Effect of genetic and paratypical factors on milk production in cattle

O E Lihodeevskaya, G A Lihodeevskiy, O V Gorelik, O G Lorets and V V Stepanova
Ural State Agricultural University, 42 Karl Liebknecht str., Yekaterinburg, 620075, Russian Federation

E-mail: lixodeevskaya@mail.ru

Abstract. The influence of genetic and paratypical factors on milk productivity has been studied for many years, but the influence of genotypes was considered as belonging to a genealogical line, which does not give correct results in terms of genetics, the influence of homo- and heterozygosity. The relevance of the work lies in the fact that for the first time a genome-wide study was carried out on the breeding stock of cattle in the Sverdlovsk region and the relationships between the influence of some genes and the paratype on milk productivity were established. In order to maximally reveal the influence of genotypes, farms with the same technological conditions, but giving a significant difference in milk productivity and quality, were selected. Analysis of complete genotypes by the method of principal components showed the genetic homogeneity of the sample, which forms a large general cluster, corresponding to the general trend of Holsteinization of black-and-white cattle. The influence of the environmental factor on productivity indicators turned out to be significant, with the exception of the proportion of protein in the third lactation, all of them were statistically significantly different \( P < 0.01 \). Six genes were selected for this study: diacylglycerol-O-acyltransferase 1, leptin, growth hormone receptor gene, β-lactoglobulin, k-casein AB and CE. A high coefficient of homozygosity was revealed for two genes CSN_AB and CSN_CE (0.705 and 0.897, respectively), which may lead in the near future to purely linear animals in herds. Despite the fact that no significant difference in the influence of the genotype on productive qualities was revealed for most indicators, the actual difference in them for agricultural organizations

1. Introduction

Modern conditions of breeding work dictate selection of animals not only for phenotypic traits, but for the combined effect of genotype and paratype on the phenotype.

Before the introduction of genomic research, the influence of the genotype was considered solely by the belonging of the animal to the genealogical line. Such studies have been carried out over the years. Various patterns of influence of linear affiliation were revealed. According to Russian scientists, conducted in different regions of the country, no priority superiority of any genealogical line was revealed [1-4]. Depending on the conditions of the economy, bulls of different lines showed themselves as improvers in milk yield, fat and protein-milk yield [5-7]. All the data presented indicate that each of the genealogical lines has both outstanding representatives and completely mediocre ones, or those who have fallen into inappropriate conditions for the disclosure of their potential [8,9].
The first attempts at genotyping cattle for productivity genes were carried out using PCR [10-13]. This is a rather laborious technique, especially if the purpose of the work is to identify several polymorphisms. Today, DNA chip technologies and genome-wide association study (GWAS) have greatly simplified the task of identifying genes that play an important role in the manifestation of a trait [14, 15].

The enzyme diacylglycerol-O-acetyltransferase 1 (DGAT1) is associated with milk composition and milk yield [16,17]. Leptin (LEP) has broad effects on metabolism, food intake and assimilation, height and body weight [18-20]. The leptin marker Exon 2FB also makes it possible to determine homo- and heterozygous allelic forms associated with perinatal mortality of heifers [21]. The growth hormone receptor (GHR) gene also affects growth and metabolism [22, 23], and its marker ARS-BFGL-NGS-118998 is associated with some indicators of milk production [17, 24]. The two important protein components of milk, k-casein and β-lactoglobulin, are encoded by the CSN_AB, CSN_CE and LGB genes, respectively. They are decisive in assessing milk productivity [25]. The BB variant of the β-lactoglobulin gene is associated with a lower concentration of its protein product [26], which makes it desirable since β-lactoglobulin is the main allergen in milk [27]. Variant B k-casein is associated with a higher proportion of protein in milk compared to the A allele [26].

2. Materials and method
The research was carried out in four breeding organizations of the Sverdlovsk region. To identify the significant influence of the genotype on the phenotypic manifestation of the trait, farms were selected with the same technology for keeping and milking cows, but with a significant difference between the groups of animals selected for genotyping by productive qualities. The data for comparison was the IAS "SELEX-Dairy Cattle" database. The oldest individuals (born in 2007-2013) were selected in the groups of animals for genomic research in the amount of 48 heads, which amounted to 5.0 to 10.0% of the livestock of each of the farms.

The data obtained were analyzed using the R stats language package [28]. The results were visualized in the R environment using the graphics and ggplot2 packages [29].

3. Results
First of all, the analysis of the frequency of genotypes by the $\chi^2$ method was carried out to determine the deviations of the observed values from the expected ones (table 1).

For the genes of kappa caseins, there is a high proportion of homozygous individuals with the form of the A allele, which determines higher rates of whole milk yield. 70% and 90% of individuals are homozygous for markers CSN_AB and CSN_CE, respectively. Allele A of β-lactoglobulin (LGB) is responsible for increased milk yield. The frequency of occurrence of homozygous individuals for this allele is 23%.

Table 1. Assessment of the genotypic state of the studied sample of cattle of the Holsteinized black-and-white breed (n = 192).

| Gene | Genotypes | Allele frequency | $\chi^2$ | p  | Ca |
|------|-----------|-----------------|---------|----|----|
|      | AA        | AB              | BB      |    |    |
|      | O         | E               | O       | E  | A  | B  |
| GHR  | 0.40      | 0.31            | 0.41    | 0.49| 0.19| 0.19| 0.60| 0.40| 7.86| 0.020*| 0.522|
| CSN_AB| 0.70     | 0.58            | 0.24    | 0.36| 0.06| 0.06| 0.82| 0.18| 12.32| 0.002*| 0.705|
| CSN_CE| 0.90     | 0.86            | 0.10    | 0.14| 0.01| 0.01| 0.95| 0.05| 2.18 | 0.34  | 0.897|
| LEP  | 0.15      | 0.15            | 0.45    | 0.47| 0.39| 0.37| 0.39| 0.61| 0.44 | 0.80  | 0.525|
| DGAT1| 0.45      | 0.45            | 0.44    | 0.44| 0.11| 0.11| 0.67| 0.33| 0.00 | 1.00  | 0.557|
| LGB  | 0.23      | 0.33            | 0.59    | 0.49| 0.18| 0.18| 0.53| 0.47| 9.48 | 0.009*| 0.502|

O and E - observed and expected genotype frequencies, Ca - Robertson homozygosity coefficient, * - p-value <0.05
The proportion of cows with the desired BB genotypes for the DGAT1 gene, the product of which is associated with the proportion of fat in milk, was 11%. Variants of genotypes for genes of growth hormone receptor (GHR) and leptin (LEP), with the desired homozygous form, were 40% and 39%, respectively. The variability of homozyogosity was determined according to Robertson, the degree of homozygosity for the studied genes ranged from 0.502 to 0.897. Checking the frequencies of genotypes by the \( \chi^2 \) method showed a significant (p <0.05) deviation from the expected distribution in three genes GHR, CSN_AB and LGB.

The analysis of the complete genotypes (the totality of all studied genes) by the principal component method showed the genetic homogeneity of the sample (figure 1). All individuals are formed into a large common cluster, which corresponds to the general trend of Holsteinization of the black-and-white stock.

These samples were normally distributed (p> 0.05) for five indicators of milk yield for the first and third lactation, the mass fraction of protein for the first and third lactation, and the mass fraction of fat for the third lactation. The indicators of the mass fraction of fat for the first lactation did not correspond to the normal distribution and the influence of factors on this indicator was assessed by nonparametric methods.

In order to exclude other factors and reliably assess the relationship between the environment and / or genotype on economic indicators in the studied sample, we checked the influence of the animal's birth year and its bloodiness, having previously got rid of outliers. A total of 192 cows born in 2007-2013 were genotyped. For each individual farm, the share of the second group of cows (born in 2011-2013) was over 71%. The total share of these cows in the data was 83.8%. Due to the too small (from 1 to 18 head) subsample of the first group born in 2007-2010, the obtained results are difficult to interpret. A similar situation was revealed in the course of testing the hypothesis about the influence of blood on economically useful traits. So in the first farm, 94% of individuals are in the 88-100 blood group, while in other farms this indicator does not exceed 20%.

![Figure 1](image1.png)

**Figure 1.** PCA results, location of samples in the plane: a - first and second principal components; b - second and third.

The farm, as an environmental factor, significantly affects milk production. With the exception of the percentage of protein in the third lactation, all indicators were statistically significantly different, P <0.01 (table 2). Farm 1 turned out to be similar to the others only in the proportion of lactation protein 3. The second and third farms were statistically indistinguishable in terms of the percentage of protein for the third lactation and the percentage of fat for the first and third. Also, the indicators of milk yield for 3 lactations, and the proportion of protein in both lactations were similar in the second and fourth farms. The highest values of milk yield were observed in the third farm, the lowest in the second, the
second farm differs from the others in the same wide scatter of data in all productivity indicators relative to the other farms (figure 2).

The highest proportions of fat are observed in the first farm, the lowest correspond to the second and third. Protein shares varied with the largest average in the third farm, but by the third lactation, no differences were observed.

**Figure 2.** Distribution of the milk production traits, depending on the farms. Black solid line - medians, white dotted line - mean values, black dots – outliers.
Table 2. Result of pairwise comparison of productivity traits depending on the farm.

| Farm | Milk yield per 305 days 1 lactation, kg | Milk yield per 305 days 3 lactation, kg | Protein per 305 days 1 lactation, % | Protein per 305 days 3 lactation, % | Fat per 305 days 1 lactation, % | Fat per 305 days 3 lactation, % |
|------|---------------------------------------|----------------------------------------|--------------------------------------|--------------------------------------|------------------------------|------------------------------|
| 1–2  | < 0.001                               | < 0.001                                | 0.0027                               | 0.2171*                              | < 0.001                     | < 0.001                     |
| 1–3  | < 0.001                               | < 0.001                                | 0.2135*                              | 0.4505*                              | < 0.001                     | < 0.001                     |
| 1–4  | < 0.001                               | < 0.001                                | 0.0001                               | 0.0680*                              | < 0.001                     | < 0.001                     |
| 2–3  | < 0.001                               | < 0.001                                | 0.0011                               | 0.5790*                              | 0.1393*                     | 0.8386*                     |
| 2–4  | 0.0042                                | 0.7509*                                | 0.5598*                              | 0.9714*                              | 0.0028                      | < 0.001                     |
| 3–4  | < 0.001                               | < 0.001                                | 0.0001                               | 0.4518*                              | 0.1144*                     | < 0.001                     |

*a* - p-values > 0.01 The p-values calculated using the Student’s t-test are presented. Fat at 305 days of first lactation was checked using the Mann-Whitney U-test.

When comparing phenotypic and genotypic traits in the studied group of animals, a significant difference in the influence of genotypes within a gene on productive characteristics was not shown for most of the genetic variants presented in table 3 (a, b). Only the associations CSN_AB with milk yield and milk fat of the first lactation turned out to be significant, variant BB determines higher milk yield, and heterozygote, for this gene, has a high proportion of protein.

The result we have obtained runs contrary to the data obtained by other researchers. Another gene that has an impact, DGAT1 heterozygotes of which give an increased milk yield. The AB genotype of the CSN_CE and LEP gene significantly increases the protein level in milk relative to homozygotes. The search for associations between productivity genes and useful traits within each individual farm showed the absence of general patterns in the correlation. An increase in milk yield was observed in cows with a homozygous form of BB growth hormone receptor and homozygous BB leptin gene in the first farm. The same genotype of leptin also increases the mass fraction of fat in milk.

For the DGAT1 gene, an effect on milk yield and its increase was found in the AA form of this gene in comparison with BB. High rates of milk yield were observed in the second farm in individuals of the BB gene β-lactoglobulin, relatively homo- and heterozygotes. Homozygotes of the LGB gene showed a significant difference in the proportion of fat and protein in milk; in both cases, the values were reduced in the AA genotype. Milk protein was also lower in the AA homozygote of the CSN_AB gene, relative to the other homozygote. In a fourth farm, it was found that the AB genotype of the DGAT1 gene significantly increases the proportion of fat in milk, and its BB form decreases milk protein. In genotypes that did not show a significant difference, however, there is a significant for economic use variation in the choice of animals for solving on-farm tasks.

In accordance with the data obtained, it can be concluded that in modern conditions, when forming highly productive herds, it is necessary to select animals for specific genotypes, to take into account those environmental factors in which the genetic potential will be fully realized.

Table 3. Economically useful milk productivity traits depending on the genotype.

| Traits          | GHR |             | CSN_AB |             | CSN_CE |             |
|-----------------|-----|-------------|--------|-------------|--------|-------------|
|                 | AA  | AB          | BB     | AA          | AB     | BB          |
| Milk yield      | Mean. kg | 6686.5 | 6165 | 7032 | 6223 | 7069 | 7688 | 6457 | 6669 | 8946 |
| per 305 days 1  | p-value | 0.61 | 0.01* | 0.23 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 |
| lactation       |     |             |        |             |        |             |
| Milk yield      | Mean. kg | 8074 | 8096.5 | 8764 | 8047 | 8552.5 | 8794 | 8213 | 8142 | 9111 |
| per 305 days 3  |     |             |        |             |        |             |
| days 1 lactation|     |             |        |             |        |             |
### Acknowledgments

The studies were carried out within the framework of the research topic "Genotyping and phenotyping of the cattle population of the Sverdlovsk region by genes associated with productivity" registration number of EGISU R&D AAAA-A20-120110300008-2.

---

#### Table A

| Traits                              | LEP | DGAT1 | LGB |
|-------------------------------------|-----|-------|-----|
|                                     | AA  | AB    | BB  | AA  | AB   | BB  | AA  | AB    | BB  |
| Milk yield per 305 days 1 lactation |     |       |     |     |      |     |     |       |     |
| Mean. kg                           | 7284 | 6541  | 6356| 6464.5| 6511 | 6281| 6256| 6484  | 6867 |
| p-value                            | 0.69 | 0.69  | 0.69| 0.69  | 0.69 | 0.69| 0.69| 0.69   | 0.69 |
|                                   | 1905.8| 1572.4 | 1661.4| 1619 | 1734.6| 1492.2| 1795.7| 1633.4| 1566.2|
|                                   | 0.62 | 0.36  | 0.56| 0.56  | 0.56 | 0.56| 0.56| 0.56   | 0.56 |
| Milk yield per 305 days 3 lactation |     |       |     |     |      |     |     |       |     |
| Mean. kg                           | 1857.9| 1559.4 | 1784.8| 1575.8| 1812.7| 1390.9| 1616.9| 1680.7| 1860.9|
| p-value                            | 0.6  | 0.6   | 0.6 | 0.6   | 0.6  | 0.6 | 0.6 | 0.6    | 0.6  |
|                                   | 3.11 | 3.11  | 3.11| 3.11  | 3.11 | 3.11| 3.11| 3.11   | 3.11 |
|                                   | 0.15 | 0.15  | 0.15| 0.15  | 0.15 | 0.15| 0.15| 0.15   | 0.15 |
|                                   | 0.02*| 0.02* | 0.02| 0.02* | 0.02*| 0.02*| 0.02*| 0.02*  | 0.02*|
| Protein per 305 days 3 lactation   |     |       |     |     |      |     |     |       |     |
| Mean. %                            | 3.11 | 3.11  | 3.11| 3.11  | 3.11 | 3.11| 3.11| 3.11   | 3.11 |
| p-value                            | 0.14 | 0.14  | 0.14| 0.14  | 0.14 | 0.14| 0.14| 0.14   | 0.14 |
|                                   | 0.12 | 0.12  | 0.12| 0.12  | 0.12 | 0.12| 0.12| 0.12   | 0.12 |
|                                   | 0.78 | 0.78  | 0.78| 0.78  | 0.78 | 0.78| 0.78| 0.78   | 0.78 |
| Fat per 305 days 1 lactation       |     |       |     |     |      |     |     |       |     |
| Mean. %                            | 3.97 | 4.06  | 4   | 4.03  | 4    | 4.11| 4.05| 4.02   | 3.98 |
| p-value                            | 0.3  | 0.28  | 0.25| 0.28  | 0.27 | 0.25| 0.27| 0.28   | 0.25 |
|                                   | 0.4  | 0.4   | 0.4 | 0.4   | 0.4  | 0.4 | 0.4 | 0.4    | 0.4  |
| Fat per 305 days 3 lactation       |     |       |     |     |      |     |     |       |     |
| Mean. %                            | 3.91 | 3.97  | 3.95| 3.95  | 3.96 | 3.95| 3.95| 3.95   | 3.95 |
| p-value                            | 0.23 | 0.21  | 0.23| 0.24  | 0.22 | 0.17| 0.19| 0.22   | 0.22 |
|                                   | 0.4  | 0.88  | 0.83| 0.88  | 0.83 | 0.83| 0.83| 0.83   | 0.83 |

* - p-values > 0.05
References

[1] Tselishcheva O N 2009 Influence of blood and linearity on milk productivity of cows Agrarian Russia 10 31-3
[2] Khimich, N G, Nesterenko N N and Kochneva M L 2012 Productivity of black-spotted cows of the Priavsk type depending on the lineage Achievements of science and technology of the agro-industrial complex 3 46-8
[3] Tokova F M and Ulimbashev M B 2016 Realization of the genetic potential of dairy productivity of Holstein cattle of different lineage News of Altai State Agrarian University 3(137) 108-11
[4] Gorelik O V, Fedoseeva N A and Knyshev I V 2019 Milk productivity of cows of Holstein lines of black-and-white cattle News of St. Petersburg State Agrarian University 3(56) 99-105
[5] Kazantseva E S 2015 Indicators of productive longevity of black-and-white cows depending on the lineage Agrarian News of the Urals 6(136) 51-53
[6] Randelin A V, Kaidulina A A, Barmina T N and Surkova S A 2018 Milk productivity of Danish cows of different lineage News of the Nizhnevolzhsky agro-university complex: science and higher professional education 2(50) 241-5
[7] Kakhikalo V G and Nazarchenko O V 2012 Pedigree and productive qualities of daughters of bulls-producers of Holstein lines in the conditions of the Trans-Urals Agrarian News of the Urals 4(96) 11-4
[8] Kokhanov M A, and Ignatov A V 2018 Milk productivity of cows of different lines Agrarian News of the Urals 9 94-5
[9] Pyankova S U 2014 Influence of sires of different genotypes on economically useful qualities of black-and-white cows of the Ural type News of the Orenburg State Agrarian University 6(50) 118-20
[10] Bleck G T and Bremel R D 1993 Correlation of the alpha-lactalbumin (+15) polymorphism to milk production and milk composition of Holsteins Journal of dairy science 76(8) 2292-8
[11] Zadworny D and Kühnlein U 1990 The identification of the kappa-casein genotype in Holstein dairy cattle using the polymerase chain reaction Theoretical and applied genetics 80(5) 63-4
[12] Ng-Kwai-Hang K F, Zadworny D, Hayes J F and Kühnlein U 1991 Identification of kappa-casein genotype in Holstein sires: a comparison between analysis of milk samples from daughters and direct analysis of semen samples from sires by polymerase chain reaction Journal of dairy science 74(8) 241-5
[13] Lucy M C, Hauser S D, Eppard P J, Krivi G G, Clark J H, Bauman D E, and Collier R J 1993 Variants of somatotropin in cattle: gene frequencies in major dairy breeds and associated milk production Domestic animal endocrinology 10(4) 32-3
[14] Gurgul A, Jasielczuk I, Szmatola T, Pawlina K, Ząbek T, Żukowski K and Bugno-Poniewierska M 2015 Genome-wide characteristics of copy number variation in Polish Holstein and Polish Red cattle using SNP genotyping assay Genetica 143(2) 14-55
[15] Liu L, Zhou J, Chen C J, Zhang J, Wen W, Tian J, Zhang Z and Gu Y 2020 GWAS-Based Identification of New Loci for Milk Yield, Fat, and Protein in Holstein Cattle Animals: an open access journal from MDPI 10(11)
[16] Fang M, Fu W, Jiang D, Zhang Q, Sun D, Ding X, and Liu J. 2014 A multiple-SNP approach for genome-wide association study of milk production traits in Chinese Holstein cattle PloS one 9(8)
[17] Grisart B, Coppingers W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mni M, Reid S and Simon P. 2002 Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition Genome Res. 12 22-31
[18] Nkrumah J D, Li C, Yu J, Hansen C, Keisler D H and Moore S S 2005 Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feeding behavior, and measures of carcass merit J Anim Sci. 83 20-8
[19] Lusk J 2007 Association of single nucleotide polymorphisms in the leptin gene with body weight
and backfat growth curve parameters for beef cattle J Anim Sci. 85 1865-72

[20] Banos G, Woolliams J, Woodward B, Forbes A and Coffey M 2008 Impact of single nucleotide polymorphisms in leptin, leptin receptor, growth hormone receptor and diacylglycerol acyltransferase (DGAT1) gene loci on milk production feed and body energy traits of UK dairy cows J Dairy Sci. 91 3190-200

[21] Brickell J S, Pollott G E, Clempson A M, Otter N and Wathes D C 2010 Polymorphisms in the bovine leptin gene associated with perinatal mortality in Holstein-Friesian heifers Journal of dairy science 93(1) 340-7

[22] Sherman E L, Nkrumah J D, Murdoch B M, Li C, Wang Z, Fu A and Moore S S 2008 Polymorphisms and haplotypes in the bovineneuropeptide Y growth hormone receptor ghrelin insulin-like growth factor 2 and uncoupling proteins 2 and 3 genes and their associations with measures of growth performance feed efficiency and carcass merit in beef cattle J. Anim. Sci. 86 1-16

[23] Di Stasio L, Destefanis G, Brugiapaglia A, Albera A and Rolando A 2005 Polymorphism of the GHR gene in cattle and relationships with meat production and quality Anim. Genet. 36 138-40

[24] Jiang L, Liu J, Sun D, Ma P, Ding X, Yu Y, and Zhang Q 2010 Genome wide association studies for milk production traits in Chinese Holstein population PloS one 5(10)

[25] Curi R, Oliveira H, Silveira A and Lopes C 2005 Effects of CSN3 and LGB gene polymorphisms on production traits in beef cattle Genetics and Molecular Biology 28(2) 262-6

[26] Heck J M, Schennink A, van Valenberg H J, Bovenhuis H, Visker M H, van Arendonk J A and van Hooijdonk A C 2009 Effects of milk protein variants on the protein composition of bovine milk. Journal of dairy science 92(3) 1192-202 R A Wood et al. 2013 The natural history of milk allergy in an observational cohort The Journal of allergy and clinical immunology 131(3) 805-12

[27] Team Core R 2020 R language and environment for statistical computing Foundation for Statistical Computing (Vienna, Austria) Retrieved from: https://www.R-project.org/

[28] Wickham H 2016 ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag) p 182