Research Article

ISSN 2771-9014

Journal of Biotechnology and its Applications

Weighted Single-Step Genome-Wide Association Study of Milk Production Traits in the Russian Black-and-White cattle

Khatib Abdulrahman[1]*, Prokhortchouk B Egor[2,3], Sharko S Fedor[2], Mazur Alexander[2,3] and Zarkawi Moataz[1]

[1]*Atomic Energy Commission of Syria (AECS), Department of Agriculture, P.O. Box 6091, Damascus, Syria.*

[2]*Laboratory of vertebrate genomics and epigenomics, Federal Research Centre "Fundamentals of Biotechnology" of the Russian Academy of Sciences, Moscow, 119071, Russian Federation.*

[3]*Laboratory I-Gene, ZAO "Genoanalytics", Moscow, 119234, Russian Federation.*

***Correspondence:**

Khatib Abdulrahman, Atomic Energy Commission of Syria (AECS), Department of Agriculture, P.O. Box 6091, Damascus, Syria.

Received: 23 Aug 2025; **Accepted:** 26 Sep 2025; **Published:** 10 Oct 2025

Citation: Abdulrahman K, Egor PB, Fedor SS, et al. Weighted Single-Step Genome-Wide Association Study of Milk Production Traits in the Russian Black-and-White cattle. J Biotechnology App. 2025; 4(2); 1-8.

ABSTRACT

Weighted single-step genome-wide association study (WssGWAS) has proven effective for exploring genomic regions associated with quantitative traits in dairy cattle. In this study, we conducted WssGWAS to identify candidate genes linked to three of the most frequently assessed milk production traits: daily milk yield (MY), milk fat percentage (FP) and milk protein percentage (PP). The phenotypic and pedigree dataset used in this study was obtained from the national genetic evaluation program in Russia for Black-and-White dairy cattle. Genotypic information for 644 animals (427 bulls and 217 cows) was used to estimate the effects of 52,445 SNPs using WssGWAS. SNPs with call rate < 0.90 and those with minor allele frequency < 0.05 were excluded from the analysis. The results were reported as the proportion of total genetic variance explained by 20 adjacent SNP windows. Only windows that accounted for the largest share of variance (>0.5%) were chosen for additional candidate gene analysis. We identified 12 genomic windows on chromosomes 7, 8 and 12 associated with MY, one genomic window on chromosomes 4 associated with PP and six non-overlapping genomic windows on chromosomes 10 and 14 associated with FP. These windows collectively explained 6.33%, 0.51 % and 6.22% of the total genetic variance for MY, FP and PP, respectively. Different genes were identified within these regions and are believed to be associated with the studied traits. These genes play a role in numerous biological processes, such as metabolism, growth, the preservation of genomic stability, and the control of gene expression.

Keywords

WssGWAS, Milk production traits, Candidate gene, QTL, Black-and-White cattle.

Introduction

Investigating the genetic architecture of intricate quantitative traits, such as milk production characteristics, is essential for enhancing genetic merit in livestock species [1,2].

Conventional breeding techniques utilising phenotypic data and pedigree records have effectively improved economically important traits in livestock for a long time [3,4]. However, the primary disadvantage of these methods is their expense and time consumption [5].

Advancements in genotyping technologies, quantitative trait loci (QTL) mapping, and genetic analysis techniques have elucidated the genetic architecture of numerous significant quantitative traits across various cattle breeds and populations [6].

The identification of genetic markers associated with desirable traits improves genetic enhancement programs by replacing

traditional selection methods with selection at the genomic level [7].

The majority traits of importance in dairy cattle are quantitative, controlled by multiple genes, influenced by environmental factors, and display continuous distributions [8].
The identification of QTL, which alleles influence the phenotypic variation of complex traits, can facilitate the early detection of animals with enhanced genetic merit. These QTL can be incorporated into marker-assisted selection programs, allowing for the simultaneous evaluation of the cumulative effects of multiple QTL on breeding objectives [9].

In recent years, genome-wide association studies (GWAS) have been recognized as a potent tool for this purpose. GWAS link phenotypic and genotypic records to explain the genetic basis of phenotypic variation among animals [10,11]. A more promising method, known as Weighted single-step Genome-Wide Association Study (WssGWAS), has recently been utilized in many studies for the identification of QTL [12,13]. In this method, genotypes, phenotypes, and pedigree information are combined to compute genomic estimated breeding values (GEBVs), which are subsequently used to assess the impact of single nucleotide polymorphisms (SNPs) on trait variation [14]. This procedure involves three main steps. First, GEBVs are predicted using single-step genomic best linear unbiased prediction (ssGBLUP). Then, the weights of all markers are obtained by back-calculating the effects of SNPs, which are then utilized to create the genomic relationship matrix (G) in the subsequent steps, as part of an iterative process [15]. By considering the unequal variances present among markers, WssGWAS facilitates a more accurate estimation of the effects of SNPs and, as a result, allows for a more precise identification of QTL. This method is also effective when only limited genotypic or phenotypic data are available [16]. Recent studies have successfully used this approach for identification a range of economically important traits in farm animals, and it has been successfully applied in dairy cattle populations [17-21].

In Russia, several studies have previously been conducted using GWAS to integrate genomic selection into traditional breeding programs and to detect QTL associated with economically important traits in dairy cattle [22-24]. This study represents a large-scale effort to apply WssGWAS for identification genomic regions and candidate genes linked to three milk production traits: daily milk yield (MY), milk fat percentage (FP), and milk protein percentage (PP) in Russian Black-and-White dairy cattle.

Materials and Methods

Animals, Phenotypes, Genotypes and Pedigree

The phenotypic and genotypic data used in this study were obtained from our previous research [25], which was performed for genomic estimated breeding value of Russian Black-and-White dairy cattle. Briefly, a unified database of phenotypic and pedigree information was developed for animals from 523 farms across 12 regions of the Russian Federation. The established database included phenotypic and pedigree records for 69,131 bulls and 2, 551, 529 cows born between 1990 and 2017. A Multi-step system was implemented to verify the reliability of the data and exclude inaccurate records. The six main steps used for data validation are described in detail in our previous work [25]. The final validated phenotypic database included information on 1,597,426 cows with 4,771,366 completed lactations. The pedigree database contained information on 1,983,031 animals including 51,810 bulls. Genotyping data for 644 animals (427 bulls and 217 cows) were used in this study. DNA was extracted from blood and skin samples using the standard QIAamp® DNA protocol. Genotyping was performed using the BovineSNP50 v3 DNA Analysis BeadChip microarray (Illumina, USA). Only genotypes with a call rate > 90% were included in the study. SNP markers with a minor allele frequency < 5% were excluded from the analysis. In total, 38,054 SNPs remained for further analysis after quality control.

Statistical Analysis

In this study, we used WssGWAS method. This approach involves calculating SNP effects based on GEBVs. The GEBVs for animals were calculated using ssGBLUP method [26]. All analyses were conducted using the BLUPF90 programs [27]. Variance components were initially estimated using AIREMLF90, and these estimates were then used in BLUPF90 to predict the GEBVs. Subsequently, SNP effects were calculated using the postGSf90 software. For WssGWAS, the following single-trait statistical model was applied to each of the studied traits:

$$y = X\,b + Z\,a + W\,p + e$$

Where, y: vector of observations for the studied traits (MY, FP, PP), b: vector of fixed effects, including age of the animal, region-year-season of calving, lactation number, and test-day, a: vector of random animal effects with $a \sim N(0, H\sigma_a^2)$, where σ_a^2 is the additive genetic variance, p: vector of random permanent environmental effects with $p \sim N(0, I\sigma_p^2)$, where σ_p^2 is the permanent environmental variance, e: vector of residual effects with $e \sim N(0, I\sigma_e^2)$, where σ_e^2 is the residual variance and I is the identity matrix. H: the relationship matrix incorporating both pedigree and genomic information, and its inverse is [28]:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

Where A^{-1} is the inverse of the numerator relationship matrix for all animals, G^{-1} is the inverse of the genomic relationship matrix and A_{22}^{-1} is the inverse of the pedigree-based relationship matrix for genotyped animals. The genomic relationship matrix G was constructed using the method of VanRaden [29]: $G = ZDZ'/2\sum_{i=1}^{M} P_i(1-P_i)$, where Z is a matrix relating to the genotypes of each locus, M is the number of SNPs, P_i is the frequency of the second allele at the i_{th} locus. D is a diagonal matrix of SNPs weights (initially $D = I$). The procedure for estimating SNPs weights and effects was carried out as follows [30]:

1. First iteration (t=1), D=I; $G_t = ZD_tZ'/2\sum_{i=1}^{M} P_i(1-P_i)$;
2. GEBVs were estimated using ssGBLUP for the entire dataset;

3. SNP effects were calculated using postGSf90 software with the formula: $\hat{u}_t = \lambda D_t Z G_t^{-1} \sigma_g^2$, where $\hat{u}_t$ is the vector of SNP effects, σ_g^2 is GEBVs for genotyped animals;
4. SNP weights were calculated for each SNP for the next iteration as follows: $D_{i(t+1)} = \hat{u}_{i(t)}^2 2P_i(1-P_i)$, where i is the i-th SNP;
5. Normalizing SNP weights to keep the total genetic variance constant: $D_{(t+1)} = \frac{tr(D_{(1)})}{tr(D_{(t+1)})} D_{(t+1)}$;
6. Recalculating G matrix for the next iteration: $G_{(t+1)} = ZD_{(t+1)}Z'/2\sum_{i=1}^{M} P_i(1-P_i)$
7. $t = t+1$ and repeat from step 2 or 3.

In our study, the procedure was run for two iterations. The results of the analysis were reported as the proportion of genetic variance explained by sliding windows of 20 adjacent SNPs, calculated using the following formula [14,30]:

$$\frac{var(a_i)}{\sigma_a^2} \times 100\% = \frac{var(\sum_{j=1}^{n} z_j \hat{u}_j)}{\sigma_a^2} \times 100\%$$

where a_i is the genetic value of the i-th 20-SNP sliding window; σ_a^2 a is the total additive genetic variance; z_j is the vector of the genotypes of the j-th SNP for all individuals; and $\hat{u}_j$ is the genetic effect of the j-th SNP within the i-th 20-SNP window.

Windows Analysis and Candidate Genes Identification

The candidate QTL regions were evaluated based on the proportion of genetic variance explained by sliding windows across the chromosomes. Windows that explained 0,5% or more of the total genetic variances were considered candidate QTL regions. Within these regions, candidate genes were identified using the NCBI database: http://www.ncbi.nlm.nih.gov.

Results

The results of the study are illustrated in Figures 1 to 3. Each point in the Manhattan plots represents the proportion of genetic variance attributed to a window of 20 adjacent SNPs. The analysis enabled us to identify 12 genomic windows located on chromosomes 7, 8 and 12, each explaining more than 0.5% of the total genetic variance for MY. Collectively, these windows explained 6.33% of the total genetic variance for MY.

For PP, only one window was identified that exceeded the 0.5% threshold. This region located on chromosome 4, explained 0.51% of the total genetic variance for PP.

Regarding FP, six non-overlapping windows were detected. Five were located on chromosome 10, and one on chromosome 14. Together, these windows explained 6.22% of the total genetic variance for FP.

Further investigation of these windows using the NCBI database identified three genes – *XRCC4, VCAN* and *HAPLN1* – on chromosome 7, and four genes – *ABHD13, LIG4, LOC107131142* and *TNFSF13B* – on chromosome 12, all associated with MY.

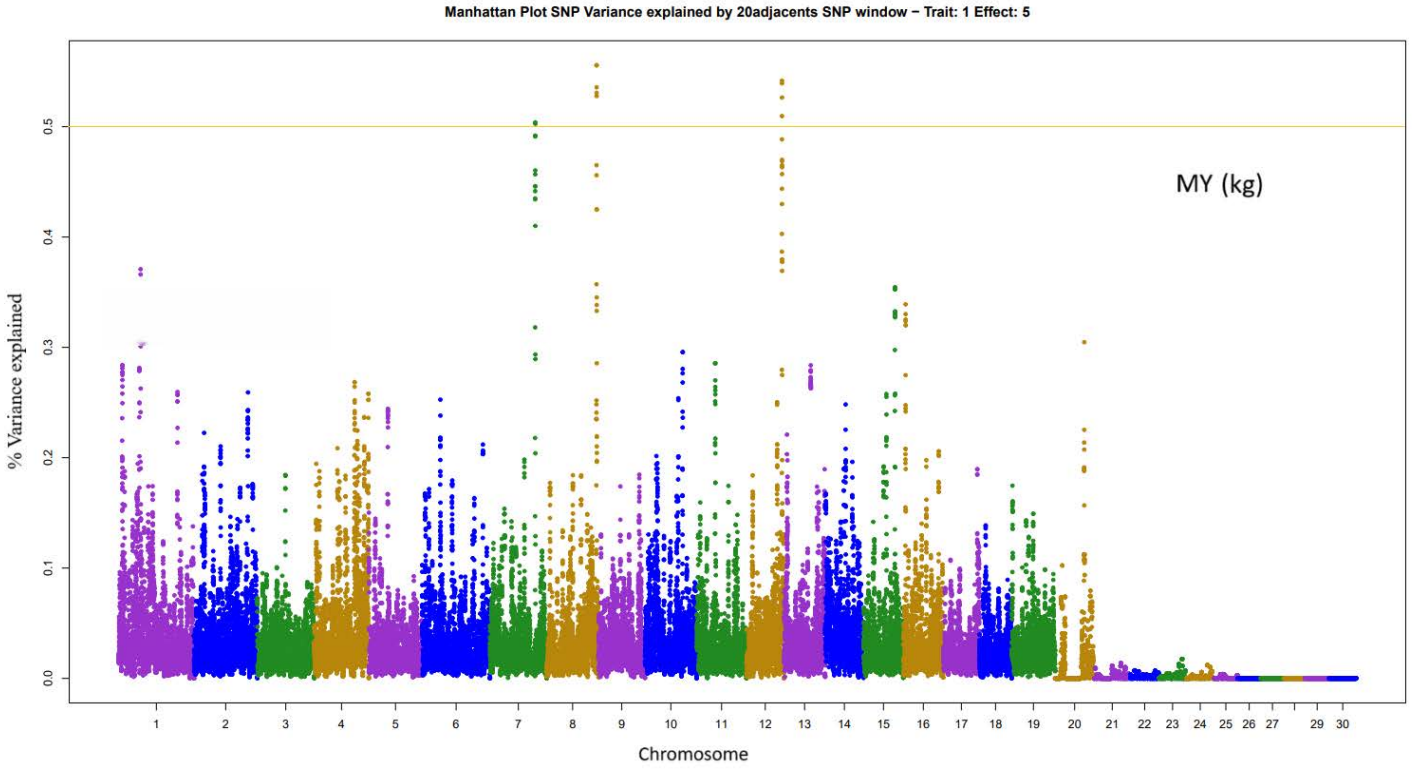

Figure 1: Manhattan plot showing the proportion of genetic variance for daily milk yield (MY) explained by sliding windows of 20 adjacent SNPs.

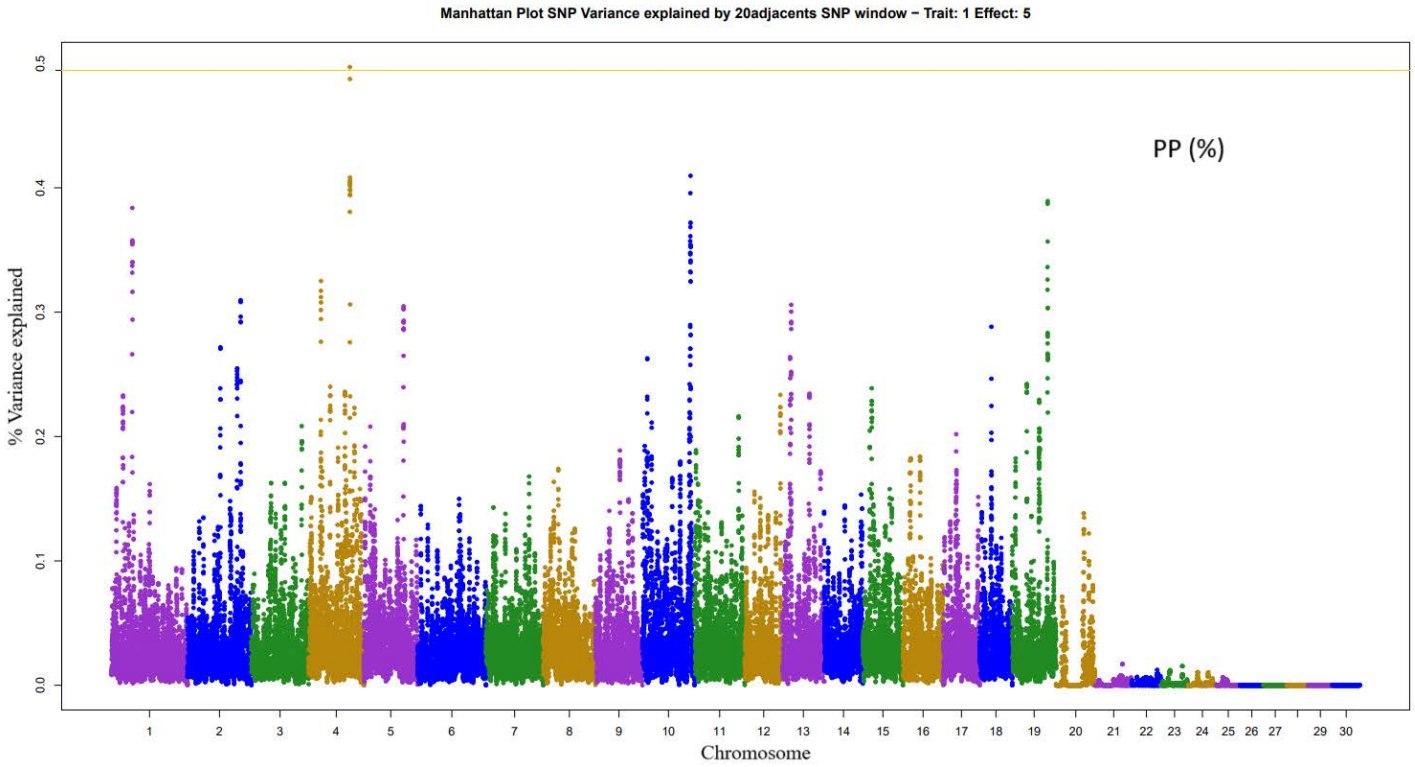

Figure 2: Manhattan plot showing the proportion of genetic variance for milk protein percentage (PP) explained by sliding windows of 20 adjacent SNPs.

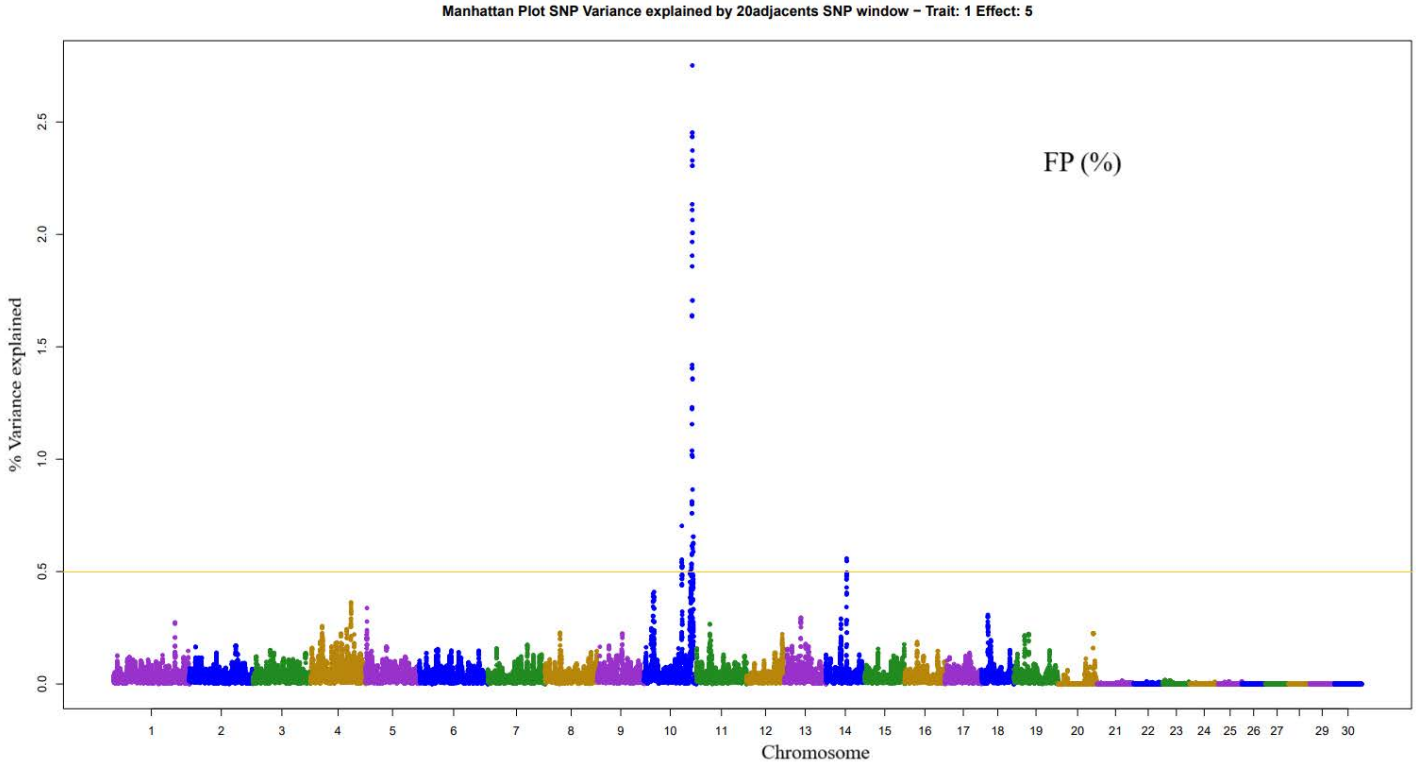

Figure 3: Manhattan plot showing the proportion of genetic variance for milk fat percentage (FP) explained by sliding windows of 20 adjacent SNPs.

Table 1: Top genomic windows explaining the highest proportion of genetic variance, with their chromosomal locations and corresponding identified genes.

Trait	Location[1]	nSNPs	gVar(%)	Genes	Location[2]
MY (kg)	Chr7:85277486-85930083	20	0.50	*XRCC4*	chr7:85267807-85556934
	Chr7:85298955-85957647	20	0.50	*VCAN*	chr7:85666058-85782896
	Chr7:85337935-85971257	20	0.50	*HAPLN1*	chr7:85845448-85924686
	Chr8:109008801-109737722	20	0.56	-	-
	Chr8:109059409-109770679	20	0.56	-	-
	Chr8:109112032-109815182	20	0.54	-	-
	Chr8:109159075-109836761	20	0.53	-	-
	Chr8:109218761-109923217	20	0.53	-	-
	Chr12:87554080-88319342	20	0.54	*ABHD13*	chr12:87630477-87631579
	Chr12:87468430-88278660	20	0.54	*LIG4*	chr12:87606118-87613800
	Chr12:87672871-88394447	20	0.53	*LOC107131142*	chr12:87654333-87671972
	Chr12:87579030-88347846	20	0.51	*TNFSF13B*	chr12:87654130-87671972
PP (%)	Chr4:92588608-93248011	20	0.51	*GCC1*	chr4:92617109-92620507
				ZNF800	chr4:92423262-92444888
				ARF5	chr4:92624397-92627378
				SND1	chr4:92679169-93108508
				LRRC4	chr4:92953572-93043887
				MIR129-1	chr4:93220889-93220960
				LEP	chr4:93249874-93266621
FP (%)*	Chr10:95629821-96695555	20	2,75	-	-
	Chr10:94786066-95524114	20	1,02	-	-
	Chr10:74479448-75685400	20	0.71	*MGC148318*	chr10:74517288-74532706
				KCNH5	chr10:75235434-75637242
	Chr10:98230479-98840028	20	0.66	-	-
	Chr10:93958415-94515340	20	0.54	*SEL1L*	chr10:93900621-93965136
	Chr14:41696729-42550365	20	0.59	*ZFHX4*	chr14:41988047-42189777
				PEX2	chr14:42315032-42330374

Location[1]: position of windows (chromosome: position (bp)), nSNPs: the number of SNPs in each window, gVar (%): the percentage of genomic variance, Location[2]: position of genes (chromosome: position (bp)), MY: daily milk yield (kg), PP: milk protein percentage, FP: milk fat percentage, *: only non-overlapping windows are shown for FP.

For pp, seven genes – 4*GCC1, ZNF800, ARF5, SND1, LRRC4, MIR129-1,* and *LEP* – were identified on chromosome 4. Additionally, three genes – *MGC148318, KCNH5* and *SEL1L* – on chromosome 10, and two genes – *ZFHX4* and *PEX2* – on chromosome 14, were found to be associated with FP trait.

The top genomic windows, which account for the highest proportion of the genetic variance, along with their chromosomal locations and the corresponding identified genes, are presented in Table 1.

Discussion

WssGWAS has demonstrated its efficacy as a valuable and promising technique for identifying QTL [31-33]. This methodology has been employed in numerous studies concerning various economically significant traits in livestock, including milk yield [34], fat [35], protein content [17], somatic cell score [36] and fertility traits [37-39]. This study employed WssGWAS to identify genomic regions and genes linked to three milk production traits: daily milk yield (MY, kg), milk protein percentage (PP, %), and milk fat percentage (FP, %) – in Russian Black-and-White dairy cattle. A threshold of 0.5% of the total genetic variance was employed to identify genomic windows that accounted for the greatest proportion of variance.

For MY, 12 windows were identified on chromosomes 7, 8, and 12. On chromosome 7, three genes – *XRCC4, VCAN,* and *HAPLN1* – were identified. *VCAN* helps build and maintain the matrix that surrounds cells, making it important for cell adhesion, growth, movement, and the formation of new blood vessels. These actions are important for creating tissues and keeping the matrix proteins stable [40,41]. *HAPLN1* encodes hyaluronan/proteoglycan-binding protein 1, which binds to hyaluronic acid – a substance well-known for holding lots of water, and it is integral to the composition of the extracellular matrix in various tissues [42]. The essential function of the third gene – *XRCC4* – is DNA double-strand break repair, especially in the non-homologous end-joining (NHEJ) pathway, however, it is not directly associated with metabolic or lactation processes [43,44].

On chromosome 12, four additional genes associated with MY were identified: *ABHD13*, *LIG4*, *LOC107131142*, and *TNFSF13B.* In a previous study, *ABHD13* was identified within 1 Mb distance from the marker OARCP38. This marker was found to be associated

with milk yield in Lesvos sheep [45]. However, there is currently no direct evidence linking the other three genes to milk production traits, but some research indicates that *TNFSF13B* influences lipid metabolism in adipocytes [46], and *LIG4* is indirectly involved in metabolism through cellular homeostasis, stress responses, and mitochondrial integrity, although its primary function is DNA repairing and maintaining genomic stability [47,48].

One genomic window on chromosome 4 (Chr4:92588608-93248011) was identified as having the greatest association with genomic variance for PP. This window included seven genes: *GCC1*, *ZNF800*, *ARF5, SND1*, *LRRC4*, *MIR129-1*, and *LEP*. Prior research has associated *SND1*, *LEP*, and *LRRC4* with diverse milk production characteristics in cattle. Mammary epithelial cells' nuclei express a highly conserved nuclear transcription factor that is encoded by *SND1*. Prior studies have demonstrated that during lactation, *SND1* expression rises in mammary tissue. Additionally, it is a crucial transcription factor in milk synthesis and has been shown to increase the expression of the milk protein gene β-casein [49].

The hormone leptin, which is encoded by *LEP*, is crucial for controlling energy balance, fertility, and milk production in cattle [50,51]. According to research by J. Citek [52], the W allele of *LEP* is linked to a higher content of milk protein, whereas the MM allele is linked to a lower content ($p < 0.05$).

LRRC4 (leucine-rich repeat-containing protein 4) participates in the negative modulation of the Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway. This gene has been previously linked to 305-day milk yield in Brazilian Girolando cattle [53]. Our study identified it as a candidate gene affecting milk protein content.

Five additional genes were found to be associated with FP – three on chromosome 10 (*MGC148318*, *KCNH5,* and *SEL1L)* and two on chromosome 14 (*ZFHX4* and *PEX2)*. Prior research has shown that the *ZFHX4* gene affects milk fat content in Simmental cattle [54]. A study by Z. Cai et al. demonstrated a correlation between the *SEL1L* gene and milk protein content in dairy cattle [54]. However, to our knowledge, the other three genes – *GC148318*, *KCNH5* and *PEX2* – have not previously been reported to be associated with milk production traits.

Conclusions

In this study, we conducted a WssGWAS to identify genomic regions and candidate genes associated with 3 milk production traits – daily milk yield (MY, kg), milk protein percentage (PP, %), and milk fat percentage (FP, %) – in Russian Black-and-White dairy cattle. We identified 18 genomic regions associated with the largest proportion of genetic variance for these traits. Within these regions, we identified seven candidate genes on chromosomes 7 and 12 related to MY, seven candidate genes on chromosome 4 associated with PP, and five candidate genes on chromosomes 10 and 14 associated with FP. These results may offer a contribution in exploring the genetic structure of complex traits in dairy cattle and aid in the enhancement of genomic progress in dairy cattle breeding programs.

Acknowledgments

The authors would like to thank the staff members of I Gen and Genoanalytica in Russia. The study was partially supported by the Russian Foundation for Basic Research under the Agreement No. 14.579.21.0147, Unique Project Identifier: RFMEFI57917X0147 and grant No. 17-29-08033. Primary results of the study were presented on 26 February 2022 at 6th International Conference on Bioscience and Biotechnology, Jour of Micr & Biote Rep, Volume: 05.

Funding

The study was partially supported by the Russian Foundation for Basic Research within the framework of the project under Agreement No. 14.579.21.0147, with the unique project identifier: RFMEFI57917X0147 and grant No. 17-29-08033.

References

1. Iung LHS, Petrini J, Ramirez-Diaz J, et al. Genome-wide association study for milk production traits in a Brazilian Holstein population. J Dairy Sci. 2019; 102: 5305-5314.
2. Liu JJ, Liang AX, Campanile G, et al. Genome-wide association studies to identify quantitative trait loci affecting milk production traits in water buffalo. J Dairy Sci. 2018; 101: 433-444.
3. Weigel KA, VanRaden PM, Norman HD, et al. A 100-Year Review: Methods and impact of genetic selection in dairy cattle-From daughter-dam comparisons to deep learning algorithms. J Dairy Sci. 2017; 100: 10234-10250.
4. Kwon DH, Gim GM, Yum SY, et al. Current status and future of gene engineering in livestock. BMB Rep. 2024; 57: 50-59.
5. Kushanov FN, Turaev OS, Ernazarova DK, et al. Genetic Diversity, QTL Mapping, and Marker-Assisted Selection Technology in Cotton (Gossypium spp.). Front Plant Sci. 2021; 12: 779386.
6. Zhou J, Liu L, Chen CJ, et al. Genome-wide association study of milk and reproductive traits in dual-purpose Xinjiang Brown cattle. BMC Genomics. 2019; 20: 827.
7. Ibtisham F, Zhang L, Xiao M, et al. Genomic selection and its application in animal breeding. Thai J Vet Med. 2017; 47: 301-310.
8. Hayes B. QTL mapping, MAS, and genomic selection. A short-course Animal Breeding & Genetics Department of Animal Science Iowa State University. 2007; 1: 3-4.
9. McClure MC, Morsci NS, Schnabel RD, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim Genet. 2010; 41: 597-607.
10. Oliveira HR, Cant JP, Brito LF, et al. Genome-wide association for milk production traits and somatic cell score in different lactation stages of Ayrshire, Holstein, and Jersey dairy cattle. J Dairy Sci. 2019; 102: 8159-8174.

11. Susmitha P, Kumar P, Yadav P, et al. Genome-wide association study as a powerful tool for dissecting competitive traits in legumes. Front Plant Sci. 2023; 14: 1123631.
12. Zhu R, Zhang Y, Jiang Y, et al. Weighted single-step genome-wide association study identified genomic regions and candidate genes for growth and reproductive traits in Wenchang chicken. Poult Sci. 2025; 104: 104733.
13. Botelho ME, Lopes MS, Mathur PK, et al. Weighted genome-wide association study reveals new candidate genes related to boar taint compounds 1. Livestock Science. 2022; 257: 104845.
14. Marques DBD, Bastiaansen JWM, Broekhuijse M, et al. Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. Genet Sel Evol. 2018; 50: 40.
15. Araujo AC, Carneiro PLS, Alvarenga AB, et al. Haplotype-Based Single-Step GWAS for Yearling Temperament in American Angus Cattle. Genes (Basel). 2021; 13: 17.
16. Park J, Na CS. Weighted single-step genome-wide association study to reveal new candidate genes for productive traits of Landrace pig in Korea. J Anim Sci Technol. 2024; 66: 702-716.
17. Raschia MA, Nani JP, Carignano HA, et al. Weighted single-step genome-wide association analyses for milk traits in Holstein and Holstein x Jersey crossbred dairy cattle. Livestock Science. 2020; 242: 104294.
18. Peripolli E, Temp LB, Fukumasu H, et al. Identification of candidate genes and genomic prediction for early heifer pregnancy in Nelore beef cattle. Livestock Science. 2024; 289: 105582.
19. Khazaei-Koohpar H, Gholizadeh M, Hafezian SH, et al. Weighted single-step genome-wide association study for direct and maternal genetic effects associated with birth and weaning weights in sheep. Sci Rep. 2024; 14: 13120.
20. Pan R, Qi L, Xu Z, et al. Weighted single-step GWAS identified candidate genes associated with carcass traits in a Chinese yellow-feathered chicken population. Poult Sci. 2024; 103: 103341.
21. Vahedi SM, Salek Ardestani S, Karimi K, et al. Weighted Single-Step GWAS for Body Mass Index and Scans for Recent Signatures of Selection in Yorkshire Pigs. J Hered. 2022; 113: 325-335.
22. Sermyagin A, Gladyr E, Plemyashov K, et al. Genome-wide association studies for milk production traits in Russian population of Holstein and black-and-white cattle. Proceedings of the Scientific-Practical Conference «Research and Development 2016». Springer, Cham. 2017; 591-599.
23. Plemyashov K, Krutikova A, Belikova A, et al. Genome-Wide Association Study (GWAS) for Left Displaced Abomasum in Highly Productive Russian Holstein Cattle. Animals (Basel). 2024; 14: 2795.
24. Igoshin AV, Yurchenko AA, Belonogova NM, et al. Genome-wide association study and scan for signatures of selection point to candidate genes for body temperature maintenance under the cold stress in Siberian cattle populations. BMC Genet. 2019; 20: 26.
25. Abdulrahman K. Genomic estimated breeding value of milk performance and fertility traits in the Russian Black-and-White cattle population. Acta Naturae. 2022; 14: 109-122.
26. Forni S, Aguilar I, Misztal I. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet Sel Evol. 2011; 43: 1.
27. Misztal I, Tsuruta S, Strabel T, et al. Proceedings from the 7th World Congress on Genetics Applied to Livestock Production. 2002.
28. Aguilar I, Misztal I, Legarra A, et al. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J Anim Breed Genet. 2011; 128: 422-428.
29. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008; 91: 4414-4423.
30. Wen H, Luo H, Yang M, et al. Genetic parameters and weighted single-step genome-wide association study for supernumerary teats in Holstein cattle. J Dairy Sci. 2021; 104: 11867-11877.
31. Zhao B, Luo H, Huang X, et al. Integration of a single-step genome-wide association study with a multi-tissue transcriptome analysis provides novel insights into the genetic basis of wool and weight traits in sheep. Genet Sel Evol. 2021; 53: 56.
32. Li S, Chang Y, Zhang H, et al. Estimation of Genetic Parameters and Weighted Single-Step Genome-Wide Association Studies for Fertility Traits in Holstein Cattle. Agriculture. 2025; 15: 1763.
33. Vahedi SM, Ardestani SS, Do DN, et al. 149 Integration of Selection Signatures Analyses and Weighted Single-step GWAS to Prioritize Candidate Genes for Body Conformation Traits in Pigs. J Anim Sci. 2021; 99: 76.
34. George L, Alex R, Gowane G, et al. Weighted single step GWAS reveals genomic regions associated with economic traits in Murrah buffaloes. Anim Biotechnol. 2024; 35: 2319622.
35. Atashi H, Chen Y, Wilmot H, et al. Single-step genome-wide association for selected milk fatty acids in Dual-Purpose Belgian Blue cows. J Dairy Sci. 2023; 106: 6299-6315.
36. Mohammadi H, Farahani AHK, Moradi MH, et al. Weighted Single-Step Genome-Wide Association Study Uncovers Known and Novel Candidate Genomic Regions for Milk Production Traits and Somatic Cell Score in Valle del Belice Dairy Sheep. Animals (Basel). 2022; 12: 1155.
37. Stafuzza NB, Costa e Silva EVd, Silva RMdO, et al. Genome-wide association study for age at puberty in young Nelore bulls. J Anim Breed Genet. 2020; 137: 234-244.

38. Lett BM, Kirkpatrick BW. Identifying genetic variants and pathways influencing daughter averages for twinning in North American Holstein cattle and evaluating the potential for genomic selection. J Dairy Sci. 2022; 105: 5972-5984.

39. Stafuzza NB, Silva RMO, Fragomeni BO, et al. A genome-wide single nucleotide polymorphism and copy number variation analysis for number of piglets born alive. BMC Genomics. 2019; 20: 321.

40. Harten IA, Kaber G, Agarwal KJ, et al. The synthesis and secretion of versican isoform V3 by mammalian cells: A role for N-linked glycosylation. Matrix Bio. 2020; 89: 27-42.

41. Islam S, Watanabe H. Versican: A Dynamic Regulator of the Extracellular Matrix. J Histochem Cytochem. 2020; 68: 763-775.

42. Dovedytis M, Liu ZJ, Bartlett S. Hyaluronic acid and its biomedical applications: A review. Engineered Regeneration. 2020; 1: 102-113.

43. Junop MS, Modesti M, Guarne A, et al. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. EMBO J. 2000; 19: 5962-5970.

44. Hammel M, Rey M, Yu Y, et al. XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair. J Biol Chem. 2011; 286: 32638-32650.

45. Mastranestasis I, Kominakis A, Hager-Theodorides A, et al. Associations between genetic polymorphisms and phenotypic traits in the Lesvos dairy sheep. Small Ruminant Research. 2016; 144: 205-210.

46. Chan CC, Harley ITW, Pfluger PT, et al. A BAFF/APRIL axis regulates obesogenic diet-driven weight gain. Nat Commun. 2021; 12: 2911.

47. Sallmyr A, Bhandari SK, Naila T, et al. Mammalian DNA ligases; roles in maintaining genome integrity. J Mol Biol. 2024; 436: 168276.

48. Ando K, Suenaga Y, Kamijo T. DNA Ligase 4 Contributes to Cell Proliferation against DNA-PK Inhibition in MYCN-Amplified Neuroblastoma IMR32 Cells. Int J Mol Sci. 2023; 24: 9012.

49. Lu L, Gao X, Li Q, et al. Comparative phosphoproteomics analysis of the effects of L-methionine on dairy cow mammary epithelial cells. Canadian Journal of Animal Science. 2012; 92: 433-442.

50. Giblin L, Butler ST, Kearney BM, et al. Association of bovine leptin polymorphisms with energy output and energy storage traits in progeny tested Holstein-Friesian dairy cattle sires. BMC Genet. 2010; 11: 73.

51. Safina N, Sharafutdinov G, Akhmetov T, et al. Association of LEP gene polymorphism with biochemical parameters of lipid metabolism and milk productivity of Holstein cattle. E3S Web Conf. 2021; 254.

52. Čítek J, Brzáková M, Hanusová L, et al. Gene polymorphisms influencing yield, composition and technological properties of milk from Czech Simmental and Holstein cows. Anim Biosci. 2021; 34: 2-11.

53. Otto PI, Guimaraes SEF, Calus MPL, et al. Single-step genome-wide association studies (GWAS) and post-GWAS analyses to identify genomic regions and candidate genes for milk yield in Brazilian Girolando cattle. J Dairy Sci. 2020; 103: 10347-10360.

54. Ignatieva LP, Sermyagin AA, Nikitin S, et al. PSXI-2 Genome-Wide association study and possibilities for genomic selection of Simmental cattle breed in Russia. Journal of Animal Science. 2021; 99: 246-247.

55. Cai Z, Dusza M, Guldbrandtsen B, et al. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. Genet Sel Evol. 2020; 52: 19.

© 2025 Abdulrahman K, et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License