

Article

Genome-Wide Association Studies for Milk Somatic Cell Score in Romanian Dairy Cattle

Daniela Elena Ilie ^{1,2,*}, Alexandru Eugeniu Mizeranschi ^{1,†}, Ciprian Valentin Mihali ¹, Radu Ionel Neamț ¹, George Vlad Goilean ¹, Ovidiu Ionuț Georgescu ², Daniela Zaharie ³, Mihai Carabaș ⁴ and Ioan Huțu ²

¹ The Molecular Research Department, Research and Development Station for Bovine Arad, Bodrogului Street, No. 32, 310059 Arad, Romania; alex.mizeranschi@gmail.com (A.E.M.); mihaliciprian@yahoo.com (C.V.M.); neamtr@yahoo.com (R.I.N.); vladgoilean@yahoo.com (G.V.G.)

² Faculty of Veterinary Medicine, Banat's University of Agricultural Sciences and Veterinary Medicine "King Michael I of Romania" from Timisoara, Calea Aradului No. 119, 300645 Timisoara, Romania; georgescu_ovidiu.ionut@yahoo.com (O.I.G.); ioan.hutu@fmvt.ro (I.H.)

³ Faculty of Mathematics and Computer Science, West University of Timișoara, 300223 Timișoara, Romania; daniela.zaharie@e-uvt.ro

⁴ Faculty of Automatic Control and Computer Science, Politehnica University of Bucharest, 060042 București, Romania; mihai.carabas@roedu.net

* Correspondence: danailie@animalsci-tm.ro

† These authors contributed equally to this work.

Abstract: Mastitis is one of the most frequently encountered diseases in dairy cattle, negatively affecting animal welfare and milk production. For this reason, contributions to understanding its genomic architecture are of great interest. Genome-wide association studies (GWAS) have identified multiple loci associated with somatic cell score (SCS) and mastitis in cattle. However, most of the studies have been conducted in different parts of the world on various breeds, and none of the investigations have studied the genetic architecture of mastitis in Romanian dairy cattle breeds up to this point in time. In this study, we report the first GWAS for SCS in dairy cattle breeds from Romania. For GWAS, we used an Axiom Bovine v3 SNP-chip (>63,000 Single Nucleotide Polymorphism -SNPs) and 33,330 records from 690 cows belonging to Romanian Spotted (RS) and Romanian Brown (RB) cattle. The results found one SNP significantly associated with SCS in the RS breed and 40 suggestive SNPs with $-\log_{10}(p)$ from 4 to 4.9 for RS and from 4 to 5.4 in RB. From these, 14 markers were located near 12 known genes (*AKAP8*, *CLHC1*, *MEGF10*, *SATB2*, *GATA6*, *SPATA6*, *COL12A1*, *EPS8*, *LUZP2*, *RAMAC*, *IL12A* and *ANKRD55*) in RB cattle, 3 markers were close to *ZDHHC19*, *DAPK1* and *MMP7* genes, while one SNP overlapped the *HERC3* gene in RS cattle. Four genes (*HERC3*, *LUZP2*, *AKAP8* and *MEGF10*) associated with SCS in this study were previously reported in different studies. The most significant SNP (rs110749552) associated with SCS was located within the *HERC3* gene. In both breeds, the SNPs and position of association signals were distinct among the three parities, denoting that mastitis is controlled by different genes that are dependent according to parity. The current results contribute to an expansion in the body of knowledge regarding the proportion of genetic variability explained by SNPs for SCS in dairy cattle.

Keywords: genome-wide association study (GWAS); mastitis; Romanian cattle; somatic cell score



Citation: Ilie, D.E.; Mizeranschi, A.E.; Mihali, C.V.; Neamț, R.I.; Goilean, G.V.; Georgescu, O.I.; Zaharie, D.; Carabaș, M.; Huțu, I. Genome-Wide Association Studies for Milk Somatic Cell Score in Romanian Dairy Cattle. *Genes* **2021**, *12*, 1495. <https://doi.org/10.3390/genes12101495>

Academic Editor: Ottmar Distl

Received: 1 September 2021

Accepted: 22 September 2021

Published: 24 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent decades, the prevalence of clinical mastitis in cattle has increased significantly [1] in most of developed agricultural countries where farming practices are more intensive. This tendency is manifested in parallel with a higher selection pressure for high milk yields in modern dairy breeds. Therefore, improving the understatement of the genetic base of mastitis and its indicator traits represents a major goal for the dairy cattle breeding industry as it can improve udder health and increase the milk quality, while also reducing early involuntary culling, discarded milk, veterinary services and labor costs.

In dairy cattle, many studies have been conducted in different parts of the world on various breeds in order to investigate the genetic base of mastitis. Research results from more than two decades have shown the association between the major histocompatibility complex (MHC) and susceptibility or resistance to intramammary infection [2]. In the beginning, most of the studies focused on MHC genes, particularly the DRB region, for which various associations between allelic variants and immune response and mastitis resistance have been reported by several authors [3–6]. Later, quantitative trait loci (QTL) responsible for clinical mastitis and somatic cell counts (SCC) have been identified in the majority of the chromosomes. Among these, QTL on chromosome 1, 3, 7, 8, 14, 18, 21 and 23 were confirmed in different separate studies [1]. Moreover, the technological developments of molecular genetics from the last decade, followed by sequencing of the *Bos taurus* genome in 2009 [7], enabled the identification of several molecular markers responsible for susceptibility or resistance to mastitis. The appearance of high-throughput genotyping technologies allowed genome-wide association studies to be carried out in cattle [8]. By the application of new genetic analysis methods and genome-wide association studies (GWAS), exploring the genetic architecture of the important traits in different cattle breeds [9] and detecting the significant single nucleotide polymorphisms (SNPs) or genomic regions associated with mastitis in dairy cattle became possible. For instance, Meredith et al. [8] used 773 Holstein-Friesian AI sires with progeny genotyped by using the Illumina BovineSNP50 Genotyping Beadchip and identified nine SNPs that were significantly associated with somatic cell score (SCS), of which three were located on chromosomes 6 and 10 within known QTL regions for SCS, and six SNPs placed outside known QTL regions for SCS were located on chromosomes 6, 15 and 20. One year later, the same author [10], by using 702 Holstein-Friesian sires that were genotyped for 777,962 SNPs on the Illumina High-density Bovine BeadChip, detected 28 QTL regions associated with SCS, with 138 SNPs located across 15 chromosomes (1, 3, 4, 5, 6, 9, 10, 13, 17, 20, 21, 22, 23, 24, 25 and 26). Strillacci et al. [11] conducted a GWAS for SCS in Valdostana Red Pied cattle and found genes involved in mastitis resistance or variation of SCS in QTL on chromosomes 9, 13, 15, 17, 19, 21 and 22. Another study [12] analyzed 544 Holstein and Holstein × Jersey cows and identified six SNPs on chromosomes 1, 5, 10, 18 and 26 that were associated with traits derived from SCC. Recently, a genome-wide association study [9] found three significant SNPs on chromosomes 5, 8 and 22 that were associated with SCS in 2410 Xinjiang Brown cows. According to the current release of the Cattle Quantitative Trait Loci (QTL) Database of Animal QTLdb (The Animal Quantitative Trait Loci Database, <https://www.animalgenome.org/cgi-bin/QTLdb/BT/index>, accessed on 20 May 2021 [13]), a total of 2401 QTLs, spread on most bovine chromosomes, were found to be associated with mastitis.

Mastitis resistance is a complex trait and although important progress has been made in this area, the research efforts should be intensified in order to better understand the genetic architecture of this important trait and, also, to find new possible markers in order to improve the possibilities of genetic progress and enable more accurate selection against mastitis. Moreover, although GWAS was widely used to identify the genetic variants associated with mastitis and its indicator traits such as SCS, to the best of our knowledge, none of the studies investigated the genetic architecture of SCS in Romanian dairy cattle breeds. To address this issue, in the current study, we used the GWAS approach in 723 Romanian Spotted (RS, national name: Baltata Romaneasca) breed (Breeds of Livestock, Department of Animal Science; Breeds of Livestock—Baltata Romaneasca Cattle (<http://afs.okstate.edu/breeds/cattle/baltataromaneasca/index.html/>, accessed on 25 August 2021 [14]) and Romanian Brown (BR, national name Bruna de Maramures) cattle breeds from three distinct herds to identify candidate genes and genomic regions associated with SCS.

2. Materials and Methods

2.1. Animals and Phenotypes

This study was conducted on 723 dairy cattle out of which 601 were RS and 122 were RB cattle. The animals involved were managed in three breeding herds: two experimental herds belonging to the Romanian Academy of Agricultural Science (the Research and Development Station for Bovine Arad and the Research and Development Station for Bovine Sighetu Marmatiei) and one commercial herd located in Berliste. All cows involved in the study were kept under comparable housing and feeding conditions, with similar milking and sanitary conditions, and were mechanically milked twice a day. All the animals had at least 3 monthly test-day records per lactation. Only data corresponding to the first three lactations were included in the analysis.

Phenotypic data consisted of 33,330 SCC records, of which 24,295 were for RS and 9035 were for RB cattle. For SCC determination, milk records were conducted every 28 days by the Official Dairy Control service. The milk samples were analyzed for SCC determination in the laboratory of the Milk Quality Control Foundation (Cluj-Napoca, Romania) by using CombiFoss™ FT + integrating MilkoScan™ FT + and Fossomatic™ FC.

In order to achieve a normal distribution of the data, the values of SCC were transformed to SCS according to the formula of Wiggans and Shook [15] as $SCS = \log_2(SCC/100,000) + 3$, specified by the international standard (Interbull Code of Practice. <https://interbull.org/ib/codeofpractice>, accessed on 25 August 2021 [16]).

2.2. Sampling and Genotyping

Biological samples needed for genotyping consisted of whole blood and hair roots that were collected by trained veterinarians. Blood samples ($n = 627$) were collected from the tail vein in vacutainers containing K3EDTA as an anticoagulant, and the hair roots (approximately 20–40 hairs with follicles attached) were collected from the tail of the animals ($n = 96$). All protocols used for collection of the biological samples were approved by the Ethics Committee of the Research and Development Station for Bovine Arad (Statement no 88/04.10.2019). After collection, the samples were transferred to the laboratory and stored at 4 °C, and they were subsequently sent to IFN Schönow GmbH (Bernau bei Berlin, Germany) for DNA extraction and genotyping using an Axiom Bovine v3 microarray (based on the reference genome *Bos_taurus_UMD_3.1.1*), which includes >63,000 SNPs. The used SNP chip included a high percentage of SNPs that are found distributed on all chromosomes. The SNPs on the array include 44,000 markers from the Council on Dairy Cattle Breeding (CDCB), International Society for Animal Genetics (ISAG) core parentage markers, VIP trait-associated SNPs and SNPs optimized for STR imputation. In total, 723 samples were genotyped.

2.3. Quality Control (QC)

Data quality control was performed in order to remove SNPs and individuals with insufficient genotyping quality. QC procedures were carried out by using PLINK v1.90b6.16 64-bit [17]. The SNPs located on the sex chromosome or those with unknown positions were removed from the current study. Individuals with a call rate of less than 95% were excluded. The SNPs with call rates <95% or with minor allele frequency (MAF) < 0.05 and SNPs with genotypes not in accordance with the Hardy–Weinberg equilibrium ($p > 10^{-6}$) were eliminated.

2.4. Association Analysis

Principal component analysis (PCA) was performed for each breed-specific subpopulation in order to check for population stratification. In both cases, the first 20 principal components were computed by using the R package *argyle* v0.2.2 [18]. Based on the proportion of variation explained by principal components, we noted that the first 8 principal components explained the most significant amount of genetic variability (23.86% of the total variance in RB and 17.49% in RS, respectively) and, thus, captured any possible effects

of population stratification (Supplementary Figure S1). For this reason, we included the first 8 principal components into the GWAS study as covariates for both breeds.

The GWAS was performed for the SCS trait by using two-stage analysis. First, a random-regression test-day model using only phenotypes and pedigree data was implemented in *Blupf90* [19] by using the combined variable herd-year-season as the fixed effect in order to represent contemporary groups and by using fourth-order Legendre polynomial coefficients for both fixed and random effects (additive genetic and permanent environmental effects) in order to represent components of the lactation curve. The variance components were computed by using the *remlf90* command, after which the estimated breeding values (EBVs) and prediction error variances and covariances (PEVs) were computed using the *blupf90* command with the *store_peg_pec* option. PEVs were used to compute EBV reliabilities by using the following formula:

$$Rel = 1 - PEV/V_A, \quad (1)$$

where V_A is the additive genetic variance.

EBV reliabilities were then used to compute de-regressed proofs (DRPs) by using the R package *DRP* [20], which implements the approach proposed by Garrick et al. [21]. More specifically, the *wideDRP* function was used by setting the heritability to its estimated value and using the default value of 0.5 for the *c* parameter, which represents the fraction of genetic variance not explained by markers.

For the second GWAS stage, DRPs were used as phenotypes in the *rrBLUP* package v.4.6.1 [22] by using the *GWAS* function, with only the first 8 principal components included as fixed effects. SNP effects were computed within the same R package by using the *mixed.solve* function. A genome-wide significance threshold of $p < 1.00 \times 10^{-4}$ was used, and SNPs exceeding this threshold were considered as candidate variants. Manhattan plots were created by using the R package *qqman* v.0.1.4 [23].

Linkage disequilibrium analyses for some of the significant SNPs on chromosome 7 and 9 in RB breed and chromosome 6 in RS breed and SNPs from the surrounding regions ($\pm 500,000$ Kbp relative to the SNP's position) were performed by using the *IntRegionalPlot* function of the R package *IntAssoPlot* v.0.99.21 [24]. This function produces a figure consisting of two layers. The top layer represents a scatter plot of the $-\log_{10}(p)$ values for all the SNPs in a selected region, with the chromosomal position on the x-axis. The bottom layer depicts a heat map of LD blocks for each of the pairs of SNPs within the region. In addition, the function adds lines between the two layers to show which SNP corresponds to which LD block.

Finally, significant SNPs were annotated with overlapped and downstream/upstream genes within a ± 1 MBp range (2 MBp total) by using the R package *VariantAnnotation* v.1.28.13 [25], and annotations were obtained via the R package *GenomicFeatures* v.1.34.8 [26] from the UMD3.1.94 GTF file downloaded from Ensembl.

3. Results

3.1. Descriptive Statistics

A total number of 690 cows with 33,330 test-day milk records were used in this study. The number of records for second (31.66%) and third parity (24.71%) was lower than those for the first lactation (43.63%). For the two breed-specific subpopulations by parity (L_1 – L_3), the corresponding percentages were as follows: 45.61%, 30.78% and 23.61% in RS and 38.32%, 34.00% and 27.68% in the RB breed, respectively. The descriptive statistics of the observed phenotypes for each breed and lactation are presented in Table 1.

According to our results, the average SCSs among parities (L_1 – L_3) were 2.81, 3.06 and 3.27 in RS and 4.53, 4.51 and 4.62 in RB cattle, respectively. The maximum level of SCS was 8.84; nevertheless, the number of cows with this level of SCS was low in both breeds. The overall mean SCS for RS was 2.99 ± 1.82 and 4.55 ± 1.89 in RB cattle.

Table 1. Descriptive statistics for somatic cell score (SCS) in Romanian dairy cattle.

Breed	Trait	N	Mean	SD	Min	Max	h^2	V_P	V_A
RS	LSCS ₁	11,081	2.81	1.72	−1.64	8.84	0.08	2.9916	0.2564
	LSCS ₂	7479	3.06	1.89	−2.06	8.82	0.10	3.4975	0.3686
	LSCS ₃	5735	3.27	1.88	−2.06	8.83	0.11	3.5347	0.3966
RB	LSCS ₁	3462	4.53	1.92	−0.84	8.82	0.03	2.7739	0.0857
	LSCS ₂	3072	4.51	1.86	−1.36	8.84	0.07	2.9304	0.2067
	LSCS ₃	2501	4.62	1.87	−1.64	8.82	0.06	3.0061	0.2048

RS, Romanian Spotted breed; RB, Romanian Brown breed; LSCS_{1 to 3}, lactation average of somatic cell scores for 1st to 3rd lactations; N, number of test-day records; SD, standard deviation; Min, minimum score; Max, maximum score; h^2 , heritability; V_P , phenotypic variance; V_A , additive genetic variance.

Estimates of heritability for the SCS showed different variation between breeds over the 305 days of lactation. The range of SCS heritability values for L₁–L₃ in RS (0.08, 0.10 and 0.11) was higher than compared to RB (0.03, 0.07 and 0.06). The highest heritability was recorded for RS cattle in the third lactation (0.11). Between breeds, a distinct pattern of variation in heritability values was observed. The shape of the SCS heritability curves (Supplementary Figure S2) was similar for L₁–L₃ in the RS breed, with the highest values in the first and last parts of the lactation. Conversely, the analysis of daily SCS heritabilities in the RB breed revealed different shapes of curves between the first, second and third lactation. This could be due to the smaller number of RB animals included in the study, which means that possible outlier individuals had a larger impact on the breed-specific population average. The SCS heritabilities in RB decreased in the first part of the lactation (up to 110 days in milk) and then remained on the plateau in mid-lactation (up to 200 days in milk), after which it gradually increased in the last part of the lactation (up to 305 days in milk). This trend appears similar to that of the incidence of mastitis.

3.2. Quality Control

After computing descriptive statistics for the phenotypic data, quality control checking of genotyping data was the second step in our data analysis. Since the success of a genome-wide association study is considerably influenced by appropriate study design and data analysis workflow [27], extra attention was paid towards data quality control of genotyping data. After eliminating markers from unplaced contigs and sex chromosomes, markers with unknown chromosomes or positions within the chromosome, multiallelic markers, indels and duplicate SNPs from the original loci of the microarray, a number of 59,915 markers for 723 animals entered the QC phase. From these, a number of animals and SNPs were discarded as follows: 11 animals that had greater than 5% missing SNP calls; 795 SNPs with call rates <95%; 17,701 SNPs with MAF <0.05; and 1114 SNPs with genotypes that are not in accordance with the Hardy–Weinberg equilibrium ($p > 10^{-6}$). Finally, after the quality control analysis, a data set of 40,305 SNPs across the entire bovine genome from 690 cows (571 RS and 119 RB) was used in the subsequent analysis to carry out genotype association tests.

3.3. Principal Component Analysis

The animals involved in this study came from two breeds and three breeding herds. Thus, in order to assess population stratification, we performed principal component analysis. The PCA plot revealed a clear population structure for the animals in the two cattle breeds included in the study, showing a clear separation between individuals, according to breed (Figure 1). Clusters of the same color represent individuals from the same breed. All animals from the RB breed clustered together in all pairwise scatter plots of the first three principal components, but they were separated from the RS individuals. From the three plots in Figure 1, we observed that PC1 separates the individuals from the two breeds, whereas PC2 and PC3 divide individuals within the RS breed. The individuals from the RS breed split into several subgroups that were best observed in the scatterplot of PC2 and PC3 (Figure 1b). We have conducted a separate investigation into the origins of these

different clusters (data not shown) and concluded that one subgroup consists of cows with paternity from one single sire and the rest of the subgroups consist of cows from several different sires.

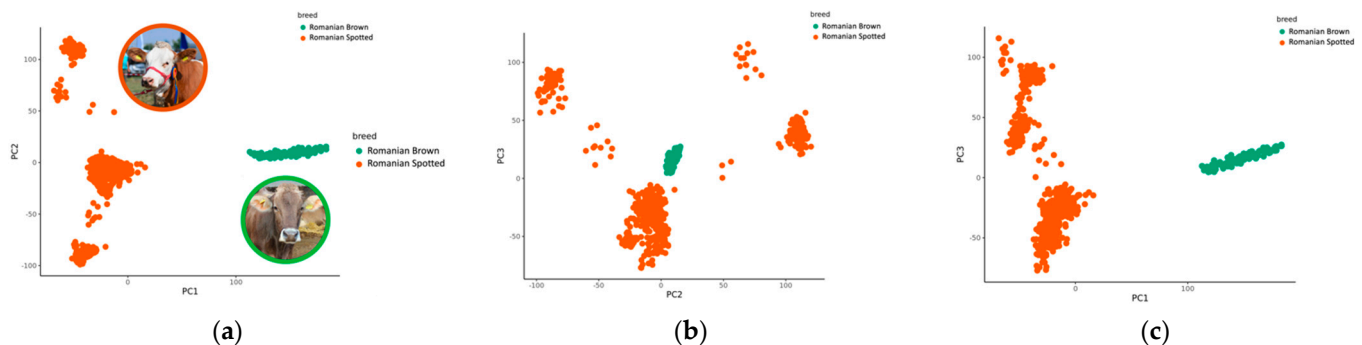


Figure 1. Population structure from the principal component analysis of the 40,305 single nucleotide polymorphisms (SNPs) and 690 cattle. Population structure is presented as pairwise scatter plots (a–c) of the first three principal components (PC), with green and orange dots representing the two breeds (Romanian Spotted in orange and Romanian Brown in green). (a) PC1 vs. PC2; (b) PC2 vs. PC3; (c) PC1 vs. PC3.

3.4. Significant SNPs Associated with SCS

To investigate the genetic variation that underlies the SCS, GWAS was performed in order to identify the associated SNP loci in cattle. The Manhattan plots for SCS in the RS and RB breeds are shown in Figure 2. The upper horizontal lines represent the Bonferroni-adjusted genome-wide significance threshold $-\log_{10}(p) \geq 5.89$, and the inferior horizontal lines represent the suggestive threshold of $-\log_{10}(p) \geq 4.00$, which was set to report further significant associations that were not observed under Bonferroni corrections. The suggestive threshold was chosen according to previous studies where researchers defined a suggestive threshold between $p < 10^{-5}$ and $p < 10^{-3}$ [28,29]. The candidate genes were detected by verifying whether significant or suggestive SNPs overlapped a gene or were located within 1 MBp upstream or downstream from a gene. The genomic positions were based on the UMD3.1 genome assembly of *Bos taurus* [7].

The genome-wide SNPs detected for SCS among parities (L_1 – L_3) are presented in Table 2 for the RS breed and Table 3 for the RB breed. A total number of 41 SNPs were detected in the RS and RB breeds, of which 40 SNPs passed the suggestive threshold and, ultimately, one SNP passed the significance threshold after Bonferroni correction and was associated with SCS in the RS breed in L_3 (Figure 2c, left side). As shown in Table 2, for the RS breed, 3 out of 15 SNPs were located near three known genes, and one SNP overlapped the *HERC3* gene. Of the fifteen SNPs in the RS breed, AX-106761943 (rs110749552) was located on chromosome 6 within the *HERC3* gene (HECT and RLD domain containing E3 ubiquitin protein ligase 3) and was the most significant SNP for SCS ($\log_{10}(p) = 6.37$). In the RB breed, the genome-wide analysis detected 26 SNPs that reached genome-wide significance for association with SCS (Table 3). Among these SNPs, 14 out of 26 SNPs were located near 12 known genes; two SNPs, AX-106741653 and AX-115114947, were located near the *AKAP8* gene, and two other SNPs, AX-106735825 (rs43585636) and AX-117085949 (rs43585209) were located near the *COL12A1* gene, respectively. None of the 15 detected SNPs in the RS breed were in common with the detected 26 SNPs that reached genome-wide significance for association with SCS in the RB breed.

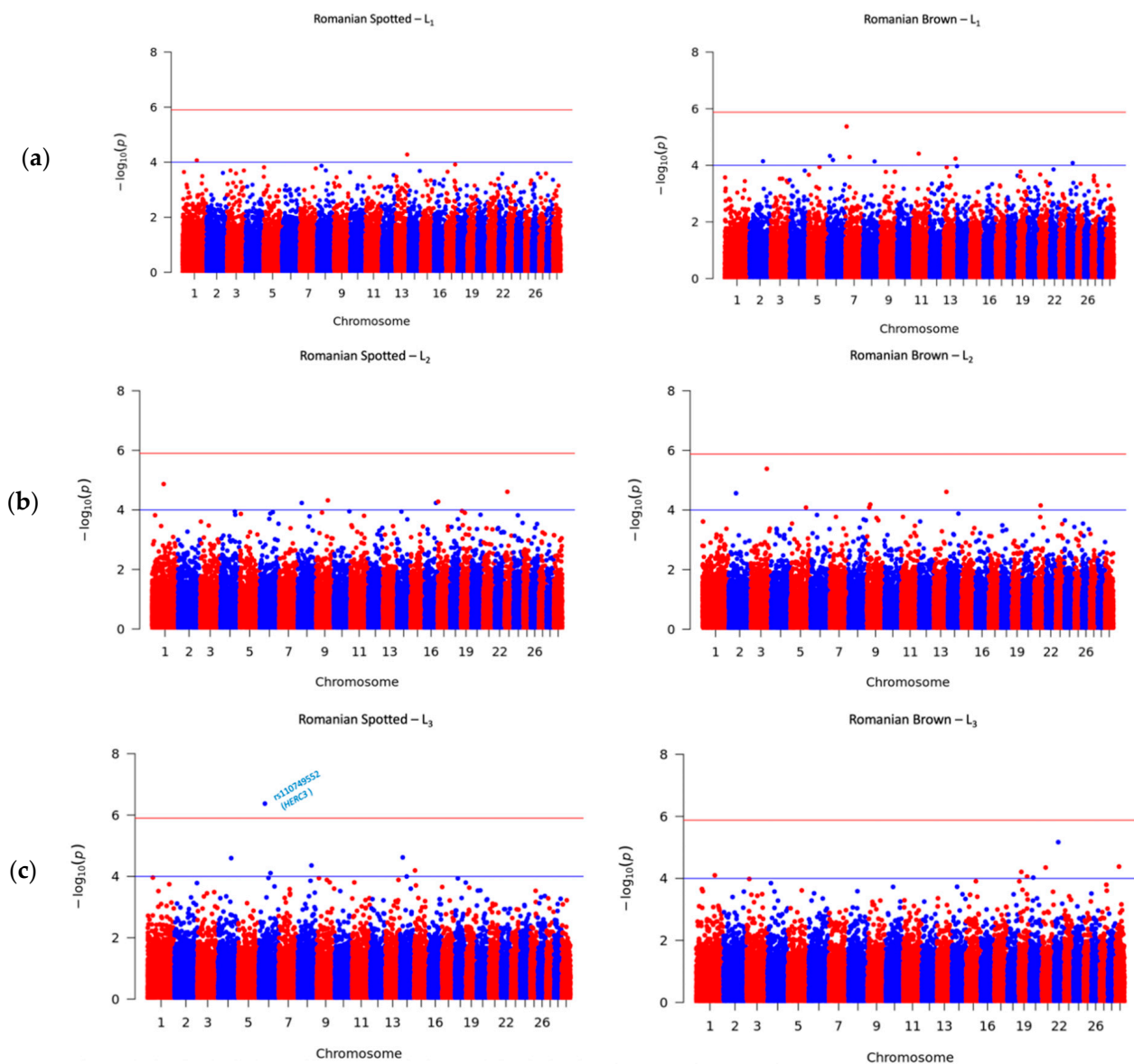


Figure 2. Manhattan plots for somatic cell score (SCS) in the Romanian Spotted (left) and Romanian Brown breed (right). The blue line indicates the suggestive p value threshold of $-\log_{10}(p) \geq 4.00$. The red line indicates the Bonferroni genome-wide significance p value threshold at $-\log_{10}(p) \geq 5.89$. The y -axis shows the $-\log_{10}(p)$ of 40,305 SNPs, and the x -axis shows the chromosomal positions. (a) First lactation; (b) Second lactation; (c) Third lactation.

Table 2. List of single nucleotide polymorphisms (SNPs) in the *Bos taurus* autosomes showing significant ($-\log_{10}(p) \geq 4.00$) associations with SCS in the Romanian Spotted breed.

Parity ¹	Informative SNP ²	SNP rsID	Chr:Position ³	A1	A2	SNP Effect ⁴	$-\log_{10}(p)$	Nearest Gene ⁵	Distance (bp)
L ₁	AX-115117070	NA	13:78383148	G	A	-4.74×10^{-8}	4.28	NA	NA
	AX-117085597	rs137805472	1:95119031	C	T	5.34×10^{-8}	4.07	NA	NA
L ₂	AX-106721594	NA	1:71370844	T	C	3.07×10^{-8}	4.87	ZDHHC19	25909
	AX-106755404	NA	23:17013312	A	G	2.21×10^{-8}	4.61	NA	NA
	AX-106740205	rs109232438	9:71884731	C	T	5.01×10^{-8}	4.32	NA	NA
	AX-171465786	rs110140732	17:3030264	G	A	9.27×10^{-8}	4.28	NA	NA
	AX-117088706	rs43209122	16:70728849	A	G	-2.50×10^{-8}	4.23	NA	NA
	AX-106761299	NA	8:25435054	C	T	-1.70×10^{-8}	4.23	NA	NA

Table 2. Cont.

Parity ¹	Informative SNP ²	SNP rsID	Chr:Position ³	A1	A2	SNP Effect ⁴	$-\log_{10}(p)$	Nearest Gene ⁵	Distance (bp)
L ₃									
	AX-106761943	rs110749552	6:37526622	G	A	5.47×10^{-8}	6.37	HERC3	within
	AX-106728871	rs29021886	14:17674401	C	A	-1.20×10^{-8}	4.62	NA	NA
	AX-124381671	NA	4:79254806	A	G	1.20×10^{-8}	4.6	NA	NA
	AX-106740778	rs42627158	8:81747455	C	T	-7.00×10^{-8}	4.36	DAPK1	459784
	AX-115112140	NA	15:5927942	A	G	3.73×10^{-8}	4.19	MMP7	462217
	AX-185121504	rs209378984	6:70369168	A	G	5.97×10^{-8}	4.11	NA	NA
	AX-115108867	NA	14:41169239	A	G	6.76×10^{-9}	4	NA	NA

¹ Parity 1 (L₁), 2 (L₂) and 3 (L₃). ² SNPs are sorted in descending order with respect to the significance of the association analysis between each SNP and SCS. ³ SNP position based on the UMD3.1 genome assembly of *Bos taurus*. ⁴ SNP effect is calculated relative to the A2 allele (if the effect is positive, then A2 increases the phenotype relative to A1). ⁵ Nearest gene: genes downstream/upstream relative to the SNP, within a ± 1 MBp range (2 MBp total). NA: not available; SNP: single nucleotide polymorphism; Chr: chromosome; A1 and A2: allele 1 and 2.

Table 3. List of SNPs in the *Bos taurus* autosomes showing significant ($-\log_{10}(p) \geq 4.00$) associations with SCS in the Romanian Brown breed.

Parity ¹	Informative SNP ²	SNP rsID	Chr:Position ³	A1	A2	SNP Effect ⁴	$-\log_{10}(p)$	Nearest Gene ⁵	Distance (bp)
L ₁									
	AX-106741653	NA	7:8272794	C	A	6.14×10^{-9}	5.37	AKAP8	514640
	AX-115114947	NA	7:8276425	G	A	6.14×10^{-9}	5.37	AKAP8	511009
	AX-106739297	rs108991944	11:37261133	A	C	1.54×10^{-9}	4.41	CLHC1	526855
	AX-124375018	rs42894728	6:20276795	G	A	7.39×10^{-9}	4.33	NA	NA
	AX-171466608	rs42228650	7:27601185	G	A	5.65×10^{-10}	4.29	MEGF10	219169
	AX-106727011	NA	13:76008898	G	A	-3.96×10^{-9}	4.24	NA	NA
	AX-117090125	rs110642171	6:39688028	A	G	-2.65×10^{-9}	4.19	NA	NA
	AX-117084655	rs109189476	2:88162563	C	T	6.31×10^{-10}	4.14	SATB2	86169
	AX-106731475	NA	8:75798051	T	C	2.67×10^{-9}	4.14	NA	NA
	AX-124381310	NA	24:34453446	A	G	-1.98×10^{-9}	4.08	GATA6	96211
L ₂									
	AX-124381305	rs43358795	3:97560648	A	G	-3×10^{-10}	5.38	SPATA6	809734
	AX-169513997	rs133252983	13:77447650	G	T	5.07×10^{-9}	4.61	NA	NA
	AX-117080179	rs110709131	2:46155190	G	A	3.47×10^{-9}	4.56	NA	NA
	AX-117083483	rs133489631	2:46126141	T	C	3.47×10^{-9}	4.56	NA	NA
	AX-124350324	rs42797639	9:20734430	C	T	-9.3×10^{-10}	4.19	NA	NA
	AX-124377880	rs41981703	21:40381539	A	C	1.02×10^{-8}	4.15	NA	NA
	AX-106735825	rs43585636	9:14745626	C	A	3.02×10^{-9}	4.09	COL12A1	123884
	AX-117085949	rs43585209	9:14767890	A	G	3.02×10^{-9}	4.09	COL12A1	101620
	AX-106726354	NA	5:94517415	A	G	5.58×10^{-9}	4.08	EPS8	156339
L ₃									
	AX-115109333	NA	22:26294371	A	G	6.59×10^{-9}	5.17	NA	NA
	AX-124375014	rs110955314	29:19740081	C	A	-1.57×10^{-8}	4.38	LUZP2	519687
	AX-106753982	NA	21:23675656	A	G	-6.09×10^{-9}	4.35	RAMAC	30303
	AX-106735391	rs110359025	19:16700029	C	T	-7.19×10^{-9}	4.21	NA	NA
	AX-106735614	rs43256975	1:108154057	A	G	-3.37×10^{-9}	4.1	IL12A	131225
	AX-115115579	rs110825674	19:50547082	T	C	-9.06×10^{-9}	4.06	NA	NA
	AX-106730586	rs109619272	20:22823334	A	G	4.48×10^{-9}	4.03	ANKRD55	184056

¹ Parity 1 (L₁), 2 (L₂) and 3 (L₃). ² SNPs are sorted in descending order with respect to the significance of the association analysis between each SNP and SCS. ³ SNP position based on the UMD3.1 genome assembly of *Bos taurus*. ⁴ SNP effect is calculated relative to the A2 allele (if the effect is positive, then A2 increases the phenotype relative to A1). ⁵ Nearest gene: genes downstream/upstream relative to the SNP, within a ± 1 MBp range (2 MBp total). NA: not available; SNP: single nucleotide polymorphism; Chr: chromosome; A1 and A2: allele 1 and 2.

3.5. Linkage Disequilibrium (LD) Blocks of the Significant SNPs

We performed the linkage disequilibrium analysis (r^2 value) on chromosome 7 in the RB breed for a region of 1 Mb interval (from 7,772,794 to 8,772,794 Bp), which covers seventeen SNPs (Figure 3a). High LD blocks were considered for SNPs with $r^2 \geq 0.8$. Four high LD blocks were revealed where the most significant SNP (AX-106741653) was located in a block along with another significant SNP (AX-115114947). The two significant SNPs in block 2 were located upstream of the *AKAP8* gene at distances of 514,640 and 511,009 Bp, respectively. A second linkage disequilibrium analysis for a total of twenty SNPs on chromosome 9 in the RB breed was performed and showed three LD blocks (Figure 3b), each of them including two SNPs, of which the SNPs in block 2 were located upstream of the *COL12A1* gene. For the RS breed, we performed a single LD analysis on chromosome 6 in a region of $\pm 500,000$ Kbp relative to the position 37,526,622 bp of the SNP AX-106761943 (rs110749552), which was the only marker in our study that had a significant association according to Bonferroni correction. For the RS breed, the LD analysis revealed three high LD blocks (Figure 3c), of which the first block contained the previously mentioned SNP AX-106761943 and a non-significant SNP AX-117085324 (rs109478631), both of them located within the *HERC3* gene.

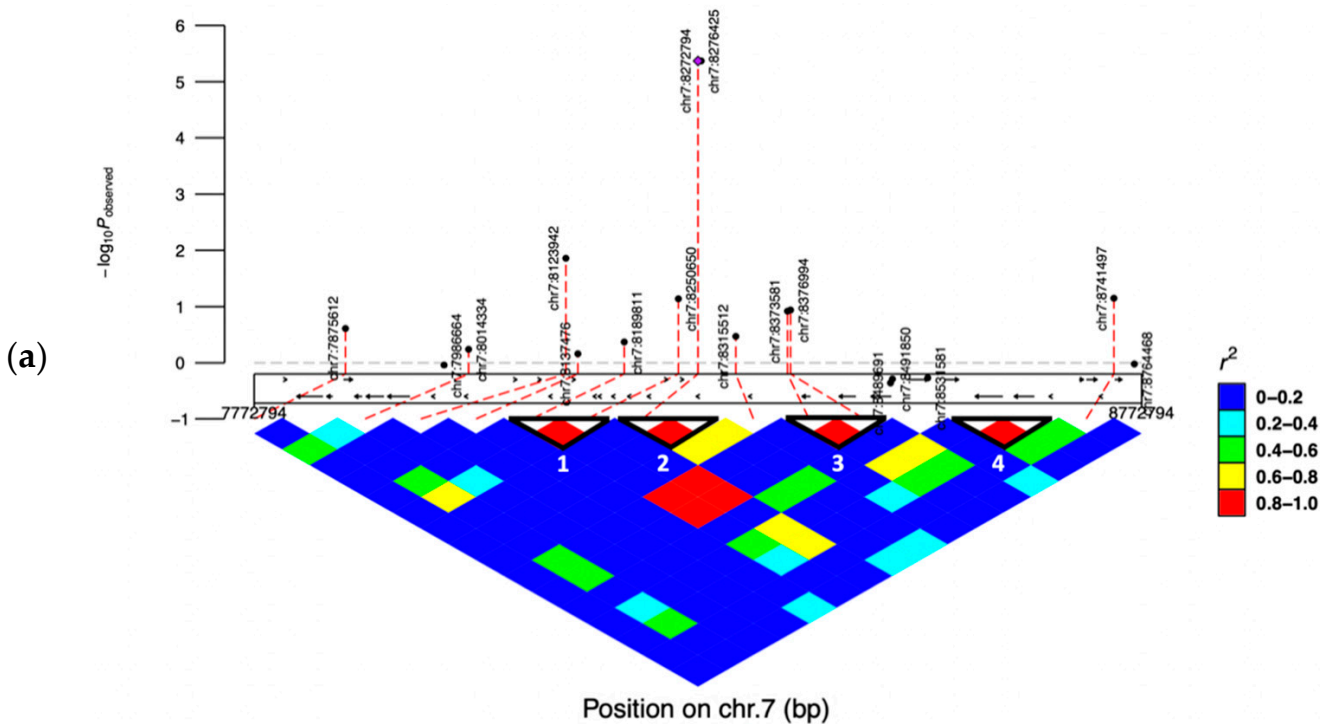


Figure 3. Cont.

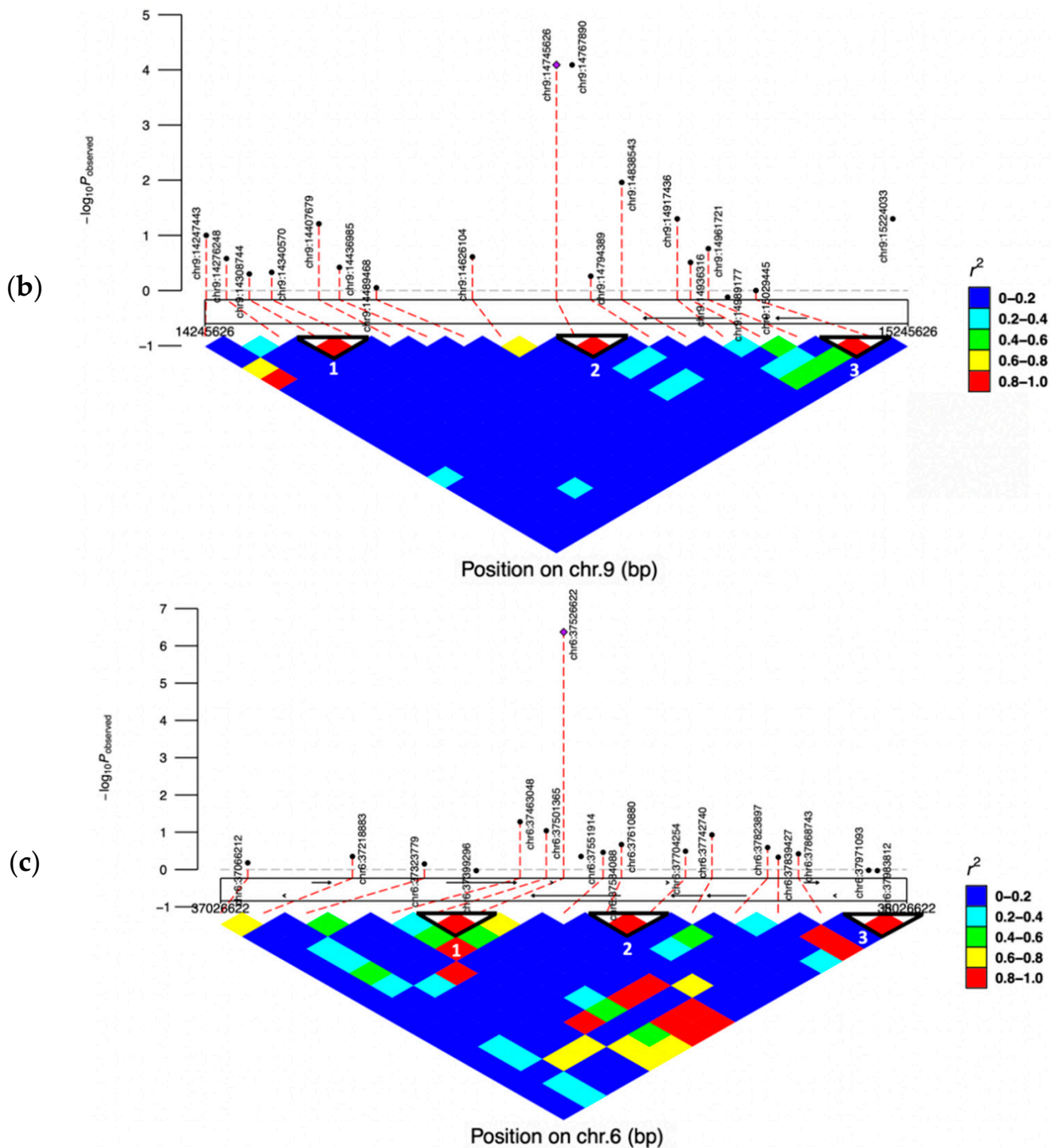


Figure 3. Linkage disequilibrium (LD) analysis. Solid line triangles refer to blocks of high linkage disequilibrium. The color of one square refers to the LD level (r^2) between two SNPs. (a) LD pattern for 17 SNPs in the region between 7,772,794 to 8,772,794 Bp on chromosome 7; (b) LD pattern for 20 SNPs in the region between 14,245,626 to 15,245,626 Bp on chromosome 9; (c) LD pattern for 17 SNPs in the region between 37,026,622 to 38,026,622 Bp on chromosome 6.

4. Discussion

Since mastitis is one of the most frequent diseases in dairy cattle negatively affecting milk production and animal welfare, contributions to understanding the genomic architecture are of great interest. Mastitis, on the other hand, is exceedingly difficult to genetically manage due to its low heritability. However, as a result of the genetic correlation between mastitis incidence and SCC, the latter indicator trait has been used as an indirect technique

for predicting udder infections [30]. Moreover, SCC is recorded in most cases through the Official Dairy Control service in all developed countries and is used as a marker to monitor the prevalence of mastitis in dairy cattle [31], thus being more accessible to farmers compared to microbiological tests. In the present study we have performed GWAS and detected SNP variants associated with SCS variation in native Romanian dairy cattle breeds. To our knowledge, this is the first research study that identified potential causative genes and reported associated SNP variants linked to mastitis in Romanian cattle breeds.

In our study, the overall mean \pm SD SCS values for the RS and RB breeds were 2.99 ± 1.82 and 4.55 ± 1.89 , respectively. These averages are in agreement with previous studies. For instance, Alam et al. [32] reported means of SCS for five lactations of Holstein cows to range between 3.50 and 4.32. Zavadilová et al. [33], using records from the first three lactations of Fleckvieh cows, reported mean SCS of 3.16 to 4.01. The average SCS in this study increased for later lactations, and the trends were similar for both breeds, with the lowest values of SCS during the first and second lactation and higher values during the third lactation. The increasing trend of SCS with the increment of lactation rank observed in our study was also reported in other articles [32,34,35]. Differences in SCS suggest that increased milk flow may contribute to an increase in SCS for later lactations.

The average heritabilities for SCS over the entire days-in-milk range for 305 days were small to moderate and varied from 0.03 (RB) to 0.11 (RS), which are in agreement with previous studies. For example, for the Simmental breed, Wei et al. [36] reported values for SCS heritability of 0.09, whereas Mancin et al. [37] found a slightly higher heritability of 0.133 in Alpine Grey cattle, a local dual-purpose breed from the Italian Alps. However, substantial variations in the heritability of SCS were reported in different studies for distinct breeds. Previous estimates of heritability were low in Jersey and Holstein cattle (0.01–0.06) [38,39], moderate in Finnish Ayrshire (0.07–0.12) [40] and Valdostana cattle (0.06) [41] and higher values in Holstein-Friesian cattle were estimated (0.19–0.21) [42,43]. Various studies have resulted in the conclusion that several factors, including populations, statistical models and estimation methods, can account for these discrepancies in heritabilities [33,44].

Genome-wide association studies provide a valuable source of information for the comprehension of genetic mechanisms that point out the various phenotypes in dairy cattle. Nevertheless, when looking at the GWAS results so far, it is clear that identifying specific SNPs and genes located around them linked to SCS is difficult. As observed in the present study (Tables 2 and 3), from a total of 41 SNPs that were found to be associated with SCS in both Romanian cattle breeds, the majority of them were dissimilar from significant SNPs reported in previous publications [8,10,11,29,45,46], while the *HERC3*, *LUZP2*, *AKAP8* and *MEGF10* genes located near significant SNPs on chromosome 6, 29 and 7 are consistent with other studies [47–49], which strongly suggests that these genes are candidate genes for SCS.

In the RS breed, the GWAS results pointed out 15 SNPs associated with SCS. Out of the significant SNPs, three markers were close to the *ZDHHC19*, *DAPK1* and *MMP7* genes, while one SNP overlapped the *HERC3* gene. The most significant SNP in our study, rs110749552, (6:g.36100172G > A on ARS-UCD1.2; $-\log_{10}(p) = 6.36$) associated with SCS in RS breed in L₃ was located within the *HERC3* gene that encodes a member of the HERC ubiquitin ligase family. Significant associations of the *HERC3* gene with different traits have been previously reported by several authors in cattle or sheep populations. This gene was previously found to be related to mastitis and somatic cell count in Ethiopian cattle populations [47]. According to Jiang [50], the *HERC3* gene was associated with milk protein percentage in Chinese Holstein cattle. In French Holstein, the *HERC3* gene had highly significant effects on α s1-casein [51]. Porto-Neto et al. [52] explored the genetic architecture of ten traits in tropical cattle breeds and reported the *HERC3* as a candidate gene for yearling weight. In addition, Al Kalaldehy [53] showed that the *HERC3* gene on *Ovis aries* chromosome 6 fell within the QTL region that underlies genetic variation for parasite resistance in sheep and also noticed that the genes from the HERC family of ubiquitin ligases are associated with antigen processing; ubiquitination and proteasome

degradation; immune system; class I MHC mediated antigen processing and presentation; and adaptive immune system pathways. However, this is the second study where this gene has been found to be involved in the genetic makeup of mastitis.

With regards to the RB breed, the GWAS results revealed 26 promising SNPs associated with SCS. Out of the significant SNPs, 14 were located near 12 known genes, including *AKAP8*, *CLHC1*, *MEGF10*, *SATB2*, *GATA6*, *SPATA6*, *COL12A1*, *EPS8*, *LUZP2*, *RAMAC*, *IL12A* and *ANKRD55*. On chromosome 7, the rs42228650 SNP was close to *MEGF10*, a gene that was found to be associated with SCS in dairy cows in a previous investigation [48] and was considered as a candidate gene associated with residuals for fecal egg counts of gastrointestinal nematodes in German Black Pied cattle in a recent study [54]. On the same chromosome, two significant SNPs (AX-106741653 and AX-115114947) in the RB breed were located close to the *AKAP8* gene, which was associated with somatic cell count in seven different populations of dairy cows in a complex study that used the combined approach of examining previous results from differential gene expression analysis and GWAS studies to identify the key pathways and candidate genes that potentially confer genetic resistance to mastitis in dairy cows [48]. Furthermore, another relevant gene revealed in our work located 519,687 Bp from the rs110955314 SNP was *LUZP2* on chromosome 29, which has been associated with SCS [48] and milk fatty acid traits in cattle [55].

In addition to the genes known to be associated with mastitis, the present study identified novel candidate genes associated with SCS in Romanian cattle. Some of these genes have been shown to be linked with other traits in ruminants. For instance, the rs109189476 SNP on chromosome 2 was close to *SATB2*, a gene that encodes a DNA binding protein that specifically binds nuclear matrix attachment regions. A previous study disclosed the *SATB2* gene to be associated with tick burden, and the author mentioned a potential link of this gene to the immune system [56]. Jiang et al. [57] conducted a large-scale GWAS using first-lactation Holstein cows and identified the 91.13–94.62 Mb region of chromosome 5, including the *EPS8* gene associated with fat yield, covered by SNP AX-106726354 located near *EPS8* in this study. The SNP AX-124381310 on chromosome 24 was close to the *GATA6* gene. The GATA gene family has six members (*GATA1*–*GATA6*) in vertebrates and plays a critical role in controlling the growth, differentiation and survival of diverse cells, as well as maintaining the body's functions [58]. Currently, mutations in *GATA6* were reported to be associated with microtia in Awassi sheep in ruminants [59].

Regarding the *COL12A1* (Collagen Type XII α 1 Chain) gene, it is known that degradation of the extracellular matrix and collagen chain trimerization is common among the related pathways of this gene. The Gene Ontology annotations associated with this gene include extracellular matrix structural constituent conferring tensile strength. Collagens consist of 28 members in vertebrates, and they are the most abundant extracellular matrix proteins [60]. Chen et al. [61] showed that gene expression levels of a collagen member (*COL4A1*) were significantly downregulated in the inflammation-associated fibroblasts from bovine mammary glands compared to normal fibroblasts, suggesting an implication of collagen members in immune response since one of the related pathways include extracellular matrix–receptor interaction. Moreover, Lu et al. [60] revealed that several collagen genes in bovine had higher expression levels during lactation, among which *COL12A1* is upregulated in late lactation, suggesting that the collagen genes had varied biological functions at different phases of lactation. In the present study, we identified two significant SNPs associated with SCS (rs43585636 and rs43585209) located close to *COL12A1*. All of the aforementioned aspects suggest that the *COL12A1* gene might be involved in SCS.

This study revealed that different markers and candidate genes were found for the two breeds in our study. In addition, all identified SNPs have been distinct among the three parities, denoting that SCS is influenced by different genes according to parity. The different sets of markers discovered in this work compared to studies reported in the literature can be attributed to various factors. We can assume that the power of detecting SNPs can be lower in the RB breed compared to the RS breed as a consequence of the comparatively lower number of genotyped RB animals used in the analysis. Furthermore,

previously known mastitis related genes such as *LUZP2*, *AKAP8* and *MEGF10* were also identified in our study as candidate genes for SCS in the RB breed. For this specific case in the RB breed, further studies using larger sample sizes should be performed in order to validate previous results. Finally, an additional factor that may influence the different sets of candidate SNPs/breed is the distinct genetic backgrounds of the two breeds as they have different levels of clustering, as shown in previous studies on Romanian cattle breeds [62,63].

The specific novel SNPs and candidate genes located around them reported in the present study can be considered as candidates involved in SCS and mastitis resistance; however, these need to be kept in perspective, and polymorphisms in those genes should be further analyzed to highlight whether they influence the ability of dairy cows to resist mastitis.

5. Conclusions

Our study represents the first GWAS for SCS in Romanian dairy cattle and, thus, provides new perspective into the genetic architecture of udder infections in these native dairy cattle. We identified 41 SNPs and detected their significant associations with variation in SCS in dairy cattle. In both breeds, the SNPs and position of association signals were distinct among the three parities, denoting that mastitis is influenced by different genes according to parity. The results contribute to an increase in knowledge regarding the proportion of genetic variability explained by SNPs for SCS in dairy cattle and, particularly, in Romanian native cattle. However, further large-scale studies of a vast number of native dairy cattle breeds are needed in order to investigate others markers and genes that could be involved in mastitis.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12101495/s1>, Figure S1: First 20 principal component analysis (PCA) of Romanian Spotted (RS) and Romanian Brown (RB) cattle. Figure S2: Heritabilities (in red) and repeatabilities (in blue) of somatic cell scores among parities (L1–L3) for Romanian Spotted (RS) and Romanian Brown (RB) cattle.

Author Contributions: Conceptualization, D.E.I. and A.E.M.; methodology, D.E.I., A.E.M. and I.H.; software, A.E.M., D.Z. and M.C.; formal analysis, D.E.I., A.E.M. and I.H.; investigation, D.E.I., A.E.M., C.V.M., R.I.N., G.V.G., O.I.G., D.Z. and I.H.; resources, D.E.I., A.E.M. and I.H.; data curation, D.E.I., A.E.M. and I.H.; writing—original draft preparation, D.E.I. and A.E.M.; writing—review and editing, D.E.I., A.E.M., C.V.M., R.I.N., G.V.G., O.I.G., D.Z., M.C. and I.H.; visualization, D.E.I. and A.E.M.; supervision, D.E.I. and I.H.; project administration, D.E.I., A.E.M. and I.H.; funding acquisition, A.E.M. and I.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the following: Romanian Ministry of Agriculture and Rural Development through the project ADER 8.1.6/2019 and UEFISCDI, grant number 7PCCDI/2018, cod PN-III-P1-1.2-PCCDI-2017-0361.

Institutional Review Board Statement: This study was approved by the Ethics Committee of Research and Development Station for Bovine (Approval number 88/04.10.2019).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be available upon request.

Acknowledgments: The authors would like to thank the Romanian Breeding Association “Băltață Românească” Simmental type (ACVBR-SIM Harman-Brasov, Romania) and Innovative Agricultural Services (Reading, UK) for providing all phenotypic data and pedigree information for Romanian Spotted and Romanian Brown breeds, respectively.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Rupp, R.; Boichard, D. Genetics of resistance to mastitis in dairy cattle. *Vet. Res.* **2003**, *34*, 671–688. [\[CrossRef\]](#)
- Aarestrup, F.M.; Jensen, N.E.; Ostergård, H. Analysis of associations between major histocompatibility complex (BoLA) class I haplotypes and subclinical mastitis of dairy cows. *J. Dairy Sci.* **1995**, *78*, 1684–1692. [\[CrossRef\]](#)
- Dietz, A.B.; Cohen, N.D.; Timms, L.; Kehrl, M.E., Jr. Bovine lymphocyte antigen class II alleles as risk factors for high somatic cell counts in milk of lactating dairy cows. *J. Dairy Sci.* **1997**, *80*, 406–412. [\[CrossRef\]](#)
- Kelm, S.C.; Detilleux, J.C.; Freeman, A.E.; Kehrl, M.E., Jr.; Dietz, A.B.; Fox, L.K.; Butler, J.E.; Kascovics, I.; Kelley, D.H. Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. *J. Dairy Sci.* **1997**, *80*, 1767–1775. [\[CrossRef\]](#)
- Gilliespie, B.E.; Jayarao, B.M.; Dowlen, H.H.; Oliver, S.P. Analysis and frequency of bovine lymphocyte antigen DRB3.2 alleles in Jersey cows. *J. Dairy Sci.* **1999**, *82*, 2049–2053. [\[CrossRef\]](#)
- Takeshima, S.; Nakai, Y.; Ohta, M.; Aida, Y. Characterization of DRB3 alleles in the MHC of Japanese Shorthorn Cattle by polymerase chain reaction-sequence-based typing. *J. Dairy Sci.* **2002**, *85*, 1630–1632. [\[CrossRef\]](#)
- Zimin, A.V.; Delcher, A.L.; Florea, L.; Kelley, D.R.; Schatz, M.C.; Puiu, D.; Hanrahan, F.; Pertea, G.; Van Tassell, C.P.; Sonstegard, T.S.; et al. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* **2009**, *10*, R42. [\[CrossRef\]](#) [\[PubMed\]](#)
- Meredith, B.K.; Kearney, F.J.; Finlay, E.K.; Bradley, D.G.; Fahey, A.G.; Berry, D.P.; Lynn, D.J. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genet.* **2012**, *13*, 21. [\[CrossRef\]](#)
- Zhou, J.; Liu, L.; Chen, C.J.; Zhang, M.; Lu, X.; Zhang, Z.; Huang, X.; Shi, Y. Genome-wide association study of milk and reproductive traits in dual-purpose Xinjiang Brown cattle. *BMC Genom.* **2019**, *20*, 827. [\[CrossRef\]](#)
- Meredith, B.K.; Berry, D.P.; Kearney, F.; Finlay, E.K.; Fahey, A.G.; Bradley, D.G.; Lynn, D.J. A genome-wide association study for somatic cell score using the Illumina high-density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front. Genet.* **2013**, *4*, 229. [\[CrossRef\]](#)
- Strillacci, M.G.; Frigo, E.; Schiavini, F.; Samoré, A.B.; Canavesi, F.; Vevey, M.; Cozzi, M.C.; Soller, M.; Lipkin, E.; Bagnato, A. Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA. *BMC Genet.* **2014**, *15*, 106. [\[CrossRef\]](#)
- Nani, J.P.; Raschia, M.A.; Poli, M.A.; Calvinho, L.F.; Amadio, A.F. Genome-wide association study for somatic cell score in Argentinean dairy cattle. *Livest. Sci.* **2015**, *175*, 1–9. [\[CrossRef\]](#)
- The Animal Quantitative Trait Loci (QTL) Database. Available online: <https://www.animalgenome.org/cgi-bin/QTLdb/BT/index> (accessed on 20 May 2021).
- Breeds of Livestock, Department of Animal Science. Breeds of Livestock—Baltata Romaneasca Cattle. Available online: <http://afs.okstate.edu/breeds/cattle/baltataromaneasca/index.html/> (accessed on 25 August 2021).
- Wiggans, G.R.; Shook, G.E. A lactation measure of somatic-cell count. *J. Dairy Sci.* **1987**, *70*, 2666–2672. [\[CrossRef\]](#)
- Interbull. Code of Practice. Available online: <https://interbull.org/ib/codeofpractice> (accessed on 25 August 2021).
- Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [\[CrossRef\]](#)
- Morgan, A.P. argyle: An R package for analysis of Illumina genotyping arrays. *G3* **2015**, *6*, 281–286. [\[CrossRef\]](#) [\[PubMed\]](#)
- Misztal, I.; Tsuruta, S.; Strabel, T.; Auvray, B.; Druet, T.; Lee, D.H. BLUPF90 and related programs (BGF90). In Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 19–23 August 2002; Volume 28.
- Lopes, F.B. DRP R Package. Available online: <https://github.com/camult/DRP> (accessed on 1 August 2021).
- Garrick, D.J.; Taylor, J.F.; Fernando, R.L. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* **2009**, *41*, 55. [\[CrossRef\]](#) [\[PubMed\]](#)
- Endelman, J.B. Ridge regression and other kernels for genomic selection with R package rrBLUP. *TPG* **2011**, *4*, 250–255. [\[CrossRef\]](#)
- Turner, S. qqman: QQ and Manhattan Plots for GWAS Data; R Package Version 0.1.4; 2017. Available online: <https://cran.r-project.org/package=qqman> (accessed on 4 August 2021).
- He, F.; Ding, S.; Wang, H.; Qin, F. IntAssoPlot: An R Package for Integrated Visualization of Genome-Wide Association Study Results With Gene Structure and Linkage Disequilibrium Matrix. *Front. Genet.* **2020**, *11*, 260. [\[CrossRef\]](#)
- Obenchain, V.; Lawrence, M.; Carey, V.; Gogarten, S.; Shannon, P.; Morgan, M. VariantAnnotation: A Bioconductor package for exploration and annotation of genetic variants. *Bioinformatics* **2014**, *30*, 2076–2078. [\[CrossRef\]](#)
- Lawrence, M.; Huber, W.; Pagès, H.; Aboyoun, P.; Carlson, M.; Gentleman, R.; Morgan, M.; Carey, V. Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* **2013**, *9*, e1003118. [\[CrossRef\]](#)
- Marees, A.T.; de Kluiver, H.; Stringer, S.; Vorspan, F.; Curis, E.; Marie-Claire, C.; Derks, E.M. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int. J. Methods Psychiatr. Res.* **2018**, *27*, e1608. [\[CrossRef\]](#)
- Parker Gaddis, K.L.; Null, D.J.; Cole, J.B. Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *J. Dairy Sci.* **2016**, *99*, 6420–6435. [\[CrossRef\]](#) [\[PubMed\]](#)
- Welderufael, B.G.; Løvendahl, P.; de Koning, D.J.; Janss, L.L.G.; Fikse, W.F. Genome-Wide Association Study for Susceptibility to and Recoverability from Mastitis in Danish Holstein Cows. *Front. Genet.* **2018**, *9*, 141. [\[CrossRef\]](#) [\[PubMed\]](#)
- Boettcher, P. Breeding for improvement of functional traits in dairy cattle. *Ital. J. Anim. Sci.* **2005**, *4*, 7–16. [\[CrossRef\]](#)

31. Alhussien, M.N.; Dang, A.K. Milk somatic cells, factors influencing their release, future prospects, and practical utility in dairy animals: An overview. *Vet. World* **2018**, *11*, 562–577. [[CrossRef](#)]
32. Alam, M.; Cho, C.I.; Choi, T.J.; Park, B.; Choi, J.G.; Choy, Y.H.; Lee, S.S.; Cho, K.H. Estimation of Genetic Parameters for Somatic Cell Scores of Holsteins Using Multi-trait Lactation Models in Korea. *Asian-Australas J. Anim. Sci.* **2015**, *28*, 303–310. [[CrossRef](#)]
33. Zavadilová, L.; Wolf, J.; Štípková, M.; Němcová, E.; Jamrozik, J. Genetic parameters for somatic cell score in the first three lactations of Czech Holstein and Fleckvieh breeds using a random regression model. *Czech J. Anim. Sci.* **2011**, *56*, 251–260. [[CrossRef](#)]
34. Martins, A.M.; Silvestre, A.M.; Petim-Batista, M.F.; Colaço, J.A. Somatic cell score genetic parameter estimates of dairy cattle in Portugal using fractional polynomials. *J. Anim. Sci.* **2011**, *89*, 1281–1285. [[CrossRef](#)]
35. Heins, B.J.; Hansen, L.B. Short communication: Fertility, somatic cell score, and production of Normande×Holstein, Montbéliarde×Holstein, and Scandinavian Red × Holstein crossbreds versus pure Holsteins during their first 5 lactations. *J. Dairy Sci.* **2012**, *95*, 918–924. [[CrossRef](#)]
36. Wei, C.; Luo, H.P.; Wang, Y.C.; Huang, X.X.; Zhang, M.H.; Zhang, X.X.; Wang, D.; Ge, J.J.; Xu, L.; Jiang, H.; et al. Analyses of the genetic relationships between lactose, somatic cell score, and growth traits in Simmental cattle. *Animal* **2021**, *15*, 100027. [[CrossRef](#)]
37. Mancin, E.; Sartori, C.; Guzzo, N.; Tuliozi, B.; Mantovani, R. Selection Response Due to Different Combination of Antagonistic Milk, Beef, and Morphological Traits in the Alpine Grey Cattle Breed. *Animals* **2021**, *11*, 1340. [[CrossRef](#)]
38. Rincón, F.J.; Zambrano, A.J.; Echeverri, J. Estimation of genetic and phenotypic parameters for production traits in Holstein and Jersey from Colombia. *Rev. MVZ Córdoba* **2015**, *20*, 4962–4973. [[CrossRef](#)]
39. DeGroot, B.J.; Keown, J.F.; Van Vleck, L.D.; Kachman, S.D. Estimates of genetic parameters for Holstein cows for test-day yield traits with a random regression cubic spline model. *Genet. Mol. Res.* **2007**, *6*, 434–444. [[PubMed](#)]
40. Negussie, E.; Strandén, I.; Mäntysaari, E.A. Genetic association of clinical mastitis with test-day somatic cell score and milk yield during first lactation of Finnish Ayrshire cows. *J. Dairy Sci.* **2008**, *91*, 1189–1197. [[CrossRef](#)]
41. Sartori, C.; Guzzo, N.; Mantovani, R. Genetic correlations of fighting ability with somatic cells and longevity in cattle. *Animal* **2020**, *14*, 13–21. [[CrossRef](#)] [[PubMed](#)]
42. Romano, G.S.; Pinto, L.F.B.; Valloto, A.A.; Horst, J.A.; Pedrosa, V.B. Genetic parameters between somatic cell score and production traits for Holstein cattle in Southern Brazil. *Rev. Colomb. Cienc. Pecu.* **2020**, *33*, 60–70. [[CrossRef](#)]
43. Samoré, A.B.; Groen, A.F.; Boettcher, P.J.; Jamrozik, J.; Canavesi, F.; Bagnato, A. Genetic correlation patterns between somatic cell score and protein yield in the Italian Holstein-Friesian population. *J. Dairy Sci.* **2008**, *91*, 4013–4021. [[CrossRef](#)] [[PubMed](#)]
44. Torshizi, M.E.; Farhangfar, H. The use of dijksra mechanistic model for genetic analysis of the lactation curve characteristics and their relationships with age at first calving and somatic cell score of Iranian dairy cows. *Acta Sci. Anim. Sci.* **2020**, *42*, e50181. [[CrossRef](#)]
45. Abdel-Shafy, H.; Bortfeldt, R.H.; Tetens, J.; Brockmann, G.A. Single nucleotide polymorphism and haplotype effects associated with somatic cell score in German Holstein cattle. *Genet. Sel. Evol.* **2014**, *46*, 35. [[CrossRef](#)]
46. Kurz, J.P.; Yang, Z.; Weiss, R.B.; Wilson, D.J.; Rood, K.A.; Liu, G.E.; Wang, Z. A genome-wide association study for mastitis resistance in phenotypically well-characterized Holstein dairy cattle using a selective genotyping approach. *Immunogenetics* **2019**, *71*, 35–47. [[CrossRef](#)]
47. Pan, D.; Zhang, S.; Jiang, J.; Jiang, L.; Zhang, Q.; Liu, J. Genome-wide detection of selective signature in Chinese Holstein. *PLoS ONE* **2013**, *8*, e60440. [[CrossRef](#)]
48. Chen, X.; Cheng, Z.; Zhang, S.; Werling, D.; Wathes, D. Combining Genome Wide Association Studies and Differential Gene Expression Data Analyses Identifies Candidate Genes Affecting Mastitis Caused by Two Different Pathogens in the Dairy Cow. *Open J. Anim. Sci.* **2015**, *5*, 358–393. [[CrossRef](#)]
49. Wang, X.; Ma, P.; Liu, J.; Zhang, Q.; Zhang, Y.; Ding, X.; Jiang, L.; Wang, Y.; Zhang, Y.; Sun, D. Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC Genet.* **2015**, *16*, 111. [[CrossRef](#)] [[PubMed](#)]
50. Jiang, L.; Liu, J.; Sun, D.; Ma, P.; Ding, X.; Yu, Y.; Zhanh, Q. Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. *PLoS ONE* **2010**, *5*, e13661. [[CrossRef](#)] [[PubMed](#)]
51. Sanchez, M.P.; Govignon-Gion, A.; Croiseau, P.; Fritz, S.; Hozé, C.; Miranda, G.; Martin, P.; Barbat-Leterrier, A.; Letaïef, R.; Rocha, D.; et al. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genet. Sel. Evol.* **2017**, *49*, 68. [[CrossRef](#)] [[PubMed](#)]
52. Porto-Neto, L.R.; Reverter, A.; Prayaga, K.C.; Chan, E.K.; Johnston, D.J.; Hawken, R.J.; Fordyce, G.; Garcia, J.F.; Sonstegard, T.S.; Bolormaa, S.; et al. The Genetic Architecture of Climatic Adaptation of Tropical Cattle. *PLoS ONE* **2014**, *9*, e113284. [[CrossRef](#)] [[PubMed](#)]
53. Al Kalaldehy, M.; Gibson, J.; Lee, S.H.; Gondro, C.; van der Werf, J.H.J. Detection of genomic regions underlying resistance to gastrointestinal parasites in Australian sheep. *Genet. Sel. Evol.* **2019**, *51*, 37. [[CrossRef](#)]
54. Wolf, M.J.; Yin, T.; Neumann, G.B.; Korkuc, P.; Brockmann, G.A.; König, S.; May, K. Genome-Wide Association Study Using Whole-Genome Sequence Data for Fertility, Health Indicator, and Endoparasite Infection Traits in German Black Pied Cattle. *Genes* **2021**, *12*, 1163. [[CrossRef](#)] [[PubMed](#)]

55. Li, C.; Sun, D.; Zhang, S.; Wang, S.; Wu, X.; Zhang, Q.; Liu, L.; Li, Y.; Qiao, L. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. *PLoS ONE* **2014**, *9*, e96186. [[CrossRef](#)]
56. Turner, L.B.; Harrison, B.E.; Bunch, R.J.; Porto Neto, L.R.; Li, Y.; Barendse, W. A genome wide association study of tick burden and milk composition in cattle. *Anim. Prod. Sci.* **2010**, *50*, 235–245. [[CrossRef](#)]
57. Jiang, J.; Ma, L.; Prakapenka, D.; VanRaden, P.M.; Cole, J.B.; Da, Y. A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle. *Front. Genet.* **2019**, *10*, 412. [[CrossRef](#)]
58. Sun, Z.; Pang, S.; Cui, Y.; Yan, B. Genetic and Functional Variants Analysis of the GATA6 Gene Promoter in Acute Myocardial Infarction. *Front. Genet.* **2019**, *10*, 1100. [[CrossRef](#)] [[PubMed](#)]
59. Jawasreh, K.; Al-Omari, H.D. Duplication in ECR near HMX1 and a SNP in GATA6 Genes Regulate Microtia in Awassi Sheep. *Genes* **2020**, *11*, 597. [[CrossRef](#)] [[PubMed](#)]
60. Lu, X.; Duan, A.; Liang, S.; Ma, X.; Deng, T. Genomic Identification, Evolution, and Expression Analysis of Collagen Genes Family in Water Buffalo during Lactation. *Genes* **2020**, *11*, 515. [[CrossRef](#)] [[PubMed](#)]
61. Chen, Q.; He, G.; Zhang, W.; Xu, T.; Qi, H.; Li, J.; Zhang, Y.; Gao, M.Q. Stromal fibroblasts derived from mammary gland of bovine with mastitis display inflammation-specific changes. *Sci. Rep.* **2016**, *6*, 27462. [[CrossRef](#)] [[PubMed](#)]
62. Ilie, D.E.; Cean, A.; Czisster, L.T.; Gavojdian, D.; Ivan, A.; Kusza, S. Microsatellite and Mitochondrial DNA Study of Native Eastern European Cattle Populations: The Case of the Romanian Grey. *PLoS ONE* **2015**, *10*, e0138736. [[CrossRef](#)] [[PubMed](#)]
63. Mizeranschi, A.E.; Neamț, R.I.; Mihali, C.V.; Goilean, V.; Carabaș, M.; Ilie, D.E. Comparative Analysis of Romanian and Swiss Bovine Populations Using Whole Genome Sequencing and SNP Microarrays. *Sci. Pap. Anim. Sci. Biotechnol.* **2020**, *53*, 93–99.