Ann. Anim. Sci., Vol. 23, No. 2 (2023) 419–430 *DOI: 10.2478/aoas-2022-0091*

THE ASSOCIATION OF GENE POLYMORPHISMS WITH MILK PRODUCTION AND MASTITIS RESISTANCE PHENOTYPIC TRAITS IN DAIRY CATTLE*

Emilia Bagnicka[1♦], Paulina Brzozowska[1,2], Kacper Żukowski[3], Renata Grochowska[4]

[1]Institute of Genetics and Animal Biotechnology Polish Academy of Sciences, Postępu 36A, 05-552 Jastrzębiec, Poland
[2]Faculty of Dietetics, Warsaw College of Engineering and Health, Bitwy Warszawskiej 1920 r. 18, 02-366 Warszawa, Poland
[3]National Research Institute of Animal Production, 32-083 Balice n. Kraków, Poland
[4]Institute of Agricultural and Food Economics – National Research Institute, Świętokrzyska 20, 00-002 Warszawa, Poland
♦Corresponding author: e.bagnicka@igbzpan.pl

Abstract

The aim of this study was to evaluate the association between gene polymorphisms (SNPs) and mastitis indicators and their relationship with milk production profitability in dairy herd. A functional analysis was also performed of five genes containing the studied SNPs and those located close by. DNA was isolated from the hair bulb of 320 dairy cows kept in three herds and SNP-microarray analysis was performed. The data on 299 cows was subjected to final statistical analysis using AI-REML method with one-trait repeatability test-day animal model and pedigree information using the DMU4 package. Five from 35 SNPs significantly associated with mastitis indicators or production traits and located within a gene or no more than 500,000 nucleotides from the gene were selected for the functional and economic analysis. A questionnaire was also developed to collect associated economic data of 219 cows from three herds, such as the value of milk production and direct costs incurred over three years; this allowed the gross margin, direct profitability index and direct costs incurred to produce one liter of milk to be determined, among others. None of the five studied SNPs were related to protein content. The rs110785912(T/A), found near *CXCR4*, and rs136813430(T/C), located in the *TLR4* gene exon, were associated with lnSCC, while rs110455063(C/G), located near *IGF-I*, was associated with milk yield, fat and total solid contents. rs109421300(T/C), associated with fat/protein content ratio, as well as fat and total solid content, is located in the *DGAT1* gene intron. rs41587003(A/C), located in the *DLG2* gene intron, was associated with lactose content. The economic analysis revealed differences between the variants of the three tested SNPs. The T/C variant of the rs136813430(T/C) SNP was characterized by the highest gross margin, the highest direct profitability index and the lowest costs incurred to produce 1 liter of milk. The T/A variant of rs110785912(T/A) was related to low lnSCC and was characterized by the highest direct profitability index. In turn, the C/C variant of the rs41587003(A/C) was related to the lowest level of lactose and the highest costs of milk production. It appears that rs136813430(T/C) may be the most promising of the tested SNPs for increasing the profitability of milk production. To our knowledge, it is the first effort to assess directly a correlation between the DNA polymorphism and economic output of a dairy enterprise.

Key words: SNP marker, intron, exon, association, mastitis indicator, dairy economics, milk production profitability

Mastitis is still one of the major diseases affecting dairy cattle, with negative effects on milk production (Davies et al., 2009). Despite the wealth of research performed on the physiological and cellular processes taking place in the mammary gland in response to infection, knowledge of many of the defense mechanisms remains incomplete (Sordillo, 2005). Mastitis is negatively correlated with milk yield, with high-yielding cows being more susceptible (Carlen et al., 2005).

Most genetic studies of mastitis indicators to date have focused on the number of somatic cells (SCC) in milk. The heritability coefficient (h^2) for SCC was estimated from 0.05 to 0.20 (Ptak et al., 2007; Sender et al., 2013). With such a low to moderate degree of heritability, breeding progress is rather slow, as genetic variability directly related to resistance accounts for only a small part of the total variance; however, this value is significant. The other trait considered as additive mastitis indicator is lactose content (Bagnicka et al., 2016; Andrei et al., 2009).

Inflammation is a complex trait controlled by many genes. As such, it is difficult to develop an appropriate selection strategy to increase immunity. Such programs are complicated by the lack of information on the genetic basis of animal immunity, the complexity of the functional interaction between host and pathogen, which may trigger a range of different immune responses, and the influence of the environment (Marogna et al., 2010).

*This study was supported by the Leading National Research Centre Scientific Consortium's "Healthy Animal – Safe Food" initiative under the Ministry of Science and Higher Education decision No. 05-1/KNOW2/2015, grant KNOW2015/CB/PRO1/52, to Emilia Bagnicka. The publication costs were co-financed by the subsidy of the Ministry of Agriculture and Rural Development (DŻW.pp.862.1.2023).

The search for genes related to the susceptibility/resistance of dairy cows to udder inflammation and production traits has recently been supported by the development of genome-wide association studies (GWAS) using phenotypic data and data on the genotypes of the studied individuals. One of the most popular methods of genotyping concerns the use of dense sets of SNP (Single Nucleotide Polymorphism) microarrays.

Taking advantage of genomics for breeding purposes raises the question of whether the use of selection methods based on SNPs influencing the immunity of the cow mammary gland is economically justified, i.e., whether this approach will reduce losses related to udder diseases and thus increase the efficiency of dairy cows, and, consequently, reduce milk production costs in a herd. The future of livestock production is not about production intensification at all costs, which requires increasing expenditure to account for more limited resources and environmental pollution. A more desirable goal is to minimize production costs thanks to the use of *inter alia* new genomic technologies, which enable the selection of individuals with specific genetic potential in terms of production and health.

Most of the research on udder health to date has focused on estimating the costs of preventing inflammation, losses due to mastitis, and effectively managing a herd to reduce such losses (Hagnestam-Nielsen and Østergaard, 2009). The use of new technologies based on SNPs has increased the efficiency of selecting economically significant traits. Hence, a number of studies have been carried out in cattle to identify relationships between gene polymorphisms and various production and functional traits (Lü et al., 2011; Yuan et al., 2013). Previous studies have only indicated the potential of economic benefits related to higher meat and milk production or lower SCC in milk; however, these traits do not guarantee greater milk production profitability, which is determined by the production volume and the costs incurred. None of these studies have examined the profitability of using individual SNPs in cattle selection based on real economic data like gross margin, direct profitability index and costs incurred to produce 1 liter of milk. Therefore, this paper attempts to fill this gap by analyzing the economic effects of SNPs in genes related to milk production and health based on information regarding the production value and direct costs obtained from farms.

It was hypothesized that the use of gene polymorphisms related to udder inflammation in dairy cows would increase milk production profitability; therefore, the aim of this study was to evaluate the association between impact of five selected polymorphisms (SNPs) and mastitis indicators or production traits as well as to investigate the impact of selected identified SNPs on the profitability of milk production in a dairy herd. A functional analysis was also made of the genes that include the studied SNPs and genes located nearby.

Material and methods

Animals and farms

The first step of the research was carried out on 320 dairy cows from three herds. All were kept in free stall barns with constant access to enclosures. The farms differed in terms of the number of cows and feeding system, resulting from diversified area of arable land, including permanent grassland. The mean herd yield was 9,393 liters of milk per lactation, with mean fat and protein contents of 4.18% and 3.49%, respectively. Cows were fed in the total mixed ration (TMR) system, balanced in accordance with the Institut national de la recherche agronomique (INRA) standards and adapted to Polish conditions (Brzóska et al., 2014), with constant access to water. The cows were milked twice a day with a mechanical milking system (DeLaval, Tumba, Sweden). All herds were assessed for dairy and breeding performance under routine performance control. Phenotypic data on the daily milk yield and the component contents was obtained from the Polish Federation of Cattle Breeders and Dairy Farmers, based on which a database with 8090 observations was prepared. Information on daily milk yield, protein, fat, dry matter, lactose contents and SCC were gathered. The yield of milk components and fat/protein ratio were also calculated. The descriptive statistic of all studied traits is presented in Table S1. The pedigree database contained information on 7,035 animals, along with their birth year, and 2,057 sires and 4,634 dams with 64 sires having more than 10 daughters (range between 10 and 61). However, the genotyped cows were descendants of 153 sires (54, 40 and 59 sires in each herd respectively) and 263 dams. Ten sires of genotyped cows had daughters in three or two herds.

Sampling and gene polymorphism identification

DNA was isolated from approximately 100 hair follicles from each cow using a kit for isolating DNA from blood and tissues (Macherey-Nagel, Düren, Germany). The procedure was performed according to the manufacturer's protocol, with the following modification: the hair root cells were disintegrated by dipping four times in liquid nitrogen, and then thawing in a water bath at 37°C and increasing the amount of added proteinase from 25 µl to 50 µl. Just before performing the microarray analysis, quantitative DNA analysis was performed using intercalating dyes capable of specifically binding to double-stranded DNA (Quant-iT PicoGreen dsDNA kit, Thermo Fisher Scientific). Fluorescence measurement was performed using a Fluoroskan Ascent FL apparatus (Thermo Fisher Scientific, Waltham, Massachusetts, USA) according to the manufacturer's protocol. Eight samples with concentrations below 50 ng DNA/µl were excluded from the further analysis.

Microarray analysis was performed on 120 ng of DNA of 312 cows from each run using the BeadChip Bovine 50k v3 kit. The results were interpreted using the GenomeStudio 1.9.4 software (Illumina, San Diego, CA, USA).

As the DNA was isolated from hair follicles, no permission was needed from the Local Ethics Committee for biological material collection.

Association study

The next step in establishing the association between individual SNPs and selected mastitis indicators and milk traits was to conduct a GWAS-based basic bioinformatic analysis using the Plink.2 program (Chang et al., 2015) for more than 50K SNPs; a linear regression model was used for this purpose, taking into account the fixed effect of SNP and population structure among individuals as covariate. Genotypes with low frequency and SNPs which were monomorphic in the studied group were excluded from the analysis. As a result, the number of SNPs tested was reduced down to 46K. All phenotypic traits were averaged using an interquartile distribution: IQR(x) = quantile(x. 3/4) - quantile(x. 1/4). In order to check the differences in the environment between the studied herds, the multidimensional scaling analysis (MDA) was performed. Genotyping rating was 0.98. All outliers, i.e. information on 13 cows, were removed from further analysis.

Following the above, a preliminary analysis of the associations between SNPs and mastitis indicators and milk traits was conducted using the PROC MIXED SAS package (SAS/STAT 2002–2012) with a mixed model using information on production of 299 genotyped cows. All SNPs which were found to be associated with udder health in the previous step of the analysis (lnSCC– a natural logarithm from SCC, lactose content) and with production traits (milk yield, protein and fat content or yield) were taken into account. Altogether, 52 SNPs were selected, of which only 35 were significantly associated with the studied traits in this analysis of variance. Finally, these 35 SNPs were analyzed by AI-REML to obtain solutions for fixed effects, using the DMU4 package (Madsen and Jensen, 2013). A one-trait repeatability test-day animal model was applied.

Phenotypic data was distributed among 21 calving years from 1997 to 2017 – each year constituted a separate class and calving season: 1 – Winter (December, January, February), 2 – Spring (March, April, May), 3 – Summer (June, July, August), and 4 – Autumn (September, October, November). Daily milking data were collected across 1996–2017. Finally, 150 classes of interaction were created between herd, calving year and season (HYS) and 537 classes of interaction between herd, year and month of milking (hym). Parity was grouped into four classes, with the fourth containing lactations higher than the third. The final statistical classification fitted HYS, parity, and SNPs along with the Legendre polynomial of order 5 regression on standardized DIM nested within parity as fixed effects and genetic additive, permanent environment, hym, and residual effects as random (File S1). Each analysis included the relationship between one or two SNPs and the tested features at a time due to computational limitation which is the shortage of the study as not all epistatic influences are considered in the analysis to estimate the differences between variants of particular SNPs. *Post hoc* test with the Bonferroni correction was used. Five from 35 SNPs significantly associated with mastitis indicators or then with production traits and located within a gene or no more than 500,000 nucleotides from the gene were selected for the functional and economic analysis.

Before statistical analysis, the total SCC were transformed to natural logarithm values (lnSCC). $\chi 2$ test was applied to determine whether SNP genotypes frequency and Hardy–Weinberg equilibrium held.

Functional analysis of genes with analyzed SNPs

Seven genes associated with studied SNPs were subjected to functional analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Kanehisa at al., 2016) and Gene Ontology Annotation (GO) (Ashburner et al., 2000).

Economic analysis

To assess the value of the SNP polymorphisms for selection purposes, a questionnaire was developed to collect the actual data on milk production of 219 cows and the direct costs incurred over three calendar years, i.e., 2015, 2016 and 2017 in three herds. Altogether, 384 records were analyzed. The analysis was based on calculations in accordance with the AGROKOSZTY methodology (Methodology, 1999; Skarżyńska, 2012).

The research covered revenues, i.e., the value of potential commercial production (assuming that the sales volume is equal to the production volume) and direct costs. The gross margin was adopted as the basic measure of the assessment of the obtained economic effects. The calculations do not take into account support with subsidies under the Common Agricultural Policy. In calculating the gross margin, only the value of the main product, i.e., milk, traded on the market was considered. The annual value of milk production of each genotyped cow was calculated based on the monthly milk yield of the genotyped cows, and the mean monthly selling prices of milk for a given farm.

Due to the inability to record information on the incurred direct costs per genotyped cow, the gross margin was calculated based on the direct costs collected for the entire dairy cow herd, and this value was converted to a single cow in the herd (Skarżyńska, 2012). For this purpose, information was obtained on the costs of feed from outside the farm, of own feed from potential commercial and non-commodity products, and of veterinary measures and services (including semen and insemination) and specialist costs (e.g., hoof correction, preparation of animals for exhibitions).

Thus, milk production profitability for the entire herd of dairy cows in individual farms was determined based on milk production value, calculated as the product of the average milk yield per cow in a year and the mean annual milk price for a given herd, and direct costs, estimated based on direct costs incurred in the entire herd in a year, per cow.

The following economic categories were considered:

gross margin (GM) = PV – DC

where: PV – production value (PLN); DC – direct costs (PLN)

direct profitability index (DP) = (PV / DC) × 100

where: PV – production value (PLN), DC – direct costs (PLN)

direct costs incurred to produce 1 liter of milk (MDC) = DC / MP

where: DC – direct costs (PLN), MP – milk production (liters)

To determine the significance of differences between the tested SNP variants, an analysis of variance with a single-factor model was performed for each economic category; in each calculation, a single SNP and the parity over the three studied years were used as fixed effects. The calculations were performed using the PROC MIXED SAS package (SAS/STAT 2002–2012).

Results

Gene polymorphism analysis

The multidimensional scaling analysis (MDA), which presents the environmental differences between the studied herds and Manhattan plots as the result of GWAS analysis for some studied traits, is presented in supplementary file (Figure S1 a and S1 b, A–E). After three-step selection, based on the present findings SNPs: rs110785912(T/A), rs110455063(C/G), rs109421300(T/C), rs41587003(A/C), and rs136813430 (T/C) were selected for the functional and economic analysis.

The outcomes of the association analysis of selected SNPs with functional and production traits of dairy cows are shown in Table 1, while the distribution of genes and alleles is presented in Table 2. Studied SNPs were associated with traits at least at P<0.05, except rs110785912 which was associated with lnSCC at P = 0.0763. For rs110785912(A/T) and rs110455063(C/G), a very high frequency of one allele was reported. The χ^2 test results indicate that the observed frequency of genotypes does not differ from that expected one (Table 2).

The first studied SNP, rs110785912(T/A), also known as ARS-BFGL-NGS-50482, associated with lnSCC in our study, is located in the intron of a gene marked as *LOC1049123* (NCBI-1). The function of the gene is unknown. The SNP identified 447,880 nucleotides from the C-X-C chemokine receptor type 4 (*CXCR4*) gene (Table 3). A/A and A/T genotypes were associated with lower lnSCC than T/T (Table S2). The lnSCC heritability coefficient estimated in the tested population was 0.13, with the investigated trait demonstrating high coefficient of variance (CV=34.5%) which reflects a total variability.

SNP rs110455063(C/G) known also as ARS-BFGL-NGS-15787, was located in the intron of gene with unknown function and marked as *LOC104972477* (NCBI-2) and, simultaneously, was located 34,194>(5') nucleotides from the insulin-like growth factor I (*IGF-I*) gene. Its frequency was 0.228 for G/G, 0.759 for C/G, and 0.013 C/C (Table 2). It was found to be associated with milk yield, fat and total solid content. G/G variant carriers had the highest milk production while the highest fat content was found in the milk of cows with the C/C genotype, and the lowest in heterozygotes. The highest dry matter content was observed in the milk of heterozygous cows and the lowest content in cows with C/C. The total variability of the milk yield expressed as CV was high as 60%, while fat was 22.5% and dry matter content only 8.37%. The heritability coefficients were 0.38, 0.17 and 0.20, respectively (Table S2).

Table 1. Relationship of the studied SNPs with the functional and production characteristics of dairy cows

Trait	Chromosome	rs number	P-value**
Production traits			
Milk yield	5	rs110455063	0.021
Fat content	5	rs110455063	<0.0001
	14	rs109421300	
Dry matter content	5	rs110455063	0.0004
	14	rs109421300	<0.0001
Health mammary gland indicators			
lnSCC	2	rs110785912	0.0763
	8	rs136813430	<0.0001
Lactose content	29	rs41587003	0.02701

**P-value after Bonferroni correction.

lnSCC – natural logarithm of the number of somatic cells in milk (SCC).

Table 2. Distribution of genotypes and alleles of analyzed SNPs

SNP	N	Distribution of genotype in the studied population			Alleles' frequency/ Genotypes' frequency	χ^2 test value/ P-value
rs110785912(A/T)	299	(A/A)=243	(AT)=55	(T/T)=1	A=0.905; T=0.095 /0.813|0.184|0.003|	1.33 0.40
rs110455063(C/G)	299	(G/G)=227	(CG)=68	(C/C)= 4	G=0.873; C=0.127 /0.228|0.759|0.013	0.19 0.83
rs136813430(T/C)	294	(C/C)=112	(T/C)=131	(T/T)=51	C=0.604; T=0.396 /0.380|0.450|0.170	1.39 0.24
rs109421300(T/C)	299	(A/A)= 140	(A/G)=137	5(G/G)=22	A=0.697; G=0.303 /0.468|0.458|0.074	1.2 0.14
rs41587003(A/C)	299	(A/A)=101	(AC)= 141	(C/C)=57	A=0.694; C=0.306 /0.410|0.570|0.020	0.39/ 0.53

N – number of genotyped cows , SNP = single nucleotide polymorphism.

Table 3. Location of the five SNPs related to the analyzed traits in the bovine genome and position in the gene/relative to the nearest gene

Chr.	SNP (nucleotides)	Trait	SNP's position in the chromosome	Symbol of the closest gene	Gene location in the chromosome	SNP location in the gene/in relation to the closest gene
8	rs136813430(T/C)	lnSCC	108829443	*TLR4*	108828899–108839913	Exon
29	rs41587003(A/C)	lactose content	10172197	*DLG2*	9997510–12157167	Intron
2	rs110785912(A/T)	lnSCC	61134245	*LOC1049123*	61096534–61148308	Intron
				CXCR4	61250082–61253877	115837>(5')
5	rs110455063(C/G)	daily milk yield, fat content, dry matter content	61134245	*LOC104972477*	66131228–66164165	Intron
				IGF1	66191602–66264083	34194>(5')
14	rs109421300(T/C)	fat to protein ratio, fat content, dry matter content	1801116	*DGAT1*	1795425–1804838	Intron
				GPIHBP1	1348153–1350911	450205<(3')
				CYP11B1	1550747–1558425	242691<(3')

Chr. = chromosome; allele: A (adenine), C (cytosyne), G (guanine), T (thymine), lnSCC – a natural logarithm of somatic cell count.
TLR4 – Toll-like receptor 4; DLG2 – Disks large homolog 2; *CXCR4* – C-X-C chemokine receptor type 4; *IGF1* – Insulin-like growth factor I; *DGAT1* – Diacylglycerol O-acyltransferase 1; *GPIHBP1* – Glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1; *CYP11B1* – Cytochrome P450 11B1.

The rs109421300(T/C), *alias* ARS-BFGL-NGS-4939, associated with fat/protein content ratio, fat and total solid content and fatty acid profile (Table 1), is located in the intron of the diacylglycerol O-acyltransferase 1 (*DGAT1*) gene (Table 3); it is also sited 450,205<(3') nucleotides from the glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1 (*GPIHBP1*) gene and 242,691>(3') from the cytochrome P450 11B1 (mitochondrial) (*CYP11B1*) gene. The lowest fat content was found in animals with the T/T variant, and the highest with the C/C genotype. The highest dry matter content was found in the milk of cows with the C/C genotype and the lowest in milk of cows with T/T. The total variability of the fat to protein ratio was found to be average (CV=20.8%), and the h^2 of this trait was 0.14 (Table S2).

SNP rs41587003(A/C) (UA-IFASA-7512) was located in the intron of the disks large homolog 2 (*DLG2*) gene (Table 3) and was associated with lactose content: the highest value was associated with the A/A variant and the lowest with C/C. The heritability of lactose content was estimated at 0.16, and the total variability (CV) of this trait was 9.23% (Table S2).

Finally, SNP rs136813430(T/C) (BTA-82770-no-rs), associated with lnSCC (Table S2), was located in the exon of the toll-like receptor 4 (*TLR4*) gene (Table 3). The milk of cows with the C/T genotype contained more somatic cells than the other variants of genotypes.

None of the analyzed SNPs were directly related to protein content or yield. The indirect relationship expressed as fat to protein ratio was mentioned above.

Functional analysis

Seven genes with known functions (*TLR4* – Toll-like receptor 4; *DLG2* – Disks large homolog 2; *CXCR4* – C-X-C chemokine receptor type 4; *IGF1* – Insulin-like growth factor I; *DGAT1* – Diacylglycerol O-acyltrans-

ferase 1; *GPIHBP1* – Glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1; *CYP11B1* – Cytochrome P450 11B1) were associated with studied SNPs and subjected to functional analysis.

In silico GO annotation analysis revealed that the protein products of two from five genes i.e., *TLR4* and *CXCR4*, whose polymorphisms were associated with lnSCC, were involved in several common biological processes connected with immune system such as homeostatic process, immune system process, signal transduction, or response to stress (Figure 1). However, they were also involved in cell differentiation, cell morphogenesis, cell motility, or cell death. Thus, further association studies are needed to check the relationships with other diseases and productivity of the cows. The GO number and definitions were gathered in Table S3 (supplementary files) based on QuickGO database (term/GO:0002376).

The next SNP – rs41587003(A/C), associated with lactose content, is localized in the intron of *DLG2* gene which is required for perception of chronic pain through N-methyl-D-aspartate (NMDA) receptor signaling pathway. The *DLG2* gene encodes a member of the membrane-associated guanylate kinase (MAGUK) family. The protein product of *DLG2* gene is involved in processes such as biosynthetic, protein transport, cellular protein modification, cytoskeleton-dependent intra-cellular transport, cell-cell signaling, cell adhesion and cellular nitrogen compound metabolic (Figure 1, Table S4).

In turn, SNP rs110455063(C/G) is located 34,194 nt 5' from the *IGF-I* gene. *In silico* GO annotation analysis revealed that IGF-I is involved in immune processes (response to stress, immune system process, signal transduction) and growth and development, cell differentiation, cell population proliferation, and anatomical structure development (Figure 1, Table S3).

In the *DGAT1* gene intron is located SNP No. rs109421300(T/C). Protein product of *DGAT1* gene is involved not only in the lipid metabolism, sulphur metabolism or biosynthesis process, but also in the process of homeostasis and the immune system (Figure 1).

Rs109421300(T/C) SNP is located 242,691<(3') from the *CYP11B1* gene and 450,205<(3') from the *GPIHBP1* gene. GO annotation analysis indicated that the protein produced by *GPIHBP1* gene is involved in lipid metabolic processes, transmembrane transport, cellular nitrogen compound metabolic process, or biosynthetic process, and also in homeostatic process (Figure 1). No hits for *GPIHBP1* gene were found in KEGG analysis. As *in silico* GO analysis showed, CYP11B1 is involved in lipid metabolic, circulatory system, homeostatic, and immune system processes (Figure 1) while according to KEGG analysis i.a. in metabolic pathways (Table S4).

Two of the studied SNPs, rs110785912(A/T) and rs110455063(C/G), are located within LOC1049123 and LOC104972477 *loci*, respectively, thus no functional analysis is possible. These SNPs were associated with lnSCC – rs110785912 (A/T), and dry matter and fat contents – rs110455063(C/G).

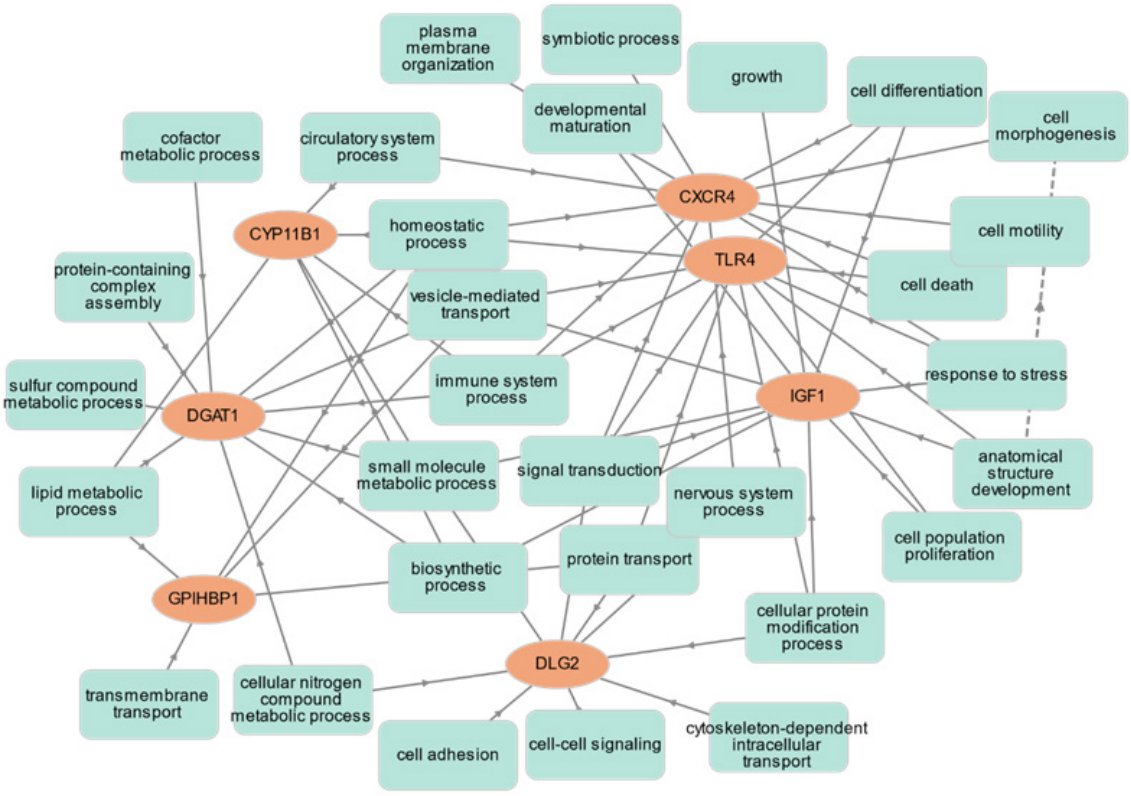

Figure 1. Common biological processes names in which protein products of the studied genes are involved – in silico GO annotation analysis

The studied genes are involved in many physiological pathways. The protein products of *IGF-I* and *TLR4* genes are involved in 35 and 28 pathways, respectively with some of them common for both genes such as proteoglycans in cancer (bta05205), HIF-1 signaling pathway (bta04066), and PI3K-Akt signaling pathway (bta04151). IGF-I and CXCR4 are involved in two common pathways (endocytosis pathway – bta04144 and pathways in cancer – bta05200), while CYP11B1 and DGAT1 in metabolic pathway (bta01100). Most of pathways in which studied genes are involved are connected with various diseases, however, some of them are involved in regulation of cellular functions such as transcription, translation, proliferation, growth, and others.

Economic analysis

The value of the studied polymorphisms for selection purposes was determined by economic analysis (Table S5). An analysis of variance was performed to determine the significance of the differences between the variants of the analyzed SNPs with regard to their obtained economic effects (Table S6). Statistically significant differences were found between the allelic variants of SNP rs136813430(T/C) for all analyzed economic categories. No such differences were found for rs110455063(C/G) or rs109421300(T/C) for any of the analyzed categories. However, for SNP rs110785912(T/A), significant differences were found between homozygous AA and heterozygous TA for the direct profitability index. Moreover, and for SNP rs41587003(A/C) significant differences were found between homozygous C/C and T/C regarding the production costs of one liter of milk (Table S3).

Based on the obtained results, SNP rs136813430(T/C) seems to be the most useful SNP for selection purposes, as it was found to have a significant influence on each of the three economic categories, with statistically significant differences observed between its variants. Cows with the T/C genotype demonstrated the highest gross margin and direct profitability index compared to other variants, and the lowest costs incurred to produce 1 liter of milk.

SNP rs110785912(T/A), related to lnSCC in milk, had a significant influence on the direct profitability index with significant differences observed between cows with the A/A and T/A genotypes. The T/A cows demonstrated the most favorable profitability index and the lowest somatic cell content.

In the case of SNP rs41587003(A/C), significant differences were found between the C/C and T/C variants regarding the production costs of 1 liter of milk. No studies have so far examined the functional features of this gene; however, our present findings suggest that this SNP is associated with lactose content in milk: variants with a higher milk lactose content demonstrated lower costs of producing one liter of milk. The lowest costs of producing 1 liter of milk were observed for the T/C heterozygotes and the highest for the C/C homozygotes.

Discussion

Animal sample

Despite the rather small sample size, as the association analysis was conducted using 299 animals with approx. 8,000 phenotypic records, our obtained values for the studied SNPs appear generally consistent with previous studies. Moreover, the standard errors (SE) of solutions for genotypes were low which also proves the low bias of the obtained results. With an average of two daughters per sire and only 12% of the dams having more than one daughter, our sample of 299 can be considered representative of the population for a polymorphism study because it is not dominated by a small number of large families (Oprządek et al., 2015).

The heritability coefficient values obtained for milk yield was within the range of estimates obtained in the Polish dairy cows' population (h^2=0.20–0.40) (Aerts et al., 2021; Rzewuska and Strabel, 2013). CV for milk yield was very high (60%) maybe because of different lactation duration.

A high CV for fat content was obtained in the present study (22.5%) and this is in line with values obtained by previous studies (Krag et al., 2013). The estimated h^2 value (0.17) was slightly lower than the coefficients obtained in the Polish active population (0.25–0.28) for the first three parities using test-day model (Rzewuska and Strabel, 2013).

The heritability coefficient for lactose content was lower (0.16) than the values reported previously (0.26–0.34) in the Polish dairy cattle population (Rzewuska and Strabel, 2013), while the low total variability was obtained in the present study (CV=9.23%). Therefore, this association analysis for the lactose content with the tested SNP should be repeated on a much greater population sample.

The heritability coefficient of lnSCC was 0.13 in the studied group of animals. Again, despite the small sample size and high total variability of lnSCC (CV=34.5%), this value did not differ from those obtained in the Polish populations of dairy cows (0.13–0.16) (Sender et al., 2013; Rzewuska and Strabel, 2013).

For dry matter content, both heritability (0.20) and variation coefficient (CV=8.37%) did not differ substantially from previous studies in Polish population (Yazgan et al., 2010).

Similarly, for fat:protein ratio, the h^2 value (0.14) and the total variability (CV=20.8%) did not differ substantially from previous values, i.e. h^2 between 0.18 and 0.40 (Rzewuska and Strabel, 2013) in a study in a big population of more than 19,000 Polish HF dairy cows (159,044 lactation records).

Hence, it can be seen that the heritability and total variation coefficients obtained in the studied group of animals did not differ actually from previous studies on Polish HF population. Moreover, as the genotyped cows were descendants of 153 sires and 263 dams, the probe from population was not overwhelmed by one or two families. As such, our findings SNPs can be regarded as representative, despite the small sample size.

Analysis of gene polymorphisms

Commercially available SNP microarrays are employed in GWAS. However, a large majority of SNPs do not correlate with breeders' interest-piqued traits. As a result, it is crucial to research all known SNPs and choose for GWAS only those that significantly affect the traits that are taken into account in the breeding goal in a particular population. Gene containing such SNPs in their structure can become the candidate for main genes or SNP is considered as a marker that indicates the quality and value of the genes linked to it.

The rs136813430(T/C) SNP, associated in our own work with lnSCC and identified in the TLR4 gene exon (ENSBTAG00000006240) (chromosome 8) is highly polymorphic: a total of 36 SNPs have been identified in 14 breeds of cattle. TLR4 is believed to initiate non-specific and adaptive immune responses that defend the body against pathogens by inducing the overexpression of the pro-inflammatory cytokines IL-1, L-6 and IL-8 involved in non-specific immunity (Wang et al., 2007). Although the identified mutation located in the exon is synonymous, it might be that the polymorphic variant is responsible for intensifying the immune response of the udder against invasion by increasing production of pro-inflammatory cytokines. This may hasten the annihilation of the pathogen without initiating the specific response and recruitment of leukocytes from the blood into the udder. Therefore, SNP rs136813430(T/C) appears to be a promising candidate in the selection of cows for resistance to mastitis, as demonstrated by the significant role played by the *TLR4* gene protein product during inflammatory processes of the udder, and our own findings regarding the relationship between the SNP and lnSCC. So far, this SNP has not been indicated as a marker of mastitis or milk composition.

Located close to the *CXCR4* (ENSBTAG 0000001060) gene (chromosome 2), the rs110785912(A/T) SNP is also associated with lnSCC. The gene is a member of the chemokine receptor family (Busillo and Benovic et al., 2007), and is involved in many developmental processes, as well as in many disease processes caused by viruses, such as bovine non-cytopathic viral diarrhoea virus (BVDV) (Weiner et al., 2012). *CXCR4* expression was found to be elevated in the secretory epithelial tissue in cow udder quarters following infection by coagulase-positive and coagulase-negative staphylococci compared to pathogen-free tissues (Kościuszuk et al., 2017). The CXCR4 receptor is believed to play a key role in lymphocyte transport and directing them to the lymph nodes (Busillo and Benovic, 2007). So far, the rs110785912(A/T) SNP has not been studied by other authors in association analyses. Based on our own results and literature data on the significant involvement of CXCR4 in the activation of the immune system, this SNP has been proposed as a marker of udder health to be used in Marked Assisted Selection (MAS). Confirmation of the close relationship between the tested SNP and SCC in milk, probably through the direct influence of the CXCR4 gene protein product on the processes of the immune system taking place in the cells of the udder tissues and leukocytes of the milk alveoli, is extremely important from the point of view of breeding work towards the selection of cows resistant to inflammation of the udder.

Further association studies are needed to check the relationships with other diseases and productivity of the cows, especially that both TLR4 and CXCR4 are involved in biological processes connected with both immune system and cell differentiation and death. All of these procedures are necessary for an effective udder to function. One of the SNP is localized in the intron of *DLG2* gene. As a heterodimer formed with a related family member, protein product of the *DLG2* gene may interact at postsynaptic sites to form a multimeric scaffold for the clustering of receptors, ion channels, and associated signaling proteins. To our knowledge, the function of DLG2 has not been studied in cattle till now. However, in humans, it has a wide range of cellular functions including those connected with nervous system, Parkinson's disease or some kinds of cancers (Zhuang et al., 2019; Shao et al., 2019). It interacts with the cytoplasmic tail of NMDA receptor subunits and with inward rectifying potassium channels (UniProt Consortium, 2019, Q15700). Until now, it has not been connected with lactogenesis. Thus, further study is needed to identify its function in bovine and its connection with lactose content, especially that DLG2 is involved in many processes including biosynthetic, protein transport, or cellular protein modification.

IGF-I gene (ENSBTAG00000011082) (chromosome 5 in bovine) protein product is also called somatomedin C (Meuwissen et al., 2002). IGF-I mediates many biological processes, increasing glucose absorption, regulating cellular differentiation, or proliferation, inhibiting apoptosis, or increasing lipid synthesis (De la Rosa Reyna et al., 2010). IGF-I is released from the liver in response to growth hormone (GH) that controls growth and lactation, thus IGF-I is thought that controlling lactation through production of milk ingredients (Lucy, 2008). It is also produced locally in a tissue-specific manner. IGF-I can be used to monitor udder health by measuring its concentration in milk (Liebe and Shams, 1998). The authors' own research showed a relationship between SNP rs110455063(C/G) located 34,194 nt 5' from the *IGF-I* gene with the dry matter, and fat contents, and daily milk yield. So far, in the association analyses, this SNP has not been associated with any trait of milk, so it is not possible to compare our results with the results of other authors. It is proposed to include the SNP found in *IGF-I* gene in MAS or genomic selection to improve cow productivity, but it is necessary to conduct further studies on greater populations of dairy cattle.

DGAT1 gene (ENSBTAG00000026356) (chromosome 14) codes for a microsomal enzyme, using diacylglycerol and fatty acetyl coenzyme A as substrates to catalyze the final stage of triacylglycerol synthesis (Cases et al., 1998). It influences fat metabolism and its yield and

percentage in milk. Many polymorphic *loci* have been found in this gene and most of them are located in the introns, promoter and untranslated regions (UTR) of exons. SNP analyzed in our study was located within this gene at position 1801,116 and was related to the fat content, dry matter content and fat to protein ratio. Meredith et al. (2013) showed a relationship between the studied SNP and the milk fat yield. Jiang et al. (2019) found a strong antagonistic pleiotropy between fat yield and milk and protein yield for this SNP; the C allele was responsible for this extremely antagonistic pleiotropy between positive fat yield and negative milk and protein yield, while the T allele had antagonistic pleiotropy between negative fat yield and positive milk and protein yield. Chen et al. (2015) showed its association with SCC during inflammation caused by *E. coli* and *S. uberis*. Many studies have found that the region that includes the *DGAT1* gene has a large impact on milk fat content and on several other production traits including milk yield, percentage of fat, and percentage of protein (Ashwell et al., 2004; Maxa et al., 2012). The fat:protein ratio was 1.15 for heterozygous cows: it should be highlighted that it is the best ratio between these traits, because it determines the best texture and taste for cheese curd. DGAT1 can be considered a candidate gene for the fat content and fatty acid profile of milk.

The *GPIHBP1* (ENSBTAG00000049125) (chromosome 14) encodes a protein belonging to the lymphocyte antigen 6 (Ly6) family. GPIHBP1 plays a key role in the transport and localization of lipoprotein lipase (LPL) synthesized by myocytes and adipocytes and creates a platform for lipolysis in endothelial cells (Yang et al., 2017). In our research rs109421300(T/C) SNP was related to the dry matter and fat contents and the fat to protein ratio. Meredith et al. (2013) showed that this SNP was related to the yield of milk fat. Fang and Pausch (2019) also proposed the inclusion of this *GPIHBP1* gene for use in selection to improve milk yield. As this SNP is located in close distance to the *GPIHBP1* gene (450,205 nt towards 3'), so it is likely to be inherited along with it.

SNP rs109421300(T/C) is closely located to *CYP11B1* gene which encodes a member of the cytochrome P450 superfamily. Proteins catalyze many reactions related to drug metabolism and the synthesis of cholesterol or steroids and other lipids. This protein localizes to the mitochondrial inner membrane and is involved in the conversion of progesterone to cortisol in the adrenal cortex. CYP11B1 affects cortisol production, androgen function, and ultimately the proliferation of mammary gland cells (Brettes and Mathelin, 2008).

Two SNPs, rs110785912(A/T) and rs110455063(C/G), located within LOC1049123 and LOC104972477 *loci* in the NCBI database are marked as genes of unknown function. These SNPs were associated with lnSC Cor dry matter, and fat contents. However, it was impossible to compare our results with the results of other authors, because there is no available literature information.

Our results for some of the analyzed SNPs differ from previous studies. These differences are probably due to the different population size or structure. It is recommended that any population used in such a study should be subjected to GWAS testing.

The analyzed SNPs were located within genes or closer than around 500,000 nucleotides (nt) from the nearest gene. As a long distance between the marker and the gene reduces the probability of inheriting the analyzed SNP with the gene, such distant loci cannot be reliable markers of the traits coded by an individual gene. Simulations (Kruglyak, 1999) and empirical analyses based on human data (Dunning et al., 2000) suggest that linkage disequilibrium (LD) extends only a few kilobases (kb) around common SNPs; however, other studies indicate that this distance can be greater than 100 kb (Abecasisi et al., 2001) or even beyond 1 Mb (Taillon-Miller et al., 2000). Nevertheless, it is suggested that due to genetic drift, admixture, selection and smaller effective population sizes, which reduce allelic heterogeneity, livestock demonstrate greater LD than humans, and LD can extend over several hundred base pairs (McRae et al., 2002).

It is possible that the analyzed SNPs which exist outside any gene sequences, may be found in enhancer (expression enhancer) or suppressor regions – they can permanently turn off expression through an epigenetic silencing mechanism; as such, they have a strong influence on gene transcription (activation) and expression. Moreover, the promoter region, a section of genomic DNA upstream of the transcription start site (TSS) of the gene commonly referred to as the +1 position, contains regulatory information for transcription initiation. Enhancer-promoter interactions precisely control the spatiotemporal timing of the activation of particular genes (Danino et al., 2015). The enhancers are cis-regulatory elements first identified in the SV40 virus genome; these can influence expression from a distance of even 1 mega base (1 Mb) and may be located within the target gene, or upstream or downstream from it. They show high specificity regarding cell type and response to stimuli. Complex regulatory genes often have multiple enhancers, with each being responsible for expression in a given cell type or situation. In addition, a single enhancer can act on many genes at once. The genes whose expression is necessary for the function of the cells in each situation may not have any enhancers or silencers. Enhancers and silencers often cooperate to maintain the balance of tissue-specific promoter activity. Most SNPs have been found in noncoding regions containing enhancers (Xia and Wei, 2019).

As the probability of inheriting a SNP with a gene is reduced when a long distance exists between marker and functional gene, loci located at a large distance cannot be reliable markers of the features encoded by a given gene, particularly when assessing the predicted breeding value of offspring. However, as genomic selection allows individuals with high breeding value to be selected based on SNP markers, all information on SNPs and their association with production traits is useful in the breeding work. Genomic information allows genetically ideal animals to

be identified at a very young age, with greater accuracy than estimates based on the average genetic value of the parents. Further studies should examine the functions of the protein products of studied genes; this would provide a better understanding of their participation in metabolic pathways and cellular processes associated with the immune system of cows and their relationship with production traits.

Economic analysis

SNP rs136813430(T/C) was found to have promise for assessing potential milk production profitability. Cows with the T/C genotype were characterized by the highest gross margin and direct profitability index, as well as the lowest costs incurred to produce one liter of milk. This SNP probably influences the number of somatic cells (SCC) possibly due to the role played by the protein encoded by the TLR4 gene, in which the SNP is located.

Many studies have confirmed the importance of Toll-like receptors (TLRs), including TLR4, in the first line of defense against invading pathogens, where they are believed to initiate the innate immune response (Kawai and Akira, 2005; De Schepper et al., 2008). Concerning the lower SCC content, the T/C variant may bestow higher resistance to inflammation of the udder. Therefore, it is an interesting candidate for the purposes of breeders.

Information about the variants of SNP rs136813430(T/C) may therefore prove useful when selecting cows with a specific genotype in programs intended to regulate the content of fatty acids in milk; proper selection will guarantee not only an appropriate fatty acid content, but also improve economic profitability.

In the case of SNP rs110785912(T/A) associated with lnSCC in milk, a significant correlation was found with the direct profitability index. The SNP lies in proximity to the *CXCR4* gene, the product of which is involved in the immune system of cattle (Weiner et al., 2012). It is therefore possible that the area within the *CXCR4* gene, and thus also in the rs110785912(T/A) locus, may be involved in the functioning of the immune system of cows; however, this requires additional research. The most favorable genotype in terms of the profitability index turned out to be the heterozygous T/A variant, which was associated with the lowest content of somatic cells. As SCC is an indirect indicator of udder health, the low cell content may mean that cows with this genotype are less susceptible to mastitis than those with other gene variants. This variant therefore offers promise in selection processes intended to increase the resistance to inflammation while ensuring high milk production profitability.

On the other hand, regarding SNP rs41587003(A/C), significant relationships were found between the C/C and T/C variants with regard to the cost of producing one liter of milk. The studied SNP is in the intron of the *DGL2* gene called also Postsynaptic Density Protein PSD-93. It encodes a protein belonging to the guanylate kinase superfamily (MAGUK). This is a huge protein complex, with more than 1,000 members, including scaffold or cytoskeletal proteins, receptors and signaling enzymes. These proteins are a part of the postsynaptic protein scaffold of excitatory synapses, and contain various domains (e.g., PDZ, GK, SH3) that enable them to bind to many of the proteins present in synapses (Zhu et al., 2016). Although no studies have examined the functional features of protein encoded by this gene, our findings indicate that the SNP lying in this gene appears to be related to lactose content in milk. Lactose level is also an indicator of udder health: its decrease often indicates inflammation of the udder, resulting in loss of production and additional costs related to treating cows (Bruckmaier et al., 2004). According to the conducted research, variants with a higher milk lactose content were characterized by a lower level of production costs per liter of milk. In this analysis, the lowest production costs for one liter of milk were recorded for T/C heterozygotes and the highest for C/C homozygotes. Therefore, there is a certain contradiction between the selection and economic goals.

To summarize, the present study examines the influence of selected SNP polymorphisms occurring in genes related to udder health and milk production on the profitability of milk production. The analysis was compared with data concerning production value and direct costs which was obtained from commercial farms. An analysis of variance revealed significant differences between allelic variants in the case of three tested SNPs. The heterozygous (T/C) variant of SNP rs136813430(T/C) was associated with a low content of lnSCC; it was also characterized by the highest gross margin, the highest direct profitability index and the lowest costs incurred to produce one liter of milk ($P = 0.01$). This therefore appears to be the most promising of the tested SNPs for selection purposes. The T/A variant of rs110785912(T/A) demonstrated low lnSCC content in milk and the highest direct profitability index, while the C/C variant of rs41587003(A/C) demonstrated lowest lactose level and the highest costs of producing one liter of milk. The obtained results indicate that there is an economic justification for including SNP variants located within, or close to, genes involved in the immune system and milk production in cattle selection.

Conclusions

While our cow sample size may seem somewhat small for this kind of study the bias can mainly regard the variance component magnitudes of the phenotypic outcome and SNP alleles' frequencies, as the entire population variation could have been narrowed in the sample. We have, however, discussed this issue with the conclusion that our sample is suitable for the undertaken research. On the other hand, the SNP-phenotypic output association analysis is much less dependent on the sample size (if at all, given our sample size).

Hence, the above let us recommend the rs136813430(T/C) SNP, located within the TLR4 gene,

as a candidate gene for explaining the variance of SCC in milk, as indicative of susceptibility to mastitis, but also for explaining the variance of the gross margin and the direct dairy profitability index. The latter is, to our knowledge, the first effort to assess directly a correlation between the DNA polymorphism and economic output of a dairy enterprise.

Acknowledgements

The authors thank prof. Marek Łukaszewicz for invaluable help in improving the quality of this work. The authors gratefully acknowledge the Polish Federation of Cattle Breeders and Dairy Farmers for phenotypic data on the daily milk yield and the component contents and pedigree file studied dairy cows.

References

Abecasis G.R., Noguchi E., Heinzmann A., Traherne J.A., Bhattacharyya S., Leaves N.I., Anderson G.G., Zhang Y., Lench N.L., Carey A., Cardon L.R., Moffatt M.F. Cookson W.O.C. (2001). Extent and distribution of linkage disequilibrium in three genomic regions. Am. J. Hum. Genet., 68: 191–197.

Aerts J., Piwczyński D., Ghiasi H., Sitkowska B., Kolenda M., Önder H. (2021). Genetic parameters estimation of milking traits in Polish Holstein-Friesians based on automatic milking system data. Animals, 11: 1943.

Andrei S., Pintea A., Bunea A., Groza I., Bogdan L., Ciupe S., Matei S., Crainic D. (2009). Non-enzymatic antioxidants concentration and lipids peroxidation level in milk from cows with subclinical mastitis. Bull. Univ. Agric. Sci. Vet. Med. Cluj. Napoca Vet. Med., 66.

Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.N., Sherlock G. (2000). Gene ontology: tool for the unification of biology. Nat. Genet., 25: 25–29.

Ashwell M.S., Heyen D.W., Sonstegard T.S., Van Tassel C.P., Da P.M., Ron M., Weller J.I. (2004). Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. J. Dairy Sci., 87: 468–475.

Bagnicka E., Łukaszewicz M., Ådnøy T. (2016). Genetic parameters of somatic cell score and lactose content in goat's milk. J. Anim. Feed Sci., 25: 210–215.

Brettes J.P., Mathelin C. (2008). Effect dual DeVuyst s androgens sur la glande mammaire. Bull. Cancer, 95: 495–502.

Bruckmaier R.M., Ontsouka C.E., Blum J.W. (2004). Fractionized milk composition in dairy cows with subclinical mastitis. Vet. Med. Czech, 49: 283–290.

Brzóska F., Kowalski Z.M., Osięgłowski S., Strzetelski J. (2014). IZ PIB-INRA Standard of ruminants' feeding: nutrient value of French and domestic fodders for ruminants (in Polish), Strzetelski J. (ed.). Foundation IZ PIB Patronus Animalium, Kraków, Poland.

Busillo J.M., Benovic J.L. (2007). Regulation of CXCR4 signaling. Biochim. Biophys. Acta, 1768: 952–963.

Carlen E., Strandberg E., Roth A. (2005). Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. J. Dairy Sci., 87: 3062–3070.

Cases S., Smith S.J., Zheng Y.W., Myyers H.M., Lear S.R., Sande E., Novak S., Collins C., Welch C.B., Lusis A.J., Erickson S.K. Farese R.V. (1998). Identification of a gene encoding an acyl CoA: diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. PNAS USA, 95: 13018–13023.

Chang C.C., Chow C.C., Tellier L.C., Vattikuti S., Purcell S.M., Lee J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience, 4: 7.

Chen X., Cheng Z., Zhang S., Werling D., Wathes D.C. (2015). 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., 5: 358–393.

Danino Y.M., Even D., Ideses D., Juven-Gershon T. (2015). The core promoter: At the heart of gene expression. BBA-Gene Regul. Mech., 1849: 1116–1131.

Davies G., Genini S., Bishop S.C., Giuffra E. (2009). An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. Animal, 3: 415–436.

De la Rosa Reyna X.F., Montoya H.M., Castrellón V.V., Rincón A.M.S., Bracamonte M.P., Vera W.A. (2010). Polymorphisms in the IGF1 gene and their effect on growth traits in Mexican beef cattle. Gen. Mol. Res., 9: 875–883.

De Schepper S., De Ketelaere A., Bannerman D.D., Paape M.J., Peelman L., Burvenich C. (2008). The toll-like receptor-4 (TLR-4) pathway and its possible role in the pathogenesis of Escherichia coli mastitis in dairy cattle. Vet. Res., 39: 5.

Dunning A.M., Durocher F., Healey C.S., Teare M.D., McBride S.E., Carlomagno F., Xu C-F., Dawson E., Rhodes S., Ueda S., Lai E., Luben R.N.., Van Rensburg E.J., Mannermaa A., Kataja V., Rennart G., Dunham I., Purvis I., Easton D., Ponder B.A.J. (2000). The extent of linkage disequilibrium in four populations with distinct demographic histories. Am. J. Hum. Genet., 67: 1544–1554.

Fang Z.H., Pausch H. (2019). Multi-trait meta-analyses reveal 25 quantitative trait loci for economically important traits in Brown Swiss cattle. BMC Genomics, 20: 695.

Hagnestam-Nielsen C., Østergaard S. (2009). Economic impact of clinical mastitis in a dairy herd assessed by stochastic simulation using different methods to model yield losses. Animal, 3: 315–328.

Jiang J., Ma L., Prakapenka D., Van Raden P.M., Cole J.B., Da Y. (2019). A large-scale genome-wide association study in U.S. Holstein cattle. Front. Genet., 10: 412.

Kanehisa M., Sato Y., Kawashima M., Furumichi M., Tanabe M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res., 44: D457–D462.

Kawai T., Akira S. (2005). Pathogen recognition with toll-like receptors. Curr. Opin. Immunol., 17: 338–344.

Kosciuczuk E.M., Lisowski P., Jarczak J., Majewska A., Rzewuska M., Zwierzchowski L., Bagnicka E. (2017). Transcriptome profiling of Staphylococci-infected cow mammary gland parenchyma. BMC Vet. Res., 13: 1–12.

Krag K., Poulsen N.A., Larsen M.K., Larsen L.B., Jans L.L., Buitenhuis B. (2013). Genetic parameters for milk fatty acids in Danish Holstein cattle based on SNP markers using a Bayesian approach. BMC Genet., 14: 1471–2156.

Kruglyak L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nature Genet., 22: 139–144.

Liebe A., Shams D. (1998). Growth factors in milk: interrelationships with somatic cell count. J. Dairy Res., 65: 93–100.

Lü A., Hu X., Chen H., Dong Y., Zhang Y., Wang X. (2011). Novel SNPs of the bovine PRLR gene associated with milk production traits. Biochem. Genet., 49: 177–189.

Lucy M. (2008). Functional differences in the growth hormone and insulin-like growth factor axis in cattle and pigs: Implications for post-partum nutrition and reproduction. Reprod. Domest. Anim., 43: 31–39.

Madsen P., Jensen J. (2013). A user's guide to DMU – a package for analysing multivariate mixed models. Version 6; release 5.2.

Marogna G., Rolesu S., Lollai S., Tola S., Leori G. (2010). Clinical findings in sheep farms affected by recurrent bacterial mastitis. Small Rumin. Res., 88: 119–125.

Maxa J., Neuditschko M., Russ I., Förster M., Medugorac I. (2012). Genome-wide association mapping of milk production traits in Braunvieh cattle. J. Dairy Sci., 92: 5357–5364.

McRae A.F., McEwan J.C., Dodds K.G., Wilson T., Crawford A.M., Slate J. (2002). Linkage disequilibrium in domestic sheep. Genetics, 160: 1113–1122.

Meredith B.K., Berry D.P., Kearney F., Finlay E.K., Fahey A.G., Bradley D.G., Lynn D.J. (2013). 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., 4: 229.

Methodology of calculating the gross margin for agricultural production activity (in accordance with the standards of the European Union (in Polish). Institute of Agricultural and Food Economics-National Research Institute (IERiGŻ), Warszawa, 1999.

Meuwissen T.H.E., Karlsen A., Lien S., Olsaker I., Goddard M.E. (2002). Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics, 161: 373–379.

NCBI-1, National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov, access May 2019.

NCBI-2, National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/genome/?term=LOC104972477, access December 14, 2021.

Oprządek J., Sender G., Pawlik A., Łukaszewicz M. (2015). Locus BoLA-DRB3 is just an ordinary site of the polygene when explaining genetic variance of somatic cell count and milk yield. J. Dairy Res., 82: 449–452.

Ptak E., Brzozowski P., Jagusiak W., Zdziarski K. (2007). Genetic parameters for somatic cell score for Polish Black-and-White cattle estimated with a random regression model. J. Anim. Feed. Sci., 16: 357–369.

Rzewuska K., Strabel T. (2013). Genetic parameters for milk urea concentration and milk traits in Polish Holstein-Friesian cows. J. Appl. Genet., 54: 473–482.

Sender G., Korwin-Kossakowska A., Pawlik A., Hammed K., Oprządek J. (2013). Genetic basis of mastitis resistance in dairy cattle – a review. Ann. Anim. Sci., 13: 663–673.

Shao Y.W., Wood G.A., Lu J., Tang Q.L., Liu J., Molyneux S., Chen Y., Fang H., Adissu H., McKee T., Waterhouse P., Khokha R. (2019). Cross-species genomics identifies DLG2 as a tumor suppressor in osteosarcoma. Oncogene, 38: 291–298.

Skarżyńska A. (2012). The impact of the milk yield of cows on the profitability of milk production (in Polish). Zagadnienia Ekonomiki Rolnej (Probl. Agricult. Econom.), 330: 90–111.

Sordillo L.M. (2005). Factors affecting mammary gland immunity and mastitis susceptibility. Livest. Prod. Sci., 98: 89–99.

Taillon-Miller P., Bauer-Sardiña I., Saccone N.L., Putzel J., Laitinen T., Cao A., Kere J., Pilia G., Rice J.P., Kwok P-Y. (2000). Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. Nat. Genet., 25: 324–328.

UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. Nucleic acids research, 47: D506–D515. https://www.uniprot.org/uniprot/Q15700, access, December 15, 2021.

Wang Y., Chen T., Han C., He D., Liu H., An H., Cai Z., Cao X. (2007). Lysosome-associated small Rab GTPase Rab7b negatively regulates TLR4 signaling in macrophages by promoting lysosomal degradation of TLR4. Blood, 110: 962–971.

Weiner C.M., Smirnova N., Webb B.T., Van Campen H., Hansen T.R. (2012). Interferon stimulated genes, CXCR4 and immune cell responses in peripheral blood mononuclear cells infected with bovine viral diarrhea virus. Res. Vet. Sci., 93: 1081–1088.

Xia J.H., Wei G.H. (2019). Enhancer dysfunction in 3D genome and disease. Cells, 8: 1281.

Yang J., Liu X., Wang D., Ning C., Wang H., Zhang Q., Jiang L. (2017). Functional validation of GPIHBP1 and identification of a functional mutation in GPIHBP1 for milk fat traits in dairy cattle. Sci. Rep., 7: 8546.

Yazgan K., Makulska J., Węglarz A., Ptak E., Gierdziewicz M. (2010). Genetic relationship between milk dry matter and other milk traits in extended lactations of Polish Holstein cows. Czech J. Anim. Sci., 55: 91–104.

Yuan Z., Li J., Li J., Gao X., Xu S. (2013). SNPs identification and its correlation analysis with milk somatic cell score in bovine MBL1 gene. Mol. Biol. Rep., 40: 7–12.

Zhu J., Shang Y., Zhang M. (2016). Mechanistic basis of MAGUK-organized complexes in synaptic development and signalling. Nat. Rev. Neurosci., 17: 209–223.

Zhuang R.J., Bai X.X., Liu W. (2019). MicroRNA-23a depletion promotes apoptosis of ovarian cancer stem cell and inhibits cell migration by targeting DLG2. Cancer Biol. Ther., 20: 897–911.

Received: 10 IV 2022
Accepted: 29 XI 2022