.32) | 3.01 (1.32) |
| FP | 0.06 (0.002) | 0.16 (0.01) | 0.03 (0.14) | 0.52 (0.23) | 0.04 (0.001) | 0.13 (0.003) | 0.04 (0.14) | 0.44 (0.20) |
| PP | 0.02 (0.001) | 0.27 (0.01) | 0.02 (0.1) | 0.05 (0.15) | 0.02 (0.000) | 0.25 (0.01) | 0.01 (0.11) | 0.52 (0.15) |
| FY | 0.01 (0.001) | 0.20 (0.01) | 0.56 (11.03) | 0.06 (0.06) | 0.01 (0.000) | 0.12 (0.005) | 1.26 (9.13) | 0.09 (0.05) |
| PY | 0.01 (0.000) | 0.25 (0.01) | 0.33 (10.5) | 0.06 (0.05) | 0.01 (0.000) | 0.18 (0.01) | 0.04 (7.85) | 0.10 (0.04) |

Abbreviations: ABCH=Aosta Black Pied-Chestnut, ARP=Aosta Red Pied, MY=milk yield, FP=fat percentage, PP=protein percentage, FY=fat yield, PP=protein yield, EBV=estimated breeding value, DRP=deregressed proof.
* n refers to the number of observations used for the evaluation of $\sigma^2_a$ and $h^2$, and the number of samples analysed for each breed for each trait EBV and DRP.

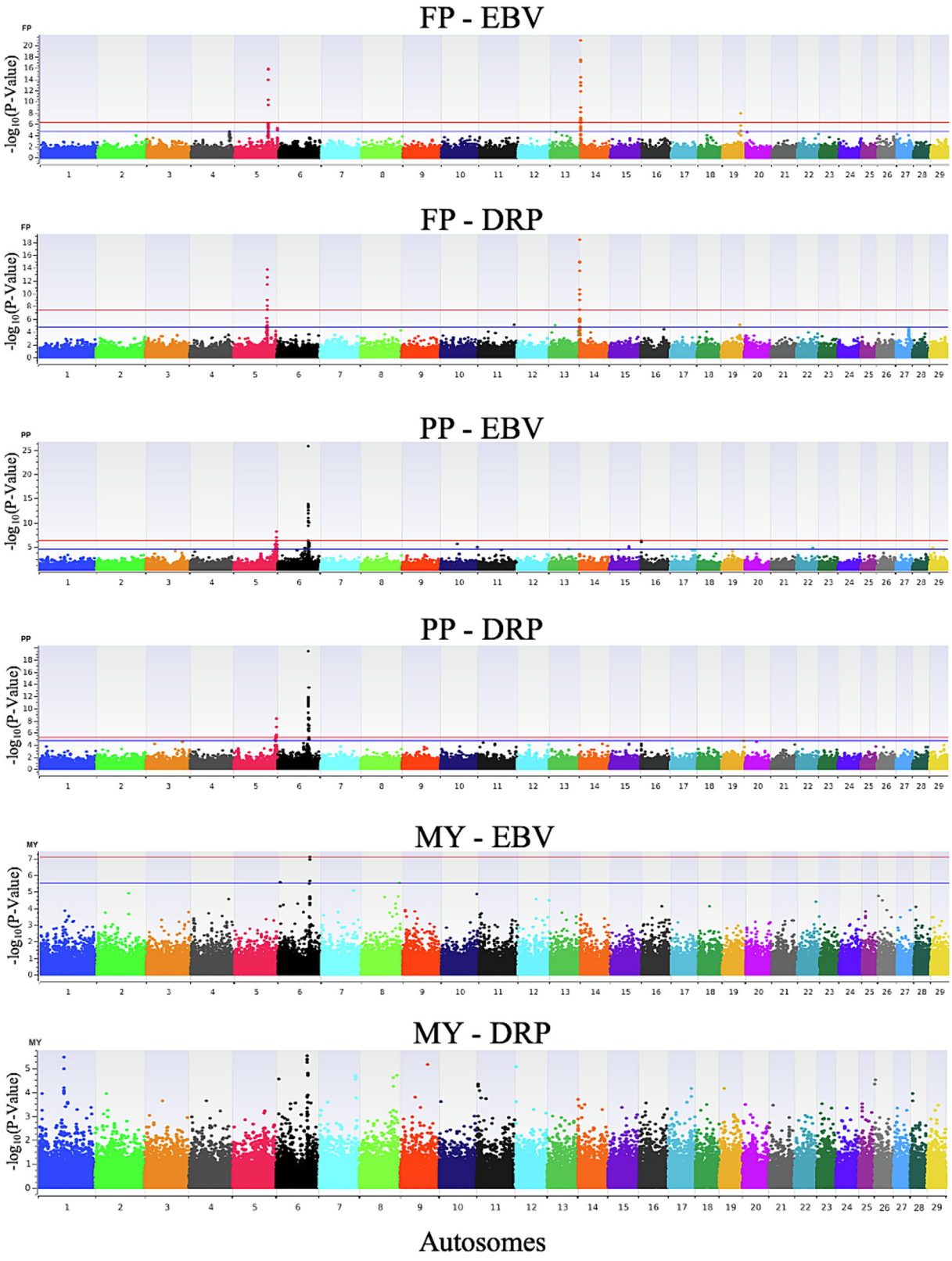


**Fig. 1.** Manhattan plots of the GWAS result for the FP, PP and for MY in the Aosta cattle breeds. Red and blue lines represent the Bonferroni and false discovery rate thresholds (both set at 5% genome-wide). Abbreviations: MY=milk yield, FP=fat percentage, PP=protein percentage, GWASs=genome wide association studies; EBV=estimated breeding value, DRP=deregressed proof.

an Isoleucine. Other previous studies found on chromosome 14 the same significant quantitative trait loci region harbouring the *CYHR1* and *VPS28* genes to explain the variability of fat content and fatty acids in milk in the studied populations (Iung et al., 2019; Jiang et al., 2019). On the same chromosome, the present study found other interesting SNPs, as follows:

**Table 2**
List of SNPs above the Bonferroni (underlined) and false discovery rate 0.05 thresholds for both EBVs and DRP for FP trait, identified in the Aosta cattle breeds. The Table reports the name of the SNP markers, the rsID number, the chromosome, the position in base pairs (bp), the *P*-values for the EBV and DPR associated with the marker, the gene in which the SNP lays, and the position of the marker respect to the gene.

| Marker | RS_SNP_ID | Chr | Position (bp) | *P*-Value EBV* | *P*-Value DRP* | Gene | Position |
|---|---|---|---|---|---|---|---|
| Hapmap47387-BTA-72195 | rs41591555 | 4 | 106144190 | 2.18E-05 | | | Intergenic |
| BovineHD0500026451 | rs135438063 | 5 | 92753124 | 2.94E-06 | | | Intergenic |
| BTA-37834-no-rs | rs109957658 | 5 | 92758201 | 5.77E-06 | 1.43E-05 | | Intergenic |
| BovineHD0500026553 | rs110579160 | 5 | 93109175 | 1.96E-05 | 7.06E-07 | | Intergenic |
| BovineHD0500026624 | rs134155693 | 5 | 93378265 | 2.41E-06 | | | Intergenic |
| BovineHD0500026635 | rs133160309 | 5 | 93413676 | 5.59E-06 | | | Intergenic |
| Hapmap32415-BTA-74556 | rs41602750 | 5 | 93441439 | 3.04E-06 | | | Intergenic |
| DB-335-seq-rs109307833 | rs109307833 | 5 | 93450091 | 6.44E-07 | | | Intergenic |
| BovineHD0500026649 | rs132674836 | 5 | 93470430 | 2.70E-06 | 3.84E-08 | | Intergenic |
| BovineHD0500026655 | rs133517677 | 5 | 93503991 | 3.30E-06 | 3.44E-06 | *MGST1* | Intronic |
| BovineHD0500026662 | rs134637616 | 5 | 93515983 | 5.52E-11 | 1.09E-09 | *MGST1* | Intronic |
| DB-337-seq-rs211210569 | rs211210569 | 5 | 93516066 | 1.57E-16 | 2.97E-13 | *MGST1* | Intronic |
| BovineHD0500026664 | rs137705840 | 5 | 93517967 | 2.29E-05 | | *MGST1* | Intronic |
| DB-339-seq-rs210744919 | rs210744919 | 5 | 93520138 | 2.01E-16 | 1.98E-14 | *MGST1* | Intronic |
| DB-340-seq-rs208014256 | rs208014256 | 5 | 93520616 | 4.12E-06 | | *MGST1* | 5′ UTR variant |
| chr5_93950333 | rs209210458 | 5 | 93520661 | 1.23E-06 | 8.89E-06 | *MGST1* | 5′ UTR variant |
| chr5_93950346 | rs210155966 | 5 | 93520674 | 3.13E-06 | | *MGST1* | 5′ UTR variant |
| BovineHD0500026666 | rs133918820 | 5 | 93521394 | 4.12E-06 | | | Intergenic |
| BovineHD0500026668 | rs135807129 | 5 | 93524134 | 3.90E-10 | 9.88E-09 | | Intergenic |
| DB-341-seq-rs209288972 | rs209288972 | 5 | 93525079 | 1.39E-14 | 3.73E-12 | | Intergenic |
| BovineHD0500034417 | rs109812511 | 5 | 117310460 | 1.51E-05 | | | Intergenic |
| BovineHD0500034461 | rs134294234 | 5 | 117440580 | 5.90E-06 | | | Intergenic |
| ARS-BFGL-NGS-23468 | rs109206555 | 11 | 100011096 | | 8.28E-06 | *C11H9orf50* | Intronic |
| ARS-BFGL-NGS-57332 | rs109338243 | 13 | 16659292 | | 9.72E-06 | | Intergenic |
| 14-1322168-C-A-rs208813903 | rs208813903 | 14 | 160524 | | 1.66E-05 | *OR10AG83* | Missense variant (C/A) |
| BovineHD1400000152 | rs110508680 | 14 | 255765 | 1.43E-09 | 3.74E-08 | | Intergenic |
| ARS-BFGL-NGS-57820 | rs109146371 | 14 | 465742 | 3.76E-07 | 1.04E-06 | | Intergenic |
| Chr14_1653693 | rs110984572 | 14 | 468124 | 1.51E-12 | 1.07E-09 | | Intergenic |
| BovineHD1400000204 | rs137727465 | 14 | 487527 | 4.71E-14 | 2.41E-11 | *CYHR1* | Intronic |
| BovineHD1400000206 | rs137472016 | 14 | 494621 | 1.60E-13 | 1.27E-10 | | Intergenic |
| ARS-BFGL-NGS-94706 | rs17870736 | 14 | 511247 | 8.42E-18 | 1.52E-15 | *VPS28* | Intronic |
| Chr14_1699016 | rs136784996 | 14 | 513203 | 4.00E-18 | 3.43E-14 | | Intergenic |
| UFL-rs134432442 | rs134432442 | 14 | 550784 | 5.44E-15 | 1.20E-15 | *CPSF1* | Missense variant (C/T) |
| Chr14_1757935 | rs211309638 | 14 | 572120 | 1.25E-21 | 3.89E-19 | | Intergenic |
| BovineHD1400000239 | rs133299034 | 14 | 663029 | 2.65E-06 | | *MROH1* | Intronic |
| BovineHD1400000241 | rs110966735 | 14 | 669738 | 3.67E-07 | | *MROH1* | Intronic |
| BovineHD1400000246 | rs137787931 | 14 | 688317 | 1.48E-05 | | *MROH1* | Intronic |
| Hapmap52798-ss46526455 | rs41256919 | 14 | 731230 | 5.81E-07 | 2.99E-06 | *MAF1* | Synonymous variant (T/C) |
| BovineHD1400000256 | rs110929299 | 14 | 751534 | 5.39E-07 | 1.89E-06 | *MAF1* | Intronic |
| BovineHD1400000271 | rs136792973 | 14 | 810116 | 1.85E-07 | 9.10E-07 | *GPAA1* | Intronic |
| UA-IFASA-6878 | rs41629750 | 14 | 810863 | 9.37E-08 | | | Intergenic |
| BovineHD1400000282 | rs136051530 | 14 | 859251 | 8.04E-07 | 2.51E-06 | *PLEC* | Intronic |
| BovineHD1400000287 | rs109662548 | 14 | 883732 | 6.65E-06 | | *PLEC* | Intronic |
| BovineHD1400000288 | rs135270011 | 14 | 891340 | 8.81E-09 | 1.32E-06 | *PLEC* | Synonymous variant (T/C) |
| BovineHD1900014321 | rs41922195 | 19 | 50607587 | 1.40E-08 | 7.33E-06 | *CSNK1D* | Intronic |
| BovineHD1900014337 | rs41922153 | 19 | 50666822 | 1.99E-06 | 9.00E-06 | *CCDC57* | Intronic |
| BovineHD1900014340 | rs135528222 | 19 | 50674342 | 1.67E-05 | | *CCDC57* | Intronic |

Abbreviations: FP=fat percentage, SNP=single nucleotide polymorphism, Chr = chromosome, EBV=estimated breeding value, DRP=deregressed proof, UTR=untranslated region.
* *P*-values have been reported only for the significative SNPs in each category.

i. three SNPs (identified only for EBV) were annotated in intron positions of the Maestro Heat Like Repeat Family Member 1 (*MROH1*) gene that has been already associated with different milk traits (Iung et al., 2019; Jiang et al., 2019; Tribout et al., 2020);
ii. two SNPs (rs41256919 and rs110929299) that map in the *MAF1* homolog, a negative regulator of RNA polymerase III (*MAF1*), a gene that has been associated with all five milk production traits and milk cholesterol content (Jiang et al., 2019; Wang et al., 2019);
iii. the rs136792973 SNP, annotated in the glycosylphosphatidylinositol anchor attachment 1 – (*GPAA1*) gene, which was already associated with protein yield (Pedrosa et al., 2021) and milk production (Raschia et al., 2020). However, Massender et al. (2023) found it to be associated with FY and FP in Canadian dairy goats using 305-day lactation milk production records as phenotypes.
iv. three SNPs located in the Plectin (*PLEC*) gene, which showed associations with milk fat percentage in numerous studies (Su et al., 2023; Wang et al., 2019; Wang et al., 2022). The Plectin gene has a pleiotropic effect on most milk production traits, explaining the results found here with the SNPs significant for FP (Bekele et al., 2023; Yang et al., 2021).

One gene in the region identified on chromosome 19, which to the best of our knowledge was never associated with FP or milk

**Table 3**
List of SNPs above the Bonferroni (underlined) and false discovery rate 0.05 thresholds for both EBVs and DRP for PP trait, identified in the Aosta cattle breeds[1].

| Marker | RS_SNP_ID | Chr | Position (bp) | *P*-Value EBV* | *P*-Value DRP* | Gene | Position |
|---|---|---|---|---|---|---|---|
| BovineHD0500033518 | rs3423212215 | 5 | 114916479 | 3.27E-06 | 1.93E-05 | | Intergenic |
| ARS-BFGL-NGS-69589 | rs109415265 | 5 | 115890490 | 2.51E-05 | | *FBLN1* | Intronic |
| ARS-BFGL-NGS-18620 | rs110810286 | 5 | 116521777 | | 3.73E-06 | *CDPF1* | Intronic |
| BovineHD0500034166 | rs135679475 | 5 | 116591924 | 2.72E-06 | 7.88E-06 | *TTC38* | Intronic |
| BovineHD0500034225 | rs109155800 | 5 | 116760590 | 1.35E-05 | 2.31E-05 | *CELSR1* | Intronic |
| ARS-BFGL-NGS-6245 | rs41593908 | 5 | 117072309 | 2.62E-05 | 5.20E-06 | *TBC1D22A* | Intronic |
| BovineHD4100004172 | rs41593907 | 5 | 117075478 | 5.47E-07 | 5.70E-09 | *TBC1D22A* | Intronic |
| ARS-BFGL-NGS-14632 | rs110117542 | 5 | 117391029 | 1.39E-07 | 2.66E-06 | | Intergenic |
| BovineHD0500034461 | rs134294234 | 5 | 117440580 | 7.23E-07 | 7.39E-06 | | Intergenic |
| BovineHD0500034507 | rs109252331 | 5 | 117532054 | 8.11E-09 | 1.36E-07 | | Intergenic |
| BovineHD0500034524 | rs110310756 | 5 | 117568705 | 5.85E-06 | | | Intergenic |
| BovineHD0500035205 | rs3423206779 | 5 | 118938470 | 3.78E-06 | | | Intergenic |
| BovineHD0500035210 | rs3423206754 | 5 | 118949961 | 1.78E-05 | | | Intergenic |
| BovineHD0600021245 | rs132642659 | 6 | 74839211 | 2.02E-05 | | | Intergenic |
| ARS-BFGL-NGS-27958 | rs110239739 | 6 | 82968804 | 3.20E-05 | | | Intergenic |
| BovineHD0600023788 | rs133627704 | 6 | 84939092 | 8.04E-07 | 4.95E-08 | | Intergenic |
| Hapmap25708-BTC-043671 | rs110063049 | 6 | 85383787 | 2.05E-14 | 1.64E-12 | | Intergenic |
| DB-429-seq-rs109193501 | rs109193501 | 6 | 85424759 | 1.95E-26 | 3.82E-20 | *CSN1S1* | Intronic |
| Hapmap33451-BTC-060559 | rs110914422 | 6 | 85446151 | 5.17E-10 | 4.55E-09 | | Intergenic |
| CSN2_4 | rs109299401 | 6 | 85451221 | 4.95E-14 | 3.01E-12 | *CSN2* | Missense variant (T/G) |
| chr6_87188128 | rs108993011 | 6 | 85457804 | 1.20E-11 | 1.93E-11 | | Intergenic |
| chr6_87202566 | rs384705370 | 6 | 85470165 | 1.58E-12 | 4.77E-11 | *HSTN* | 3′ UTR variant |
| chr6_87202599 | rs378595205 | 6 | 85470198 | 3.71E-14 | 5.28E-12 | *HSTN* | Splice region variant (G/A) |
| BovineHD0600023888 | rs136049155 | 6 | 85471455 | 6.41E-07 | 3.73E-07 | *HSTN* | Intronic |
| chr6_87204247 | rs382297554 | 6 | 85471846 | 1.81E-14 | 3.01E-12 | *HSTN* | 3′ UTR variant |
| chr6_87204311 | rs386014273 | 6 | 85471910 | 5.27E-11 | 5.34E-10 | *HSTN* | 3′ UTR variant |
| chr6_87204315 | rs379649542 | 6 | 85471914 | 1.81E-14 | 3.01E-12 | *HSTN* | 3′ UTR variant |
| chr6_87204358 | rs449985830 | 6 | 85471957 | 3.94E-14 | 7.99E-12 | *HSTN* | 3′ UTR variant |
| chr6_87204403 | rs385251021 | 6 | 85472002 | 1.81E-14 | 3.01E-12 | *HSTN* | 3′ UTR variant |
| chr6_87204870 | rs382158121 | 6 | 85472469 | 9.92E-14 | 7.19E-12 | | Intergenic |
| chr6_87204878 | rs383383092 | 6 | 85472477 | 4.59E-14 | 3.55E-12 | | Intergenic |
| chr6_87205080 | rs381311750 | 6 | 85472679 | 1.84E-14 | 1.89E-12 | | Intergenic |
| chr6_87205162 | rs384622341 | 6 | 85472761 | 2.05E-14 | 3.01E-12 | | Intergenic |
| chr6_87205336 | rs382862058 | 6 | 85472936 | 2.92E-13 | 2.77E-11 | | Intergenic |
| chr6_87205349 | rs385917248 | 6 | 85472949 | 2.84E-14 | 3.01E-12 | | Intergenic |
| BovineHD4100005320 | rs133035102 | 6 | 85578801 | | 6.11E-06 | | Intergenic |
| DB-434-seq-rs43703015 | rs43703015 | 6 | 85656736 | 9.84E-06 | 3.82E-20 | *CSN3* | Missense variant (T/C) |
| CSN3_AY380228_13104_1 | rs43703016 | 6 | 85656772 | 9.67E-06 | 5.20E-09 | *CSN3* | Missense variant (C/A) |
| CSN3_AY380228_13165 | rs110014544 | 6 | 85656833 | 1.13E-05 | 7.29E-09 | *CSN3* | Synonymous variant (G/A) |
| Hapmap52348-rs29024684 | rs29024684 | 6 | 85662466 | 3.59E-06 | 4.29E-09 | | Intergenic |
| BovineHD0600023926 | rs110312754 | 6 | 85686637 | 1.07E-10 | 3.59E-14 | | Intergenic |
| BTA-115149-no-rs | rs109581772 | 6 | 85784380 | 4.55E-06 | 1.98E-07 | | Intergenic |
| BovineHD0600023955 | rs137754062 | 6 | 85817621 | | 1.40E-05 | | Intergenic |
| ARS-BFGL-NGS-24522 | rs110064541 | 6 | 86151977 | 1.72E-06 | 8.19E-08 | *RUFY3* | Intronic |
| BovineHD0600024093 | rs110091883 | 6 | 86380145 | 2.07E-06 | 5.38E-07 | | Intergenic |
| BovineHD0600024315 | rs109452259 | 6 | 87068809 | | 1.45E-05 | | Intergenic |
| BovineHD1000013824 | rs136032713 | 10 | 46034552 | 2.78E-06 | | *DAPK2* | Intronic |
| BovineHD1000030067 | rs136238611 | 10 | 101887955 | 1.28E-05 | | *NRDE2* | Intronic |
| ARS-BFGL-NGS-116624 | rs41696761 | 13 | 54955140 | 3.28E-05 | | *OSBPL2* | Intronic |
| ARS-BFGL-NGS-107234 | rs110249976 | 15 | 52381562 | 9.22E-06 | | *FCHSD2* | Synonymous variant (C/T) |
| BovineHD1500015413 | rs135702946 | 15 | 52598144 | 9.22E-06 | | *FCHSD2* | Intronic |
| ARS-BFGL-NGS-4613 | rs110428369 | 15 | 53384296 | 1.62E-05 | | *PAAF1* | Intronic |
| BovineHD1600000346 | rs110355156 | 16 | 1543848 | 9.96E-07 | | | Intergenic |
| BovineHD1600000400 | rs109375222 | 16 | 1771270 | 5.74E-07 | | | Intergenic |
| BovineHD1900017689 | rs41931384 | 19 | 61035099 | | 2.09E-05 | | Intergenic |
| BovineHD2200013363 | rs110741058 | 22 | 45910129 | 1.76E-05 | | | Intergenic |
| DB-1451-seq-rs384691767 | rs384691767 | 29 | 9510570 | 1.78E-05 | | | Intergenic |

Abbreviations: PP=protein percentage, SNP=single nucleotide polymorphism, Chr = chromosome, EBV=estimated breeding value, DRP=deregressed proof, UTR=untranslated region.
[1] See Table 2 for further details.
* *P*-values have been reported only for the significative SNPs in each category.

production traits, was the Casein Kinase 1 delta - *CSNK1D*. This gene showed an increased expression in adipose tissue and mammary gland in postpartum cows (Wang et al., 2015); it was also identified in a genome-wide association studies related to the occurrence of clinical ketosis in first parity dairy cows (Soares et al., 2021). Cruz et al. (2019) found a significant region for milk short-chain fatty acids on chromosome 19 encoding the LOC101909618, now identified as the Tubulin Folding Cofactor D – (TBCD) gene, but it is ~830 kb far from the *CSNK1D* gene.

Three genes did not overpass the Bonferroni threshold, i.e. *C11H9orf50*, *OR10AG83* and *CCDC57*, and only the last one was found to be associated with FP and other milk traits. This gene is

**Table 4**
List of SNPs above the Bonferroni (underlined) and false discovery rate 0.05 thresholds for DRP for MY trait, identified in the Aosta cattle breeds[1].

| Marker | RS_SNP_ID | Chr | Position (bp) | *P*-Value EBV* | *P*-Value DRP | Gene | Position |
|---|---|---|---|---|---|---|---|
| BovineHD0600024355 | rs110434046 | 6 | 87184768 | 1.17E-07 | – | | Intergenic |
| BovineHD0600024357 | rs137712965 | 6 | 87187812 | 8.14E-08 | – | | Intergenic |
| BovineHD0600024093 | rs110091883 | 6 | 86380145 | 3.50E-06 | – | | Intergenic |
| BovineHD0800031775 | rs109292185 | 8 | 104696362 | 3.17E-06 | – | | Intergenic |

Abbreviations: MY=milk yield, SNP=single nucleotide polymorphism, Chr = chromosome, EBV=estimated breeding value, DRP=deregressed proof.
[1]See Table 2 for further details.
* *P*-values have been reported only for the significative SNPs in each category.

the Coiled-Coil Domain Containing 57 (*CCDC57*), here mapped by the rs41922153 and rs135528222 SNPs, which has also been associated with FY and FP as well as carcass fatty acids composition (Bouwman et al., 2014; Jiang et al., 2019; Tribout et al., 2020). The other two genes identified (*OR10AG83* and *C11H9orf50*) have never been associated with milk production traits. However, the olfactory receptor family 10 subfamily AG member 83 (*OR10AG83*) was associated with milk citrate. Milk citrate is a potential early biomarker for negative energy balance in dairy cows, and for this reason, we may speculate that this gene could be somehow involved in metabolism regulation, affecting, as a consequence, milk fat content (Chen et al., 2023).

*Protein percentage trait*

For PP, different significant regions have been identified for both EBV and DRP. A first quantitative trait loci region, defined by 13 SNPs, is located on chromosome 5. This region harbours five genes previously found as associated with milk traits in different cattle breeds, such as the TBC1 domain family member 22A (*TBC1D22A*) and cadherin EGF LAG seven-pass G-type receptor 1 (*CELSR1*), associated with PP (Tribout et al., 2020), the fibulin 1 (*FBLN1*) with PP and PY (Raven et al., 2014), and the tetratricopeptide repeat domain 38 (*TTC38*) with milk eicosapentaenoic acid content (Ibeagha-Awemu et al., 2016).

A second quantitative trait loci region, defined by 33 SNPs (about 1.4 Mbp long), was identified on chromosome 6. This region harbours five genes, three of which are part of the casein's family (the αs1-casein - *CSN1S1*, the β-casein - *CSN2*, and the k-casein gene - *CSN3*) (Table 3). As recently reported by Bernini et al. (2023), these three genes are polymorphic in the Aosta cattle breeds and show different allele frequencies for the two breeds (i.e., for the k-casein the B allele has a frequency of 0.40 for the ABCH and 0.63 for the ARP while for the αs1-casein the A allele has frequencies of 0.80 and 0.95, respectively). The most significant SNP (rs109193501) is annotated within the αs1-casein gene (*CSN1S1*) and is an intronic mutation. As in this study, Kemper et al. (2016) found an association between the rs109193501 and PP concentration in the milk of Jersey cows. For the same SNP, other authors found an association with PP but also with FP in three different cattle breeds, the Braunvieh, the Fleckvieh, and the Holstein (Pausch, Emmerling, et al., 2017; Pausch, MacLeod, et al., 2017). As discussed by Kuss et al. (2005) and Korkuć et al. (2023), polymorphisms of the regulatory region of the gene have a role in modulating transcription levels that impact the production of αs1-casein, which has an exerting influence not only on milk protein but also on milk fat content and overall milk properties.

The rs109299401 SNP was here found in association with PP variability. This SNP is a missense variant (ATG/CTG) mapping in the exon 7 of the β-casein gene (*CSN2*) determining a substitution of the Methionine (amino acid position: 143) with a Leucine responsible for the I/H2 variants of the β-casein (Chessa et al., 2020). The rs109299401 SNP marker was previously associated with PP and PY in two studies carried out in Holstein cattle, one from Fontanesi et al. (2014) in which the major allele (T) was associated with a reduction of PY, PP, and higher milk somatic cell count and the other one from Viale et al. (2017) in which the association study was done considering the minor allele that resulted in an increase of 0.056% of the PP.

The three significative SNPs in the *CSN3* gene include: (i) two missense variants, the rs43703015 (ATC/ACC), causing a substitution Ile157Thr, and the rs43703016 (GCT/GAT), causing a substitution Ala169Asp, both mapped in exon 4 of the k-casein gene (*CSN3*) encoding the variants A and B of this gene (Farrell et al., 2004); (ii) a synonymous variant, the rs110014544. Numerous studies focused on the k-casein gene because of its well-known effects in the cheese production process (Viale et al., 2017). Generally, the allele A correlates with reduced protein content and increased milk yield, while allele B is linked to high protein content, better milk quality but lower milk production (Caroli et al., 2004; Schopen et al., 2011). Schopen et al. (2011) identified in Holstein-Friesian cows a significant association with milk k-casein content, PP, β-lactoglobulin content, and casein index at the rs43703016 SNP.

In the quantitative trait loci region on chromosome 6, two more genes were in common between the results obtained with the DRP and EBV. The first one is the histatherin - *HSTN* gene which is near a regulatory element that affects the expression of the β-casein gene (Pegolo et al., 2021). In many studies, the *HSTN* gene has been found associated with PY, PP, and αs1-casein and β-casein concentration in milk (Jiang et al. 2019; Tribout et al. 2020; Pegolo et al. 2021). Even the genes mapping on chromosome 10 (death–associated protein kinase 2 - *DAPK2*) and chromosome 13 (oxysterol binding protein–like 2 - *OSBPL2*) were already associated with milk PP as well as with milk traits (*NRDE2*, necessary for RNA interference, domain containing - *NRDE2*, chromosome 10) (Jiang et al., 2019).

The Aosta cattle national breeder association reports the individual genotype at the k-casein, β-casein and β-lactoglobulin genes in the bull's catalogue. This information allowed the selection of the different casein variants for cheese yield, increasing as such also the allele frequencies of the B allele at the β-lactoglobulin, reported to be 0.57 and 0.68 in the 1980 s by (Merlin and Di Stasio, 1982) to current frequencies of 0.73 and 0.69 for the ARP and ABCH, respectively (Bernini et al., 2023). Indeed, the favourable alleles for the protein content and cheese-making proprieties are also the most present in the population (e.g., for the β-lactoglobulin 49 and 54% of subjects with BB genotype for the ABCH and ARP breed respectively, and at the k-casein gene the B allele has the highest frequency in the ARP breed with the absence of the E allele in the whole population) (Bernini et al., 2023). The differences in the genotypic frequencies may reflect the diverse evolutionary history of the two breeds, showing different origins and bred for a slightly different purpose: ARP is mainly bred for milk yield, to produce Fontina cheese, whereas the leading interest for ABCH breeders is the fighting (Sartori et al., 2020).

*Milk yield trait*

Two quantitative trait loci regions were associated with MY_EBV but not for the MY_DRP dependent variable, probably because of the bigger sample size of the EBV sample compared to the one of the DRP. Most likely, the reason why this is occurring for MY and not for PP or FP is related to the average lower genetic variance of the phenotype MY (8.32 and 9.72 for the ABCH and ARP, respectively) compared to the other two traits: FP (43.57 for the ABCH and 58.62 for the ARP) and PP (20.5 in the ABCH and 22.56 in the ARP). The significant SNP rs110434046 has been associated with the Daughter Pregnancy Rate in a study of Liang et al. (2023) on a sample of a million Holstein cows. In a study based on the same Holstein cow's dataset, the same SNP (rs110434046) resulted in an epistatic effect with the rs109421300 SNP in the DGAT1 gene (Prakapenka et al. 2024).

## Conclusion

The results of this study show that even if the majority of the quantitative trait loci identified in this study have been previously associated with milk production traits, the Aosta population owns a peculiar genetic structure that differentiates them from the cosmopolitan specialised breeds intensively selected for milk yield (Signer-Hasler et al., 2023). The population studied here has been selected with a much lower intensity for milk traits with respect to specialised dairy breeds such as the Holstein one. Being a double−purpose population with a strong aptitude for pasture in harsh mountain environments the selection programme is in fact oriented to several breeding objectives: (i) to improve the milk yield and its physical and chemical characteristics as a function of cheese yield and renneting properties; (ii) to improve the amount of beef produced and the estimated carcass quality; (iii) to maintain the genetic variability; (iv) to maintain the hardiness and longevity; these latter two are intrinsic characteristics of the population. In addition to the fact that the selection is both for meat and milk traits, the mating scheme is oriented to make the gene flow of males used in reproduction as homogeneous as possible in the female population, to maintain the genetic variability as large as possible. The ongoing breeding and selection scheme has been active for decades making the population strongly homogenous in its genomic makeup, a condition that is possibly affecting the identification of region containing quantitative trait loci under segregation for milk traits. For some of the novel identified quantitative trait loci regions, the analysis of sequence data may further explore the genomic variation here found, to determine the presence of proprietary polymorphism of the Aosta cattle with respect to other cosmopolitan and specialised populations.

## Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.animal.2024.101322.

## Ethics approval

Not Applicable.

## Data and model availability statement

The data supporting the findings of this study are available within the article and its Supplementary Materials. The raw genetic datasets generated during the current study are available from the corresponding author upon reasonable request. The data were not deposited in an official repository.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) did not use any AI and AI-assisted technologies.

## Author ORCIDs

**Francesca Bernini:** https://orcid.org/0000-0001-8430-6813.
**Enrico Mancin:** https://orcid.org/0000-0002-4130-4639.
**Cristina Sartori:** https://orcid.org/0000-0002-2091-2280.
**Mario Vevey:** N/A.
**Veruska Blanchet:** N/A.
**Roberto Mantovani:** https://orcid.org/0000-0002-1919-0559.
**Alessandro Bagnato:** https://orcid.org/0000-0003-2087-1287.
**Maria Giuseppina Strillacci:** https://orcid.org/0000-0002-8972-1730.

## CRediT authorship contribution statement

**F. Bernini:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **E. Mancin:** Writing – review & editing, Formal analysis. **C. Sartori:** Writing – review & editing, Formal analysis. **R. Mantovani:** Writing – review & editing, Resources, Data curation. **M. Vevey:** Resources, Funding acquisition. **V. Blanchet:** Resources, Data curation. **A. Bagnato:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Conceptualization. **M.G. Strillacci:** Writing – review & editing, Writing – original draft, Supervision.

## Declaration of interest

None.

## Acknowledgements

None.

## Financial support statement

Project name: "Le razze bovine a duplice attitudine: un modello alternativo di zootecnia eco-sostenibile - DUAL BREEDING FASE 2" - PSRN Programma di Sviluppo Rurale Nazionale 2014/2020 - Sottomisura 10.2 - PERIODO 2020-2023 - funded by the Ministero dell'Agricoltura della Sovranità Alimentare e delle Foreste with economic support of European Agricultural Fund for Rural Development. Grant n.: CUP_J72C21000650005. Project name: Prin 2022 "GENOmic Breeding VALue Estimation in a Native Alpine Cattle Breed" grant CUP_G53D23004040006 - Finanziato dall'Unione europea - Next Generation EU.

## References

Bakhshalizadeh, S., Zerehdaran, S., Javadmanesh, A., 2021. Meta-analysis of genome-wide association studies and gene networks analysis for milk production traits in Holstein cows. Livestock Science 250, 104605.

Bekele, R., Taye, M., Abebe, G., Meseret, S., 2023. Genomic regions and candidate genes associated with milk production traits in Holstein and its crossbred cattle: a review. International Journal of Genomics 2023, 8497453. https://doi.org/10.1155/2023/8497453.

Bernini, F., Punturiero, C., Vevey, M., Blanchet, V., Milanesi, R., Delledonne, A., Bagnato, A., Strillacci, M.G., 2023. Assessing major genes allele frequencies and the genetic diversity of the native Aosta cattle female population. Italian Journal of Animal Science 22, 1008–1022.

Bouwman, A.C., Visker, M.H.P.W., van Arendonk, J.M., Bovenhuis, H., 2014. Fine mapping of a quantitative trait locus for bovine milk fat composition on Bos taurus autosome 19. Journal of Dairy Science 97, 1139–1149.

Bouwman, A.C., Daetwyler, H.D., Chamberlain, A.J., Ponce, C.H., Sargolzaei, M., Schenkel, F.S., Sahana, G., Govignon-Gion, A., Boitard, S., Dolezal, M., 2018.

Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nature Genetics 50, 362–367.
Caroli, A., Chessa, S., Bolla, P., Budelli, E., Gandini, G.C., 2004. Genetic structure of milk protein polymorphisms and effects on milk production traits in a local dairy cattle. Journal of Animal Breeding and Genetics 121, 119–127.
Chen, Y., Hu, H., Atashi, H., Grelet, C., Wijnrocx, K., Lemal, P., Gengler, N., 2023. Genetic analysis of milk citrate predicted by milk mid-infrared spectra of Holstein cows in early lactation. Journal of Dairy Science 107, 3047–3061 10.3168/jds.2023-23903.
Chen, S.-Y., Schenkel, F.S., Melo, A.L.P., Oliveira, H.R., Pedrosa, V.B., Araujo, A.C., Melka, M.G., Brito, L.F., 2022. Identifying pleiotropic variants and candidate genes for fertility and reproduction traits in Holstein cattle via association studies based on imputed whole-genome sequence genotypes. BMC Genomics 23, 331.
Chessa, S., Gattolin, S., Cremonesi, P., Soglia, D., Finocchiaro, R., Van Kaam, J.-T., Marusi, M., Civati, G., 2020. The effect of selection on casein genetic polymorphisms and haplotypes in Italian Holstein cattle. Italian Journal of Animal Science 19, 833–839.
Cruz, V.A.R., Oliveira, H.R., Brito, L.F., Fleming, A., Larmer, S., Miglior, F., Schenkel, F. S., 2019. Genome-wide association study for milk fatty acids in Holstein cattle accounting for the DGAT1 gene effect. Animals (basel) 9, 997.
Farrell Jr, H.M., Jimenez-Flores, R., Bleck, G.T., Brown, E.M., Butler, J.E., Creamer, L.K., Hicks, C.L., Hollar, C.M., Ng-Kwai-Hang, K.F., Swaisgood, H.E., 2004. Nomenclature of the proteins of cows' milk—Sixth revision. Journal of Dairy Science 87, 1641–1674.
Fontanesi, L., Calò, D.G., Galimberti, G., Negrini, R., Marino, R., Nardone, A., Ajmone-Marsan, P., Russo, V., 2014. A candidate gene association study for nine economically important traits in Italian Holstein cattle. Animal Genetics 45 (4), 576–580.
Garrick, D.J., Taylor, J.F., Fernando, R.L., 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics Selection Evolution 41, 1–8.
Glantz, M., Månsson, H.L., Paulsson, M., Stålhammar, H., 2012. Genomic selection in relation to bovine milk composition and processability. Journal of Dairy Research 79, 53–59.
Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., 2002. The Ensembl genome database project. Nucleic Acids Research 30, 38–41.
Ibeagha-Awemu, E.M., Peters, S.O., Akwanji, K.A., Imumorin, I.G., Zhao, X., 2016. High density genome wide genotyping-by-sequencing and association identifies common and low frequency SNPs, and novel candidate genes influencing cow milk traits. Scientific Reports 6, 31109.
Iung, L.H.S., Petrini, J., Ramírez-Díaz, J., Salvian, M., Rovadoscki, G.A., Pilonetto, F., Dauria, B.D., Machado, P.F., Coutinho, L.L., Wiggans, G.R., 2019. Genome-wide association study for milk production traits in a Brazilian Holstein population. Journal of Dairy Science 102, 5305–5314.
Jayawardana, J., Lopez-Villalobos, N., McNaughton, L.R., Hickson, R.E., 2023. Genomic regions associated with milk composition and fertility traits in spring-calved dairy cows in New Zealand. Genes 14, 860.
Jiang, J., Ma, L., Prakapenka, D., VanRaden, P.M., Cole, J.B., Da, Y., 2019. A large-scale genome-wide association study in US Holstein cattle. Frontiers in Genetics 10, 442321.
Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Sabatti, C., Eskin, E., 2010. Variance component model to account for sample structure in genome-wide association studies. Nature Genetics 42, 348–354.
Kemper, K.E., Littlejohn, M.D., Lopdell, T., Hayes, B.J., Bennett, L.E., Williams, R.P., Xu, X.Q., Visscher, P.M., Carrick, M.J., Goddard, M.E., 2016. Leveraging genetically simple traits to identify small-effect variants for complex phenotypes. BMC Genomics 17, 1–9.
Korkuć, P., Arends, D., May, K., König, S., 2021. Genomic loci affecting milk production in German Black Pied cattle (DSN). Frontiers in Genetics 12, 640039.
Korkuć, P., Neumann, G.B., Hesse, D., Arends, D., Reißmann, M., Rahmatalla, S., May, K., Wolf, M.J., König, S., Brockmann, G.A., 2023. Whole-genome sequencing data reveal new loci affecting milk production in German Black Pied Cattle (DSN). Genes 14, 581.
Kühn, C., Thaller, G., Winter, A., Bininda-Emonds, O.R.P., Kaupe, B., Erhardt, G., Bennewitz, J., Schwerin, M., Fries, R., 2004. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics 167, 1873–1881.
Kuss, A.W., Gogol, J., Bartenschlager, H., Geldermann, H., 2005. Polymorphic AP-1 binding site in bovine CSN1S1 shows quantitative differences in protein binding associated with milk protein expression. Journal of Dairy Science 88, 2246–2252.
Liang, Z., Prakapenka, D., VanRaden, P.M., Jiang, J., Ma, L., Da, Y., 2023. A million-cow genome-wide association study of three fertility traits in US Holstein cows. International Journal of Molecular Sciences 24, 10496.
Littlejohn, M.D., Tiplady, K., Fink, T.A., Lehnert, K., Lopdell, T., Johnson, T., Couldrey, C., Keehan, M., Sherlock, R.G., Harland, C., 2016. Sequence-based association analysis reveals an MGST1 eQTL with pleiotropic effects on bovine milk composition. Scientific Reports 6, 25376.
Mancin, E., Tuliozi, B., Pegolo, S., Sartori, C., Mantovani, R., 2022. Genome wide association study of beef traits in local Alpine breed reveals the diversity of the pathways involved and the role of time stratification. Frontiers in Genetics 12, 746665.
Manzanilla-Pech, C.I.V., Løvendahl, P., Gordo, D.M., Difford, G.F., Pryce, J.E., Schenkel, F., Wegmann, S., Miglior, F., Chud, T.C., Moate, P.J., 2021. Breeding for reduced methane emission and feed-efficient Holstein cows: an international response. Journal of Dairy Science 104, 8983–9001.
Massender, E., Oliveira, H.R., Brito, L.F., Maignel, L., Jafarikia, M., Baes, C.F., Sullivan, B., Schenkel, F.S., 2023. Genome-wide association study for milk production and conformation traits in Canadian Alpine and Saanen dairy goats. Journal of Dairy Science 106, 1168–1189.
Mazza, S., Guzzo, N., Sartori, C., Mantovani, R., 2016. Genetic correlations between type and test-day milk yield in small dual-purpose cattle populations: the Aosta Red Pied breed as a case study. Journal of Dairy Science 99, 8127–8136.
McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., Cunningham, F., 2016. The ensembl variant effect predictor. Genome Biology 17, 1–14.
Merlin, P., Di Stasio, L., 1982. Study on milk proteins loci in some decreasing Italian cattle breeds. Annales De Génétique et De Sélection Animale 14, 17–28.
Misztal, I., Wang, H., Aguilar, I., Legarra, A., Tsuruta, S., Lourenco, D.A.L., Fragomeni, B., Zhang, X., Muir, W., Cheng, H.H., 2014. GWAS using ssGBLUP. Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, 17–22.
Nguyen, T.T.T., Bowman, P.J., Haile-Mariam, M., Nieuwhof, G.J., Hayes, B.J., Pryce, J.E., 2017. Implementation of a breeding value for heat tolerance in Australian dairy cattle. Journal of Dairy Science 100, 7362–7367.
Oliveira, H.R., Cant, J.P., Brito, L.F., Feitosa, F.L.B., Chud, T.C.S., Fonseca, P.A.S., Jamrozik, J., Silva, F.F., Lourenco, D.A.L., Schenkel, F.S., 2019. Genome-wide association for milk production traits and somatic cell score in different lactation stages of Ayrshire, Holstein, and Jersey dairy cattle. Journal of Dairy Science 102, 8159–8174.
Pagnacco, G., Gandini, G.C., Bagnato, A., Miglior, F., Caroli, A., 1989. Genetic gain and conservation in a small alpine cattle breed. Journal of Animal Breeding and GeneTics 106, 351–357.
Pausch, H., Emmerling, R., Gredler-Grandl, B., Fries, R., Daetwyler, H.D., Goddard, M. E., 2017a. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. BMC Genomics 18, 1–11.
Pausch, H., MacLeod, I.M., Fries, R., Emmerling, R., Bowman, P.J., Daetwyler, H.D., Goddard, M.E., 2017b. Evaluation of the accuracy of imputed sequence variant genotypes and their utility for causal variant detection in cattle. Genetics Selection Evolution 49, 1–14.
Pedrosa, V.B., Schenkel, F.S., Chen, S.-Y., Oliveira, H.R., Casey, T.M., Melka, M.G., Brito, L.F., 2021. Genomewide association analyses of lactation persistency and milk production traits in Holstein cattle based on imputed whole-genome sequence data. Genes 12, 1830.
Pedrosa, V.B., Boerman, J.P., Gloria, L.S., Chen, S.-Y., Montes, M.E., Doucette, J.S., Brito, L.F., 2023. Genomic-based genetic parameters for milkability traits derived from automatic milking systems in North American Holstein cattle. Journal of Dairy Science 106, 2613–2629.
Pegolo, S., Yu, H., Morota, G., Bisutti, V., Rosa, G.J.M., Bittante, G., Cecchinato, A., 2021. Structural equation modeling for unraveling the multivariate genomic architecture of milk proteins in dairy cattle. Journal of Dairy Science 104, 5705–5718.
Prakapenka, D., Liang, Z., Zaabza, H.B., VanRaden, P.M., Van Tassell, C.P., Da, Y., 2024. A million-cow validation of a chromosome 14 region interacting with all chromosomes for fat percentage in US Holstein cows. International Journal of Molecular Sciences 25, 674.
Raschia, M.A., Nani, J.P., Carignano, H.A., Amadio, A.F., Maizon, D.O., Poli, M.A., 2020. Weighted single-step genome-wide association analyses for milk traits in Holstein and Holstein x Jersey crossbred dairy cattle. Livestock Science 242, 104294.
Raven, L.-A., Cocks, B.G., Hayes, B.J., 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics 15, 1–14.
Sahana, G., Cai, Z., Sanchez, M.P., Bouwman, A.C., Boichard, D., 2023. Invited review: Good practices in genome-wide association studies to identify candidate sequence variants in dairy cattle. Journal of Dairy Science 106, 5218–5241. https://doi.org/10.3168/jds.2022-22694.
Sanchez, M.-P., Govignon-Gion, A., Croiseau, P., Fritz, S., Hozé, C., Miranda, G., Martin, P., Barbat-Leterrier, A., Letaïef, R., Rocha, D., 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genetics Selection Evolution 49, 1–16.
Sartori, C., Guzzo, N., Mantovani, R., 2020. Genetic correlations of fighting ability with somatic cells and longevity in cattle. Animal 14, 13–21. https://doi.org/10.1017/S175173111900168X.
Schopen, G.C.B., Visker, M., Koks, P.D., Mullaart, E., Van Arendonk, J.A.M., Bovenhuis, H., 2011. Whole-genome association study for milk protein composition in dairy cattle. Journal of Dairy Science 94, 3148–3158.
Signer-Hasler, H., Casanova, L., Barenco, A., Maitre, B., Bagnato, A., Vevey, M., Berger, B., Simčič, M., Boichon, D., Capitan, A., 2023. Genomic regions underlying positive selection in local, Alpine cattle breeds. Animal Genetics 54, 239–253.
Soares, R.A.N., Vargas, G., Duffield, T., Schenkel, F., Squires, E.J., 2021. Genome-wide association study and functional analyses for clinical and subclinical ketosis in Holstein cattle. Journal of Dairy Science 104, 10076–10089.
Strillacci, M.G., Vevey, M., Blanchet, V., Mantovani, R., Sartori, C., Bagnato, A., 2020. The genomic variation in the Aosta Cattle breeds raised in an extensive alpine farming system. Animals 10, 2385.

Strillacci, M.G., Punturiero, C., Milanesi, R., Bernini, F., Mason, T., Bagnato, A., 2023. Antibiotic treatments and somatic cell count as phenotype to map QTL for mastitis susceptibility in Holstein cattle breed. Italian Journal of Animal Science 22, 190–199.

Su, M., Lin, X., Xiao, Z., She, Y., Deng, M., Liu, G., Sun, B., Guo, Y., Liu, D., Li, Y., 2023. Genome-wide association study of lactation traits in Chinese Holstein cows in Southern China. Animals 13, 2545.

Tribout, T., Croiseau, P., Lefebvre, R., Barbat, A., Boussaha, M., Fritz, S., Boichard, D., Hoze, C., Sanchez, M.-P., 2020. Confirmed effects of candidate variants for milk production, udder health, and udder morphology in dairy cattle. Genetics Selection Evolution 52, 1–26.

van den Berg, I., Boichard, D., Lund, M.S., 2016. Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. Journal of Dairy Science 99, 8932–8945.

Viale, E., Tiezzi, F., Maretto, F., De Marchi, M., Penasa, M., Cassandro, M., 2017. Association of candidate gene polymorphisms with milk technological traits, yield, composition, and somatic cell score in Italian Holstein-Friesian sires. Journal of Dairy Science 100, 7271–7281.

Wang, P., Li, X., Zhu, Y., Wei, J., Zhang, C., Kong, Q., Nie, X., Zhang, Q., Wang, Z., 2022. Genome-wide association analysis of milk production, somatic cell score, and body conformation traits in Holstein cows. Frontiers in Veterinary Science 9, 932034.

Wang, D., Ning, C., Liu, J.-F., Zhang, Q., Jiang, L., 2019. Replication of genome-wide association studies for milk production traits in Chinese Holstein by an efficient rotated linear mixed model. Journal of Dairy Science 102, 2378–2383.

Wang, M., Zhou, Z., Khan, M.J., Gao, J., Loor, J.J., 2015. Clock circadian regulator (CLOCK) gene network expression patterns in bovine adipose, liver, and mammary gland at 3 time points during the transition from pregnancy into lactation. Journal of Dairy Science 98, 4601–4612.

Yang, Z., Lian, Z., Liu, G., Deng, M., Sun, B., Guo, Y., Liu, D., Li, Y., 2021. Identification of genetic markers associated with milk production traits in Chinese Holstein cattle based on post genome-wide association studies. Animal Biotechnology 32, 67–76.