AGRICULTURAL BIOLOGY,
2020, V. 55, № 2, pp. 257-274
(SEL'SKOKHOZYAISTVENNAYA BIOLOGIYA)

ISSN 2412-0324 (English ed. Online)
ISSN 0131-6397 (Russian ed. Print)
ISSN 2313-4836 (Russian ed. Online)

UDC 636.2:636.082:577.21

doi: 10.15389/agrobiology.2020.2.257eng
doi: 10.15389/agrobiology.2020.2.257rus

GENOMIC VARIABILITY ASSESS FOR BREEDING TRAITS IN HOLSTEINIZATED RUSSIAN BLACK-AND-WHITE CATTLE USING GWAS ANALYSIS AND ROH PATTERNS

A.A. SERMYAGIN[1], O.A. BYKOVA[2], O.G. LORETTS[2], O.V. KOSTYUNINA[1], N.A. ZINOVIEVA[1]

[1]*Ernst Federal Science Center for Animal Husbandry,* 60, pos. Dubrovitsy, Podolsk, Moscow Province, 142132 Russia, e-mail alex_sermyagin85@mail.ru (✉ corresponding author), kostolan@mail.ru, n_zinovieva@mail.ru;
[2]*Ural State Agricultural University,* 42, ul. Karl Liebknecht, Ekaterinburg, Sverdlovsk Province, 620075 Russia, e-mail olbyk75@mail.ru, rector.urgau@yandex.ru

ORCID:
Sermyagin A.A. orcid.org/0000-0002-1799-6014
Bykova O.A. orcid.org/0000-0002-0753-1539
Loretts O.G. orcid.org/0000-0002-9945-5691
Kostyunina O.V. orcid.org/0000-0001-8206-3221
Zinovieva N.A. orcid.org/0000-0003-4017-6863

The authors declare no conflict of interests

Acknowledgements:

Supported financially by the Ministry of Science and Higher Education of the Russian Federation (theme No. AAAA-A18-118021590134-3) and the Ministry of Agriculture of the Russian Federation (theme No. AAAA-A19-119031590044-3)

Received January 31, 2020

Abstract

For using the pedigree information recorded in dairy herds rational, the creation of reference groups with a high reliability the estimated breeding value of animals is required. This is a necessary requirement not only for the genetic assessment procedure, but also for the introduction of genomic selection methods. However, without high-quality phenotyping cannot be this achieved. Therefore, it seems relevant to conduct model studies using the highly productive herd of Holsteinizated Russian Black-and-White cattle in the Urals as an example to study the genomic variability of phenotypic traits in animals' different generations using data from genome-wide analysis (GWAS) and runs of homozygosity (ROH), as well as to replenish the Russian reference population. The novelty of the work is to assess the dynamics of genomic inbreeding variability (FROH) in animal generations ("dam of mother"—"mother"-"daughter") and its comparison with direct genomic value (DGV), as well as the search for new mechanisms to confirm the scientific hypothesis of using a limited experimental dataset in GWAS. The object of this research was 76 cows and heifers of Holsteinizated Russian Black-and-White breed, as well as 9 Holstein sires genotyped for 139 thousand SNPs by the Bovine GGP HD platform (Illumina/Neogen, USA). To calculate the DGVs of studied animals the Russian reference bulls and cows' group that include 591 individuals was used. GWAS and ROH analyzes were performed based on 110448 SNPs. The reliability (p-values) of genome-wide associations with direct cows' phenotypes ranged from 2.31×10^{-5} to 1.08×10^{-7}. Quantitative traits loci on autosomes BTA1, BTA2, BTA5, BTA7, BTA8, BTA10, BTA11, BTA12, BTA14, BTA16, BTA20, BTA21, and BTA26 were found. For milk yield a region on BTA14 (1.44-1.59 Mb) with the genes *ZNF16*, *ARHGAP39*, *ZNF7* associated with an increased fat milk yield was detected. For the number of inseminations found SNPs included into the genes (*ARHGAP31*) or located close to the genes (*SERPINA5*) and associated with the growth intensity to a mature state, as well as ovarian function in animals. The characteristic of ROHs depending on the length of their fragments in the genome is given. Conservative homozygous regions on BTA12, BTA14, BTA26, and BTA29 and the most significant genes entering them were identified, which are potentially associated with selection pressure in the studied population mainly by milk production traits, reproduction, and udder type measurement parameters. The value of FROH significantly ($p < 0.05$-0.001) increased in the offspring—parent generations: by +0.012 or 1.2 % for mothers, and +0.029 for daughters. The highest values of FROH = 0.135 were noted for bulls that were signed as fathers of cows (generation of "mother") and heifers (generation of "daughter"). Each subsequent generation of individuals showed an average increase in DGV for milk yield by +94.2 kg, fat milk yield by +4.4 kg and protein milk yield by +3.0 kg, reflecting clear the strategy to improve milk production traits in the herd for obtaining new cattle genotypes. Thus, the possibilities for assessing the variability of direct cattle phenotypes, as a model for studying

genomic variability in a single herd, based on the search for associations and loci in the genome under selection pressure are shown.

Keywords: cattle, GWAS, ROH, reference population, genomic inbreeding, genomic evaluation, milk production, fertility

Over the past century, methods for improving animal populations have undergone a number of changes: from selection of individuals by their ancestors and phenotype, to the offspring's quality evaluation, and further on by the totality of genetic markers. The key to the successful development of genomic selection approaches is to obtain accurate phenotypic information on economic and biological traits. Therefore, the efficiency of further measures of improving the genetic potential in dairy cattle breeding depends both on the quality of pedigree records and on the chosen strategy for animal breeding in the population. Datasets accumulated in cattle herds are increasing annually. The genomic passport is added to the pedigree background and phenotype of the animal, which allows you to clarify the origin of the individual and its genetic value [1, 2].

Studies on the genomic architecture of animal productivity traits are actively being conducted both in Russia and worldwide. The characteristic of variability of quantitative and qualitative parameters can be measured in the components that are genetic, phenotypic, and adjusted for a number of regression factors. For example, previous genome-wide associative studies (GWAS) on the holsteinized Black Pied livestock population of Moscow Oblast and Leningrad Oblast showed the possibility of using estimated breeding values (EBV) in searching for single nucleotide polymorphisms (SNPs) associated with QTL. Significant mutations were found in the genes DGAT1 ($p = 6.8\times10^{-22}$), PLEC ($p = 6.9\times10^{-20}$) and GRINA ($p = 4.2\times10^{-10}$) and they are associated with the fat content in milk [3]. GWAS among different populations of dairy cattle in the USA, Germany, Holland, Australia and China revealed a pool of common genes responsible for indicators of milk productivity of cows and the amount of milkfat in particular: DGAT1, SCD1, GHR, EPS8, GPAT4, casein cluster genes (Hapmap24184-BTC-070077) [4-7]. It is worth noting that the pleiotropic effect of a gene was observed in the Holstein cattle population of Chinese origin according to DGAT1, in terms of the content and yield of milk protein and milk yield [8].

Simultaneously, there are studies on using absolute phenotypic values or "direct" phenotypes in GWAS. Highly reliable associations with the main genes were obtained for meta-analysis of complex of milk productivity traits, fertility and body type of dairy cattle [9], and for evaluation of feed efficiency in beef cattle [10], which indicates genes' determination of phenotypic variability. Similar studies are conducted on Large White, Landrace, and Duroc pigs, showing moderate convergence in the search for quantitative trait loci (QTL) in animals belonging to different populations [11, 12, 13]. It is worth noting that the main factor of GWAS in this case is the accuracy of phenotyping of characters and strict consideration of environmental factors for each single experiment. The practice of using data from long-term observations is more typical for research herds in Europe, North America and Australia with the aim of obtaining "pure" animal phenotypes not just for standard productivity traits, but also for indicators requiring laborious accounting (methane emission, residual feed intake, fat and amino acid composition of milk and meat, etc.) [14].

Another powerful tool of genomic analysis is the assessment of animal individual autozygosity or extended homozygous nucleotide fragments — ROH (Runs of Homozygosity) patterns. Using an example of analysis of demographic events in populations of Pinzgauer, Brown Swiss and Tyrolean cattle breeds, thresholds are determined by ROH length categories, based on the probability of inheritance of genome segments from a common ancestor for 50, 25, 6 and 3

generations ago (respectively, with ROH >1, >2 , >8 and >16 Mb) [15]. Fixation of ROH with a low frequency of recombination of loci in animal generations may indicate formation of gene clusters under evolutionary pressure – primarily, positive selection for productivity traits, as well as adaptation to changing environmental conditions [16, 17]. Nevertheless, ROH patterns allow us to evaluate the level of genomic inbreeding (F_{ROH}) both in a population (herd) and for a single animal more efficiently, relative to the pedigree information. Thus, the average variability of the F_{ROH} values between four herds of Russian Black Pied cattle of the Leningrad Oblast was 5.5-8.0% for the 46-85 selected animals [18].

The abovementioned GWAS and ROH results are more relevant to the study of mechanisms of variability of the genetic architecture of quantitative and qualitative traits in animals. The practical direction of using genomic data is focused on developing methods for early selection of the best individual genotypes, i.e. the genomic selection. This approach is a routine procedure worldwide, but in Russia it is at the stage of development and implementation. The first results show that the values of the genomic forecast should be no lower than +900 kg of milk, +31 kg of milkfat and +23 kg of protein for the young bulls of Russian origin that are being evaluated in order to use them in the reproduction of herds [19].

The scientific novelty of this work, built as an experiment, is the integration of genomic analysis and forecasting methods for animals of three generations that are simultaneously in a highly productive herd, with the aim of developing a model for their selection by a set of indicators – the level of genomic inbreeding and genomic breeding value in the early stages. Approaches for testing the hypothesis about the possibility of using direct phenotypic data on cattle productivity in searching for associations with individual nucleotide mutations and determining loci in the genome under evolutionary pressure are considered. For the first time on the basis of the Russian reference group of stud bulls and cows, it has been shown that its expansion can be achieved by including a population of holsteinized Russian Black Pied cattle from Ural.

The objective of our comprehensive research was to search for genome-wide associations, as well as loci that are associated with the intensity of selection in the population of holsteinized Russian Black Pied cattle from Ural for characterizing the genomic variability of trait of milk production and fertility of animals belonging to different offspring generations.

The research aims were: i) to conduct genotyping of 138 thousand SNPs of stud bulls of the Holstein breed and Holsteinized Russian Black Pied cows and heifers in the generations of mother's mother—mother—daughter, owned by PJSC "Kamenskoye", Sverdlovsk Province. In accordance with the biochip design, determine the carriage of valuable alleles of genes associated with the qualitative and quantitative composition of milk, as well as recessive mutations that cause loss of reproductive ability; ii) to characterize GWAS with direct indicators of productivity and reproductive qualities of cows in order to search for loci of quantitative traits in the genome of Ural cattle; and iii) to evaluate the level of genomic inbreeding in generations for the following groups: mother's mothers—mothers—daughters and bull-fathers of cows and heifers, in order to clarify the information in the analysis of the selection results of parental pairs, as well as the search for regions in the genome (ROH) that are subjected to evolutionary pressure. To calculate DGV for traits of milk production and to study the dynamics of its change in different generations of descendants.

Materials and methods. The Bovine GGP 150K biochip (Illumina/Neogen, USA) with overlap of 138974 SNP was used for the animal genotyping procedure. Based on the results of quality control of reading genomic information and conducting pedigree counting, genotypes with a presence of at least 110,448 SNPs for

85 animals were selected, including 9 stud bulls of Holstein breed. (OJSC "Uralplemcentr"), Holsteinized Russian Black Pied cows and heifers, which are descendants of these bulls, in the amount of 76 animals (PJSC "Kamenskoye" in Sverdlovsk Province). The average proportion of Holstein genes in the sample of Russian Black Pied breed was 97.7% with fluctuations from 75% to 99%. In order to study the dynamics of accumulation of homozygosity and the genomic breeding value in generations, experimental groups of animals were formed and individuals were selected in three generations of the ancestors and their descendants: mothers' mothers (25 cows, 2012-2014 years of birth), mothers (30 cows, 2015-2017 years of birth) and daughters (21 heifers, 2018-2019 years of birth). The milk productivity in the mothers' mothers and mothers groups was, on average, 7666 and 8461 kg of milk per 305 days long lactation, with 3.95% and 3.99% of mass fraction of fat, 3.20% and 3.21% of mass fraction of protein, 302.0 and 338.1 kg of milkfat, as well as 245.0 and 274.4 kg of milk protein.

The editing of biochip data for the construction of adapted extension files (.ped, .map, .fam, .bed, .bim), as well as the calculation of genome-wide associations, was performed in the Plink 1.9 program [20]. The distribution diagrams of GWAS and ROH on the chromosomes were constructed using the "qqman" and "ggplot2" packages in the "R" programming and visualization language. The averaged indices of "direct phenotypes" of 44 cows for 1-5 complete lactations were used in the GWAS. Based on the biochip design (https://genomics.neogen.com/pdf/slicks/ggp_bovine150k.pdf), selection-significant point mutations included in genes and having polymorphic variants of genotypes were pre-selected: *DGAT1* (diacylglycerol O-acyltransferase 1); *LEP* (leptin); *kCSN_A(CE)* (kappa-casein (A, C, E alleles); *kCSN_AB* (kappa-casein (A, B alleles); *HH1, HH3* (fertility haplotypes); *GHR* (growth hormone gene receptor); *CSN2_I* (beta-casein (allelic variant I); *ABCG2* (ATP-binding cassette of the G subfamily); *BLG* (beta-lactoglobulin); *YellowFat* (accountable for changes in the carotene content in fat cells).

The ROH patterns were studied using the cgaTOH program [21] in a combined selection of cows (grandmothers, mothers), heifers (daughters) and bulls-fathers for calculating the genomic inbreeding coefficient in each group. The conditions were accepted according to the Ferenčaković method [15], which provided for the differentiation of ROH into groups, according to their length in connection with the moment of the occurrence of a demographic event: [1; 2], (2; 4], (4; 8], (8; 16] and > 16 Mb. The following results were visualized and the common ROH segments that were found with the highest frequency in at least 30-40% of animals were determined (due to the small experimental selection sample, the threshold was reduced to 30%).

Summary of the identified associations between SNPs and phenotypes of animals, as well as the localization of ROH patterns and genes included, was performed by the cattle genome assembly Bos_taurus_UMD_3.1.1 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000003055.6, access date 01.30.2020). The "CattleQTLdb" database [22] was used for searching the QTL and regions under evolutionary pressure on chromosomes that are interfaced with the functional characteristics of animals.

The calculation of genomic estimates of breeding value (DGV) based on the sum of SNP markers for a selected population of the three generations brood stock and bulls according to traits of milk productivity was carried out on the basis of the Russian reference group of animals in the amount of 591 animals [19], according to the GBLUP algorithm proposed by VanRaden [23].

Results. At the first stage of studies on a selected population of the herd, the frequencies of genes associated with milk composition and animal fertility, the

level of polymorphism, and their genetic equilibrium in the Holstein cattle population in Ural were determined (Table 1).

1. Estimation of occurrence frequency of breeding-valuable genetic markers in the studied cattle population of Holsteinized Russian Black Pied breed (n = 85, PJSC "Kamenskoye", Sverdlovsk Province, 2019)

Gene	Frequency	Genotypes			Allele frequency		χ^2	*Ca*
		11	12	22	1	2		
DGAT1	O	0.389±0.035	0.474±0.036	0.137±0.025	0.626	0.374	0.014	0.532
	E	0.392	0.468	0.140				
LEP	O	0.200±0.021	0.547±0.026	0.253±0.022	0.474	0.526	1.816	0.501
	E	0.224	0.499	0.277				
kCasein_A(CE)	O	0.792±0.017	0.208±0.017	0	0.896	0.104	3.894	0.813
	E	0.803	0.187	0.010				
kCasein_AB	O	0.531±0.021	0.396±0.020	0.073±0.011	0.729	0.271	0.734	0.605
	E	0.532	0.395	0.073				
HH3	O	0.990±0.004	0.010±0.004	0	0.995	0.005	0.008	0.989
	E	0.990	0.010	0				
HH1	O	0.979±0.006	0.021±0.006	0	0.010	0.990	0.032	0.979
	E	0.979	0.021	0				
GHR	O	0.646±0.035	0.333±0.034	0.021±0.010	0.813	0.187	0.849	0.695
	E	0.660	0.305	0.035				
CSN2_I	O	0.990±0.005	0.010±0.005	0	0.995	0.005	0.005	0.989
	E	0.990	0.010	0				
ABCG2	O	0.292±0.033	0.510±0.036	0.198±0.029	0.547	0.453	0.085	0.504
	E	0.299	0.496	0.205				
BLG	O	0.095±0.012	0.432±0.021	0.474±0.021	0.311	0.698	0.018	0.572
	E	0.096	0.428	0.475				
YellowFat	O	0.990±0.004	0.010±0.004	0	0.005	0.995	0.011	0.989
	E	0.990	0.010	0				

N o t e. Coding options for genotypes and allele frequencies are given here. *HH1* and *HH3* are indicated according to the chip design. Ca is Robertson homozygosity coefficient, O and E are observed and expected frequencies.

It was found that polymorphic variants of genotypes were found for all detected point mutations in the genes, with a variation in the level of homozygosity (according to Robertson) from 0.501 to 0.989. A slight deviation from genetic equilibrium (at a threshold up to $\chi^2 = 3.840$, $p < 0.05$, when equilibrium is not disturbed) was found for the kappa-casein gene of the allelic variant *A(CE)* ($\chi 2 = 3.894$). Carriage of recessive mutations in the heterozygous state with occurrence frequencies of 1.0% and 2.1% was established for the *HH1* and *HH3* fertility haplotypes, respectively.

The desired genotypes (including heterozygotes) for milk protein genes, the casein kappa form, were 20.8% (AE genotype of the kCasein_A gene(CE), 7.3% (BB genotype of the kCasein_AB gene), and only 1.0% for the beta form (A1I genotype of the CSN2_I gene) , respectively, which is generally specific for the Holstein cattle population, which was designed for production of significant volumes of whole milk, in comparison with the products of its processing (higher yield of cheese, cottage cheese, butter). The *DGAT1* and *ABCG2* genes, associated with the percentage yield of fat and the amount of milk protein, obtained a moderate frequency distribution of the desired genotypes for the alternative allele 22 (AA) — 13.7% and 19.8%, respectively, which indicates hereditarily predetermined aspects of selection according to the quantitative composition of milk components (the frequencies of the desired alleles for each of the genes were 0.374 and 0.453, respectively). The beta-lactoglobulin (BLG) gene is significant in controlling the allergenic properties of cow milk as a part of human nutrition; the occurrence frequency of the most desirable genotype 11 (AA), which was 9.5% or 31.1% for the allele 1 in the studied population, which indicates the potential for animal selection when creating herds that are specific for this gene in production of milk with low allergenic properties.

Along with indicators of milk productivity, the animal growth and development and the ability of intense fattening (bulls) are of no small importance. The

obtained data show that for livestock genotypes of the leptin gene and growth hormone gene receptor, the desired allelic variants were 20.0% (LEP_TT) and 64.6% (GHR_AA), respectively.

In order to confirm the hypothesis about the genomic dependence of the "direct" phenotypes (indicators of own productivity) on the traits of milk productivity and fertility of cows, the GWAS was performed. The division into generations of grandmothers, mothers and daughters was carried out in order to take into account the influence of the population structure with a limited set of observations, and also to identify segments in the genome that have a common identity by origin. The increase in average productivity for 1-5 lactations per generation amounted to +795 kg of milk, + 0.04% of mass fraction of fat, +0.01 of mass fraction of protein, +36.1 kg of milkfat, +29.4 kg of milk protein. Multiple analysis of variance in order to determine the influence of paratypical factors on the variability of characters showed that the year of calving, the farm and the bull-father did not have substantial significance, except for the mass fraction of protein index ($p < 0.05$ to $p < 0.001$). In this regard, it was decided to use "direct" cow phenotypes for GWAS without adjusting for differences in environmental factors.

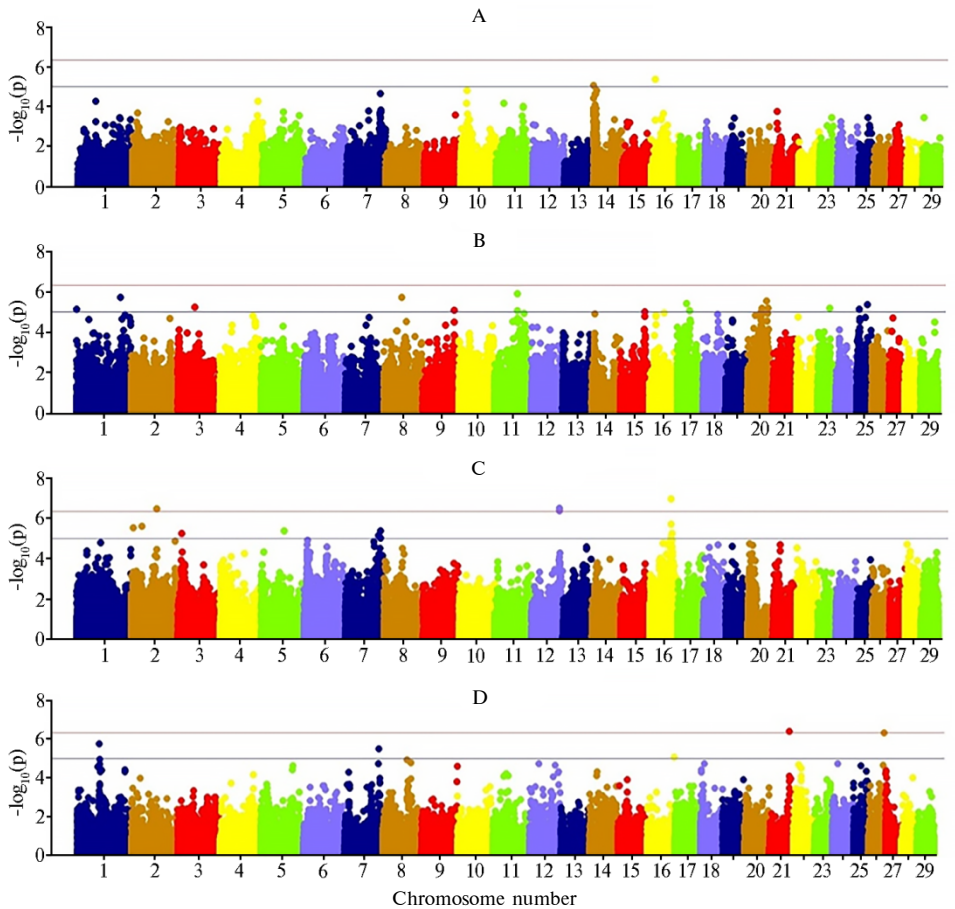

Fig. 1. Distribution of single nucleotide mutations in chromosomes of the studied cow group of Holsteinized Russian Black Pied cattle in connection with the significance level for indicators of own productivity on average per 305 days long lactation (the upper horizontal line is the significance threshold for genome-wide associations, $-\log_{10}(p) = 4{,}5\times10^{-7}$; the lower horizontal line is the significance threshold for suggestive associations $-\log_{10}(p) = 1{,}0\times10^{-5}$): A — milk yield per one lactation, B — mass fraction of fat, C — mass fraction of protein, D — conception rate ($n = 44$, PJSC "Kamenskoye", Sverdlovsk Province, 2019).

Figure 1 presents the results of the analysis of genome-wide associations with traits of milk productivity and fertility of the cows of the studied selected population. A significant correlation was established between the genotype by SNP markers and the animal's own productivity for: milk yield for 305 days long lactation on cattle chromosomes (BTA) 7, 10, 14, 16; MJ on BTA1, 8, 11, 20; CSBM on BTA2, 5, 7, 12, 16; conception rate on BTA1, 7, 8, 21, 26.

2. Significant single nucleotide mutations, associated with phenotypic traits of cows of the studied group of Holsteinized Russian Black Pied cattle (PJSC "Kamenskoye", Sverdlovsk Province, 2019)

SNP	BTA	Position, bp	p-value	Closest gene	Distance to the gene, bp
Average milk yield per 305 days long lactation, kg					
BovineHD1600003292	16	12318853	4.18×10^{-6}	*CDC73*	+325765
BovineHD1400000152	14	1439476	8.54×10^{-6}	*ZNF16*	Inside
BovineHD1400000187	14	1585385	8.54×10^{-6}	*ARHGAP39*	Inside
BTA-34956-no-rs	14	1514056	8.54×10^{-6}	*ZNF7*	+1864
BovineHD1000006525	10	19972763	1.58×10^{-5}	*HCN4*	+7449
BovineHD1400002920	14	10319796	1.58×10^{-5}	*EFR3A*	–37385
				KCNQ3	–234570
ARS-BFGL-NGS-42106	7	97772240	2.23×10^{-5}	*ELL2*	–78731
BovineHD0700028539	7	97833516	2.23×10^{-5}	*PCSK1*	+322158
BovineHD4100010843	14	7332355	2.31×10^{-5}	*KHDRBS3*	+99747
Average mass fraction of fat, %					
BovineHD1100019146	11	67734294	1.21×10^{-6}	*GFPT1*	Inside
Hapmap59899-ss46527105	11	67678534	1.21×10^{-6}	*ANTXR1*	–88003
ARS-BFGL-NGS-36975	1	128799966	1.83×10^{-6}	*ZBTB38*	–263595
BovineHD0100036417	1	128823979	1.83×10^{-6}	*SPSB4*	+6929
BovineHD0800015859	8	52869025	1.83×10^{-6}	*RFK*	–65127
BovineHD0800015899	8	52984674	1.83×10^{-6}	*PRUNE2*	Inside
BTB-00347944	8	53031744	1.83×10^{-6}	*PRUNE2*	Inside
ARS-BFGL-NGS-38258	20	56721394	2.79×10^{-6}	*MYO10*	-149181
Average mass fraction of protein, %					
ARS-BFGL-NGS-45195	16	63612072	1.08×10^{-7}	*STX6*	Inside
BovineHD4100009755	12	81991589	3.23×10^{-7}	*ITGBL1*	Inside
BTA-87771-no-rs	12	82056537	3.23×10^{-7}	*ITGBL1*	Inside
BovineHD0200021465	2	74934453	3.31×10^{-7}	*IGDCC3*	+455733
BovineHD1200023614	12	82094359	4.37×10^{-7}	*ITGBL1*	Inside
ARS-BFGL-NGS-102876	16	63521833	2.00×10^{-6}	*XPR1*	-19726
BovineHD0200009386	2	31672054	2.52×10^{-6}	*COBLL1*	Inside
BovineHD0200002141	2	7215096	2.95×10^{-6}	*COL3A1*	+102191
Hapmap53461-rs29027660	7	102855103	4.37×10^{-6}	*AP3S1*	142023
BovineHD0500018676	5	66809794	4.38×10^{-6}	*IGF1*	–206095
				PAH	+140707
Average conception rate, number					
BovineHD2100017363	21	59895529	4.13×10^{-7}	*SERPINA5*	–99716
ARS-BFGL-NGS-119213	26	47375257	4.87×10^{-7}	*DOCK1*	–79054
BovineHD0100018320	1	64773642	1.79×10^{-6}	*ARHGAP31*	Inside
BovineHD0700029413	7	100618564	3.17×10^{-6}	*CHD1*	–51472
BovineHD1600023747	16	81193491	8.92×10^{-6}	*KIF14*	–1809
				DDX59	Inside
BovineHD1600023756	16	81215628	8.92×10^{-6}	*CAMSAP2*	+62155
BovineHD0100018663	1	66151741	1.14×10^{-5}	*STXBP5L*	+61871
ARS-BFGL-NGS-108666	8	71352779	1.18×10^{-5}	*LOXL2*	Inside
BovineHD0800021457	8	71309172	1.18×10^{-5}	*LOXL2*	Inside
BovineHD0800024578	8	82536035	1.70×10^{-5}	*FBP2*	–97218

Note. For SNP, single nucleotide substitutions are presented in order of decreasing of association significance level (p-value); BTA means cattle chromosome; "+" marks the distance from the corresponding SNP to the gene, "–" marks the distance against this direction.

Based on the publicly available "CattleQTLdb" database, quantitative trait loci in the livestock genome are determined, which include genes and SNPs that are reliably associated with selection indicators in the studied Ural livestock population (Table 2). Ehe following significant polymorphisms were detected in the genes *ZNF16, ARHGAP39, EFR3A, KCNQ3, KHDRBS3*, for milk productivity, which were associated with milk yield, milk fat and protein yield, and corresponded to loci found in similar works by foreign authors [24]. The *HCN4* and

PCSK1 genes were characterized by a conjugation with growth and development parameters, as well as behavioral responses, which suggests their possible impact on milk productivity [25, 26].

Locus at position 1.44-1.59 represents the milk yield trait on BTA14, which included the three most significant genes, and at a distance of 210 kb there is the *DGAT1* gene, which plays an important role in the conversion of diacylglycerides and CoA enzymes of fatty acids into triglycerides, which are essential for synthesis of fats. Associations with a mass fraction of fat had a moderate level of significance $p = 2{,}79\times10^{-6}$ to $1{,}21\times10^{-6}$, while the *ANTXR1*, *ZBTB38*, *PRUNE2*, *MYO10* genes showed correlation with QTL, annotated with increased traits of temperament and residual feed intake, as well as growth parameters, body length and chest circumference [27-30].

More convincing results of GWAS were obtained about the mass fraction of protein, $p = 4{,}38\times10^{-6}$ to $1{,}08\times10^{-7}$. The *COL3A1, IGF1*, and *PAH* genes showed an ambiguous relationship with QTL in previous similar studies in terms of meat marbling, milk yield, percentage and yield of milk protein, live weight, conception rate, and a number of parameters for assessing the quality of the udder and legs [31-33].

At the same time, conception rate had QTL due to the functional relationship of genes with the reproductive qualities of animals responsible for: ovarian regulation and follicular atresia — *SERPINA5* ($p = 4{,}13\times10^{-7}$, including nearby variants of the super family *SERPINA1, 10, 3-2, 11*) [34]; live weight at 18 months of age (period of successful insemination) — *ARHGAP31* ($p = 1{,}79\times10^{-6}$) [35]; early preimplantation development of embryo — *CHD1* ($p = 3{,}17\times10^{-6}$) [36]; fertility of bull-daughters — *STXBP5L* ($1{,}14\times10^{-5}$) [24]. At the same time, polymorphisms in the genes were found to be correlated between the number of inseminations and the yield of milk fat — *LOXL2* ($1{,}18\times10^{-5}$) [24], the growth parameters of dairy cattle — *FBP2* ($1{,}70\times10^{-5}$) [37] and those without a link, for example, with resistance to a number of parasitic diseases with a 22.1 kb cluster on BTA16 — *KIF14, DDX59* and *CAMSAP2* ($p = 8{,}92\times10^{-6}$) [38].

The single nucleotide substitutions detected by GWAS were largely consistent with previously annotated genes and QTL for milk productivity and cattle fertility traits by a number of researchers. However, some of the SNPs only indirectly confirmed their conjugation with “direct” animal phenotypes, which indicates their pleiotropic effect, or a different nature of the combinational variability of the quantitative indicators taken into account (phenotyping accuracy, gene drift, selection intensity or lack of it).

The present-day expansion of the Holstein cattle population in one or another region of its breeding faces a number of difficulties: avoidance of close degrees of kinship, control of homozygosity in the population and lethal recessive genetic mutations, reproduction of their own breeding resources (stud bulls, rearing flocks) with given parameters of breeding values. Achieving high performance of cows is possible due to creation of an animal genotype that steadily transfers valuable alleles and/or haploblocks from generation to generation, as well as having a moderate level of homozygosity.

The pattern search algorithm is designed to scan sections of successively arranged homozygous single nucleotide substitutions; the principle is in continuous checking along the entire length of the chromosome. ROH was determined provided that 15 or more consecutive homozygous SNPs were present in the studied genome at a density of at least 1 SNP for every 100 kb, with gaps between them of no more than 1000 kb. A number of conditions were implemented as well:

starting from the ROH group with a length of (8; 16] Mb, 2 missing SNPs per segment were allowed, and then, at a level >16 Mb, 4 missing nucleotide substitutions and 1 heterozygous allele were allowed. Table 3 shows the statistics of the SNP number, included in the ROH of different lengths depending on the implemented allowances.

3. Statistical parameters of the number of single nucleotide polymorphisms (SNP) **for homozygosity patterns** (ROH) **of different lengths in the studied group of Holsteinized Russian Black Pied cattle** (n = 85, PJSC "Kamenskoye", Sverdlovsk Province, 2019)

Statistics	ROH length by groups, Mb				
	[1;2]	(2;4]	(4;8]	(8;16]	> 16
SNP number	58.6±0.4	122.0±1.1	233.8±2.0	449.8±5.0	904.2±27.0
Standard deviation	22.3	44.4	67.8	110.2	268.3
Minimum	15	15	27	253	617
Maximum	289	476	668	1050	2214
Average ROH length, Mb	1.40±0.01	2.82±0.01	5.49±0.03	10.77±0.10	22.01±0.65
Total ROH length, Mb	48.3±1.2	59.0±2.0	71.0±2.6	63.8±3.4	42.7±3.9

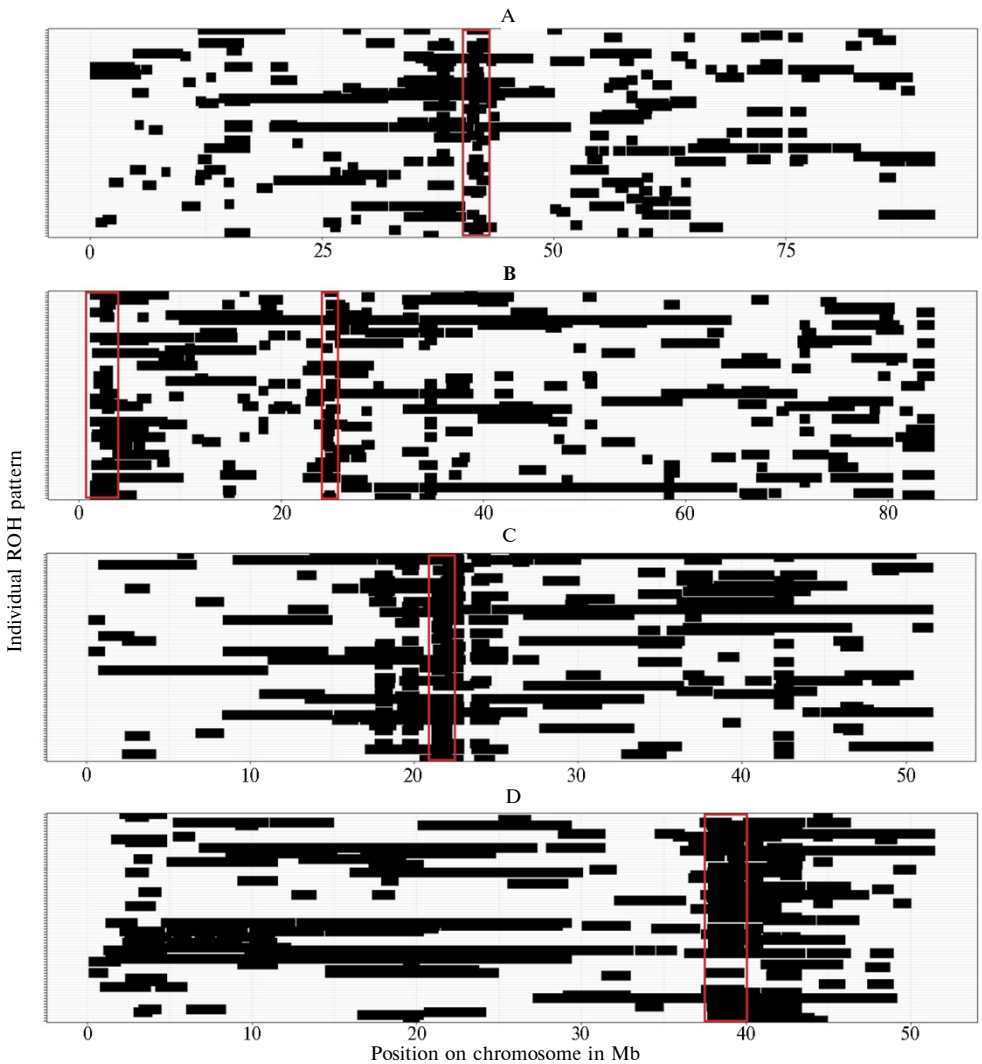

Fig. 2. Identification of ROH patterns for the breeding-valuable loci (indicated by the rectangular frame) **on the bovine chromosomes** (BTA) **12** (A)**, 14** (B)**, 26** (C)**, and 29** (D) of Holsteinized Russian Black Pied cattle (each row represents one animal; n = 76, females of different generations, PJSC "Kamenskoye", Sverdlovsk Province, 2019).

With an increase of the patterns' length, the number of SNPs entering the segment almost proportionally increased, reaching a maximum of 2214 markers. The average ROH length with increasing pattern size from 1-2 Mb to > 16 Mb increased from 1.40 to 22.01 Mb. While the total ROH length (the sum of all segments on average per animal by gradations) did not change significantly, and even tended to decrease from 48.3 Mb to 42.7 Mb.

A search for causal mutations in genes using data visualization (Fig. 2) and an analysis of the ROH patterns occurrence frequency in the studied Ural dairy cattle population (Table 4) made it possible to establish a number of conservative regions in the genome (clusters with protein-encoding genes included) which, we assume, are under the evolutionary pressure.

4. Genomic regions and the most significant protein-coding genes that are under evolutionary pressure and that are identified by ROH (extended homozygous fragments) in the studied group of Holsteinized Russian Black Pied cattle (n = 76, PJSC "Kamenskoye", Sverdlovsk Province, 2019)

BTA	ROH average detected position, Mb		Ratio of animals having ROH in the region, %	The number of genes	Most significant genes
	beginning±σ	end±σ			
12	35.258±1.967	39.747±3.407	44	25	*IL17D, PARP4*
14	1.580±0.454	5.102±2.424	35	66	*ARHGAP39*, *LRRC14*, *PPP1R16A*, *FOXH1*, *CYHR1*, *TONSL*, *CPSF1*, *ADCK5*, *SLC52A2*, *TMEM249, SCRT1*, *DGAT1*, *BOP1*, *MROH1*, *MAF1*, *OPLAH*, *SMPD5*, *SPATC1*, *PLEC*, *GRINA*, *MAPK15*, *EPPK1*, *SCRIB*, *EEF1D*, *RHPN1*, *GPIHBP1*, *GML*, *ADGRB1*, *TSNARE1*, *SLC45A4*, *DENND3*, *PTK2*, *AGO2*, *TRAPPC9*, *KCNK9*
14	24.412±1.828	26.621±2.346	59	18	*XKR4*, *PLAG1*, *CHCHD7*, *SDR16C5-SDR16C6*, *FAM110B*, *SDCBP*, *NSMAF*, *TOX*
26	22.142±1.102	23.877±1.878	73	48	*PKD2L1*, *SCD*, *BTRC*, *ELOVL3*, *GBF1*, *SUFU*, *CNNM2*
29	37.953±1.177	41.982±2.334	62	117	*PAG1-PAG20*, *FADS1-FADS3*, *INCENP*, *SLC3A2*

N o t e. BTA is bovine chromosome.

Due to limitations of the selected experimental population sample, the beginning and end of each of the detected ROH regions is shown as the average value in Mb with a deviation in sigma in order to maximize the accuracy of variation in the homozygous segment. Having analyzed the distribution of ROH across 29 chromosomes in three generations of animals, 5 regions in the genome were identified with a maximum occurrence in at least 30% of individuals on BTA12 (35.3-39.7 Mb), 14 (1.6-5.1 Mb; 24.4-26.6 Mb), 26 (22.1-23.9 Mb) and 29 (38.0-42.0 Mb). The SNP annotation was carried out in order to search for the most significant genes and the QTLs formed by them. It was found that from 35% to 73% of individuals from the total sample had a similar arrangement of ROH patterns with detectable positions.

Annotation of genes that are forming clusters of increased selection pressure was carried out after the visual search and analysis of the ROH occurrence frequency. Among the 25 genes on BTA12, only 2 were classified as the most significant – *IL17D* (interleukin 17D) [39], PARP4 (Poly(ADP-ribose)-polymerase, member of the family 11) [40]. These genes are involved in the immune processes of mammary glands, as well as adipogenesis. In general, on the basis of the functional abstract of the "Cattle QTL Database", it can be summarized that the QTLs are concentrated in this region of the chromosome 12, which are associated with the milkfat and protein yield, total number of somatic cells, duration of a productive life, lactation persistence, live weight, calving difficulties, stillbirths, conception rate, quality of the udder and limbs.

Two regions formed by ROH were identified for BTA14, which are quite

informative in terms of breeding dairy cattle. The first one included 66 genes, of which the most studied ones were the DGAT1 (diacylglycerol acyltransferase 1), the PLEC (pectin) and the GRINA (glutamine receptor), and they are involved in synthesis of lipid components — milkfat and several fatty acids [3]. QTLs were more associated with milk yield per lactation, fat, casein, phosphorus, calcium, milk fatty acids, somatic cells, mastitis, ketosis, resistance to tuberculosis and heat stress in animals. The second region on the chromosome contained 18 genes that were part of the QTL that were responsible for the growth and development parameters, live weight, average daily gain, qualitative and quantitative post-mortem characteristics of carcasses, participation in the synthesis of insulin-like growth factor 1. One of the pre-determining genes in this case is the PLAG1 (gene of pleiomorphic adenoma 1), which is important for growth and reproduction of animals, and also acts as a transcription factor in regulating the expression of insulin-like growth factors [41].

Only 18 genes were detected on BTA26, and the *SCD* (stearoyl-CoA desaturase 1), *PKD2L1, BTRC, ELOVL3, GBF1, SUFU, CNNM2* are being direct precursors for the synthesis of milk fatty acids (linolenic, stearic, myristic and palmitoleic, mono-unsaturated and other acids) [42]. Also, chromosome 26 shows the QTL series for milk yield, content of fat, protein, lactose, linoleic, lauric, myristic and conjugated linoleic fatty acids, milk casein, and qualitative characteristics of the cattle udder.

The discovered genes on BTA29, which are part of the homozygous locus, suggest a predefined effect on animal fertility: a group of glycoproteins associated with pregnancy — the *PAG1 ... PAG20*, as well as the *INCENP*, *SLC3A2* genes, which showed a connection to sperm motility and general fertility index using the QTL database [43, 44]. In turn, a cluster of fatty acid desaturase genes (*FADS1 ... 3*), which is a key enzyme in the metabolism of fatty acids of cow milk, was also identified in the ROH pattern [45]. In general, this QTL region also caused variability of traits, i.e. the interval between calving and heat (estrous), conception rate, semen quality, fertility of bull-daughters and content of protein, casein and fat in milk.

The use of homozygosity patterns to search for the selection “fingerprints” in the animal genome is not limited to this. Using the sum value of the total length of ROH per individual divided by the total length of the autosomal genome (2516398 Mb, in this particular research), it is possible to evaluate the individual autozygosity of the animal. It is assumed that by the number and sum of extended ROH segments, we can judge the level of the so-called genomic inbreeding, which should show higher accuracy compared to the pedigree background.

In this regard, ROH-based inbreeding estimates were calculated depending on their extent in animal generations, the grandmothers, mothers, daughters, including the control group of fathers (Table 5). This approach provided a more accurate assessment of the accumulation of homozygosity in a population by sequential selection of parental pairs of animals. The most accurate assessment of genomic inbreeding was given by the $F_{ROH} > 1$ Mb gradation, which shows the growth in generations, for grandmothers/mothers from 8.6% to 9.8%, or 1.2% ($p < 0.05$), and for mothers/daughters from 9.8% to 12.7%, or 2.9% ($p < 0.001$). $F_{ROH} > 1$ Mb of bull-fathers reached a value of 13.5%. There is an opinion that use of ROH > 1 Mb leads to an overestimation of the genomic inbreeding rate due to difficulties in identifying common ancestors in the pedigree background and because of high recombination frequency; therefore, some authors consider it appropriate to determine F_{ROH} by segments over 4 Mb [18]. Thus, the population of Holsteinized Russian Black Pied cattle, using the example of the highly productive herd of Ural, shows the accumulation of general homozygosity and, as a

result, the formation of a consolidated group of animals according to the genetic characteristics inherent in the Holstein breed. With an increase of ROH patterns size to more than 16 Mb, the observed value of inbreeding in animal generations was lower and amounted to 1.4%/1.4%/2.6% ($p < 0.001$). This suggests that a significant share in the homozygosity assessment is given by ROH with a relatively small extent, which is a reflection of demographic events at the breed level as a whole. The use of a limited number of stud bulls of the current livestock population (genealogical complexes — breeding lines and branches), selected to obtain a high level of milk productivity, has led to an increase in the overall level of homozygosity. Further increase of genomic inbreeding over 15.0% should be controlled in order to timely reduce the possible negative effect on reproductive traits of animals.

5. Comparative estimates of the genomic inbreeding level (F_{ROH}) **of the total length of homozygous fragments** (ROH_{total}, Mb) **in relation to the ROH length, the generation of daughters-mothers-grandmothers and the value of the direct genomic breeding value** (DGV) **in the studied group of Holsteinized Russian Black Pied cattle** ($n = 85$, $M \pm SEM$, PJSC "Kamenskoye", Sverdlovsk Province, 2019)

Parameter	Animal generations (year of birth)			
	mother's mother, $n = 25$ (2012-2014)	mother, $n = 30$ (2015-2017)	daughter, $n = 21$ (2018-2019)	cow's father, $n = 9$ (2008-2015)
Level of genomic inbreeding in the studied population sample (ROH)				
$F_{ROH} > 1$	0.086±0.004	0.098±0.004*	0.127±0.005***	0.135±0.009
$F_{ROH} > 2$	0.068±0.004	0.080±0.004*	0.105±0.005***	0.114±0.009
$F_{ROH} > 4$	0.051±0.004	0.061±0.003*	0.083±0.005***	0.092±0.009
$F_{ROH} > 8$	0.028±0.002	0.034±0.003	0.053±0.004***	0.059±0.009
$F_{ROH} > 16$	0.014±0.002	0.014±0.002	0.026±0.003***	0.032±0.005
$ROH_{total} > 1$	215.3±10.1	247.3±9.3*	319.9±12.7***	338.7±23.1
Genomic breeding value based on milk productivity (DGV)				
Milk yield, kg	448.7±37.8	645.5±39.7**	677.1±57.1	730.3±113.0
Mass fraction of fat, %	0.0005±0.0108	–0.0024±0.0101	0.0049±0.0133	–0.0295±0.0126
Milkfat, kg	18.3±1.4	24.9±1.6**	27.1±2.1	27.3±15
Mass fraction of protein, %	–0.0017±0.0032	–0.0144±0.0042*	–0.0059±0.0041	–0.0165±0.0056
Milk protein, kg	9.6±1.3	13.7±1.2*	15.5±1.9	16.7±3.1

*, **, *** Differences in sequential generations (mother—mother's mother or daughter—mother) are statistically significant at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

A preliminary analysis of the phenotypic traits of cows with their genomic inbreeding indicated a positive correlation between them: for milk yield $r = 0.464$, for mass fraction of fat $r = 0.121$, for milkfat content $r = 0.463$, for milk protein content $r = 0.389$, for conception rate $r = 0.168$, with the exception of mass fraction of protein $r = –0{,}400$. So, with an increase in the overall level of homozygosity (ROH), a simultaneous improvement in the parameters of milk yield, yield of milk fat and protein was observed, with a moderate decrease in fertility rates.

Based on the Russian reference group of dairy cattle, the DGV forecast was calculated for three generations of animals, including bull-fathers (see Table 5). It was shown that mothers significantly exceeded the generation of grandmothers for genomic breeding value in milk yield by +157 kg of milk, +6.6 kg of milkfat and +4.1 kg of milk protein, ceding by -0.003% to mass fraction of fat (unreliable) and by -0.013% to mass fraction of protein. The daughters' generation had positive growth dynamics of DGV, compared to the mothers by +32 kg of milk, +2.2 kg of fat, +1.8 kg of protein, +0.007% fat and +0.008% protein. The bull-fathers of the studied offspring showed the highest level of genomic breeding value, but like their daughters, they did not have significant superiority over the last generation of animals in the experimental sample. This may indicate an exit to the breeding

plateau in the herd through saturation of the offspring genotype with a combination of valuable alleles for traits of milk productivity obtained as a result of selection of the parental pairs of animals. Further improvement in the genetic value of individuals is possible through the use of new generations of bulls with higher DGV values, or with high accuracy and level of breeding value in the offspring quality.

Hence, it was shown that the control of monogenic genetic mutations (defects) and selection-determined DNA markers can provide efficient selection of animals at the herd level (or population as a whole) and they can be used in the selection of parental pairs for producing offspring with desired parameters of breeding qualities. Analysis of genome-wide associations with traits of milk productivity and fertility in cattle made it possible to establish causal mutations that are localized close to the genes or within the genes, of which *ZNF16*, *ARHGAP39*, *EFR3A*, *KCNQ3*, *KHDRBS3*, *IGF1*, *SERPINA5*, *ARHGAP31*, *CHD1* and *STXBP5L* are the most significant ones. Value of significant associations for point mutations was within $p < 2{,}3\times10^{-5}$ to $1{,}1\times10^{-7}$. The identified genetic polymorphisms are consistent with studies conducted in Russia on the example of Holsteinized Russian Black Pied cattle, as well as with the data on the Holstein breed in North America and several European countries. According to the ROH analysis, the regions in the genome that are subject to the highest evolutionary pressure on chromosomes 12, 14, 26, and 29, were identified, which showed the presence of QTLs that are associated with traits of milk productivity, animal fertility, fatty acid composition of cow milk, cattle body type, growth and development parameters. Experimental data were obtained on the basis of homozygosity patterns in order to clarify the level of livestock inbreeding, as an example for use in animal selection programs at the level of individual herds and the population as a whole. Genomic inbreeding showed positive dynamics of population increase, while at the same time there was an increase in the forecast value of their genomic breeding value. This shows that the complex of genomic analysis tools makes it possible to be more efficient about planning of obtaining animals with desired breeding value that would meet the economic needs of dairy cattle breeding. The replacement of the reference group of dairy cattle in Russia through the integration of regional animal populations, primarily the genotyped cows and the stud bulls, provides prospects for expanded use of prognosis of genomic breeding value for reproduction of both bull and pedigree livestock of Holsteinized Russian Black Pied and Holstein cattle.

REFERENCES

1. Zinov'eva N., Strekozov N., Yanchukov I., Ermilov A., Eskin G. *Zhivotnovodstvo Rossii*, 2015, 3: 27-30 (in Russ.).
2. Kudinov A.A., Maslennikova E.S., Plemyashov K.V. *Zootekhniya*, 2019, 1: 2-6 (in Russ.).
3. Sermyagin A.A., Gladyr E.A., Plemyashov K.V., Kudinov A.A., Dotsev A.V., Deniskova T.E., Zinovieva N.A. 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»*. K. Anisimov et al. (eds.). Springer, Cham, 2018: 591-599 (doi: 10.1007/978-3-319-62870-7_62).
4. Qanbari S., Pimentel E.C.G., Tetens J., Thaller G., Lichtner P., Sharifi A.R., Simianer H. A genome-wide scan for signatures of recent selection in Holstein cattle. *Animal Genetics*, 2010, 41(4): 377-389 (doi: 10.1111/j.1365-2052.2009.02016.x).
5. Wang X., Wurmser C., Pausch H., Jung S., Reinhardt F., Tetens J., Thaller G., Fries R. Identification and dissection of four major QTL affecting milk fat content in the German Holstein-Friesian population. *PLoS ONE*, 2012, 7(7): e40711 (doi: 10.1371/journal.pone.0040711).
6. Maurice-Van Eijndhoven M.H.T., Bovenhuis H., Veerkamp R.F., Calus M.P.L. Overlap in genomic variation associated with milk fat composition in Holstein Friesian and Dutch native dual-purpose breeds. *Journal of Dairy Science*, 2015, 98(9): 6510-6521 (doi: 10.3168/jds.2014-9196).

7. Chamberlain A.J., Hayes B.J., Savin K., Bolormaa S., McPartlan H.C., Bowman P.J., Van Der Jagt C., MacEachern S., Goddard M.E. Validation of single nucleotide polymorphisms associated with milk production traits in dairy cattle. *Journal of Dairy Science*, 2012, 95(2): 864-875 (doi: 10.3168/jds.2010-3786).
8. Fang M., Fu W., Jiang D., Zhang Q., Sun D., Ding X., Liu J. A multiple-SNP approach for genome-wide association study of milk production traits in Chinese Holstein cattle. *PLoS ONE*, 2014, 9(8): e99544 (doi: 10.1371/journal.pone.0099544).
9. Xiang R., MacLeod I.M., Bolormaa S., Goddard M.E. Genome-wide comparative analyses of correlated and uncorrelated phenotypes identify major pleiotropic variants in dairy cattle. *Scientific Reports*, 2017, 7: 9248 (doi: 10.1038/s41598-017-09788-9).
10. Zhang F., Wang Y., Mukiibi R., Chen L,, Vinsky M., Plastow G., Basarab J., Stothard P., Li G. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. *BMC Genomics*, 2020, 21: 36 (doi: 10.1186/s12864-019-6362-1)
11. Reyer H., Shirali M., Ponsuksili S., Murani E., Varley P.F., Jensen J., Wimmers K. Exploring the genetics of feed efficiency and feeding behaviour traits in a pig line highly selected for performance characteristics. *Molecular Genetics and Genomics: MGG*, 2017, 292(5): 1001-1011 (doi: 10.1007/s00438-017-1325-1).
12. Mel'nikova E.E., Bardukov N.V., Fornara M.S., Kostyunina O.V., Sermyagin A.A., Brem G., Zinov'eva N.A. The study of effect of genotypes for DNA marker on reproductive qualities of sows of Large White and Landrace breeds. *Sel'skokhozyaistvennaya biologiya* [*Agricultural Biology*], 2019, 54(2): 227-238 (doi: 10.15389/agrobiology.2019.2.227eng).
13. Belous A.A., Sermyagin A.A., Kostyunina O.V., Brem G., Zinov'eva N.A. Study of genetic architecture of feed conversion rate in duroc young boars (*Sus scrofa*) based on the genome-wide SNP analysis. *Sel'skokhozyaistvennaya biologiya* [*Agricultural Biology*], 2019, 54(4): 705-712 (doi: 10.15389/agrobiology.2019.4.705rus).
14. Pryce J.E., Johnston J., Hayes B.J., Sahana G., Weigel K.A., McParland S., Spurlock D., Krattenmacher N., Spelman R.J., Wall E., Calus M.P.L. Imputation of genotypes from low density (50,000 markers) to high density (700,000 markers) of cows from research herds in Europe, North America, and Australasia using 2 reference populations. *Journal of Dairy Science*, 2014, 97(3): 1799-1811 (doi: 10.3168/jds.2013-7368).
15. Ferenčaković M., Sölkner J., Curik I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution*, 2013, 45: 42 (doi: 10.1186/1297-9686-45-42).
16. Ignat'eva L.P., Belous A.A., Nedashkovskii I.S., Kostyunina O.V., Sermyagin A.A., Zinov'eva N.A. *Sibirskii vestnik sel'skokhozyaistvennoi nauki*, 2019, 49(6): 79-87 (doi: 10.26898/0370-8799-2019-6-9) (in Russ.).
17. Bakoev C.Yu., Getmantseva L.V. *Dostizheniya nauki i tekhniki APK*, 2019, 33(11): 63-68 (doi: 10.24411/0235-2451-2019-11114) (in Russ.).
18. Smaragdov M.G., Kudinov A.A. *Dostizheniya nauki i tekhniki APK*, 2019, 33(6): 51-53 (doi: 10.24411/0235-2451-2019-10612) (in Russ.).
19. Sermyagin A.A., Zinov'eva N.A. *Dostizheniya nauki i tekhniki APK*, 2019, 33(12): 77-82 (doi: 10.24411/0235-2451-2019-11216) (in Russ.).
20. Chang C.C., Chow C.C., Tellier L.CAM., Vattikuti S., Purcell S.M., Lee J.J. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 2015, 4: 7 (doi: 10.1186/s13742-015-0047-8).
21. Zhang L., Orloff M.S., Reber S., Li S., Zhao Y., Eng C. aTOH: Extended Approach for Identifying Tracts of Homozygosity. *PLoS ONE*, 2013, 8(3): e57772 (doi: doi:10.1371/journal.pone.0057772).
22. Hu Z.-L., Park C.A., Reecy J.M. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Research*, 2019, 47: D701-710 (doi: 10.1093/nar/gky1084).
23. VanRaden P.M. Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 2008, 91(11): 4414-4423 (doi: 10.3168/jds.2007-0980).
24. Jiang J., Ma L., Prakapenka D., VanRaden P.M., Cole J.B., Da Y. A large-scale genome-wide association study in U.S. Holstein cattle. *Front. Genet.*, 2019, 10: 412 (doi: 10.3389/fgene.2019.00412).
25. Sun J., Shan L., Zhang C., Chen H. Haplotype combination of the bovine *PCSK1* gene sequence variants and association with growth traits in Jiaxian cattle. *Journal of Genetics*, 2014, 93(3): e123-e129. Rezhim dostupa: http://www.ias.ac.in/jgenet/OnlineResources/93/e123.pdf. Data obrashcheniya: 27.01.2020 goda.
26. Friedrich J., Brand B., Ponsuksili S., Graunke K.L., Langbein J., Knaust J., Kühn C., Schwerin M. Detection of genetic variants affecting cattle behaviour and their impact on milk production: a genome-wide association study. *Anim. Genet.*, 2016, 47(1): 12-18 (doi: 10.1111/age.12371).
27. Karisa B.K., Thomson J., Wang Z., Stothard P., Moore S.S., Plastow G.S. Candidate genes and single nucleotide polymorphisms associated with variation in residual feed intake in beef cattle.

Journal of Animal Science, 2013, 91(8): 3502-3513 (doi: 10.2527/jas.2012-6170).
28. Liu Y., Zan L., Xin Y., Tian W., Li L., Wang H. *ZBTB38* gene polymorphism associated with body measurement traits in native Chinese cattle breeds. *Gene*, 2013, 513(2): 272-277 (doi: 10.1016/j.gene.2012.10.026).
29. Valente T.S., Baldi F., Sant'Anna A.C., Albuquerque L.G., Paranhos da Costa M.J. Genome-wide association study between single nucleotide polymorphisms and flight speed in Nellore cattle. *PLoS ONE*, 2016, 11(6): e0156956 (doi: 10.1371/journal.pone.0156956).
30. Lima A.O.D., Oliveira P.S.N., Tizioto P.C., Somavilla A.L., Diniz W.J.S., Silva J.V.D., Andrade S.C.S., Boschiero C., Cesar A.S.M., Souza M.M., Rocha M.I.P., Afonso J., Buss C.E., Mudadu M.A., Mourao G.B., Coutinho L.L., Regitano L.C.A. *PRUNE2* gene has a potential effect on residual feed intake in Nellore cattle. *Journal of Animal Science*, 2016, 94(E-Suppl. 5): 152-153 (doi: 10.2527/jam2016-0318).
31. Cole J.B., Wiggans G.R., Ma L., Sonstegard T.S., Lawlor T.J., Crooker B.A., Van Tassell C.P., Yang J., Wang S., Matukumalli L.K., Da Y. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics*, 2011, 12: 408 (doi: 10.1186/1471-2164-12-408).
32. Szewczuk M. Polymorphism in exon 2 encoding the putative ligand-binding pocket of the bovine insulin-like growth factor 1 receptor affects milk traits in four different cattle breeds. *Journal of Animal Breeding and Genetics*, 2017, 134(1): 34-42 (doi: 10.1111/jbg.12216).
33. Leal-Gutiérrez J.D., Elzo M.A., Johnson D.D., Hamblen H., Mateescu R.G. Genome wide association and gene enrichment analysis reveal membrane anchoring and structural proteins associated with meat quality in beef. *BMC Genomics*, 2019, 20(1): 151 (doi: 10.1186/s12864-019-5518-3).
34. Hayashi K., Ushizawa K., Hosoe M., Takahashi T. Differential gene expression of serine protease inhibitors in bovine ovarian follicle: possible involvement in follicular growth and atresia. *Reproductive Biology and Endocrinology*, 2011, 9: 72 (doi: 10.1186/1477-7827-9-72).
35. Zhuang Z., Xu L., Yang J., Gao H., Zhang L., Gao X., Li J., Zhu B. Weighted single step genome-wide association study for growth traits in Chinese Simmental beef cattle. *Genes*, 2020, 11: 189 (doi: 10.3390/genes11020189).
36. Zhang K., Rajput S.K., Wang S., Folger J.K., Knott J.G., Smith G.W. *CHD1* regulates deposition of histone variant H3.3 during bovine early embryonic development. *Biology of Reproduction*, 2016, 94(6): 140, 1-8 (doi: 10.1095/biolreprod.116.138693).
37. Pryce J.E., Hayes B.J., Bolormaa S., Goddard M.E. Polymorphic regions affecting human height also control stature in cattle. *Genetics*, 2011, 187(3): 981-984 (doi: 10.1534/genetics.110.123943).
38. Twomey A.J., Berry D.P., Evans R.D., Doherty M.L., Graham D.A., Purfield D.C. Genome-wide association study of endo-parasite phenotypes using imputed whole-genome sequence data in dairy and beef cattle. *Genetics Selection Evolution*, 2019, 51: 15 (doi: 10.1186/s12711-019-0457-7).
39. Wang X.G., Ju Z.H., Hou M.H., Jiang Q., Yang C.H., Zhang Y., Sun Y., Li R.L., Wang C.F., Zhong J.F., Huang J.M. Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus* mastitis. *PLoS ONE*, 2016, 11(12): e0167666 (doi: 10.1371/journal.pone.0167666).
40. Edea Z., Jeoung J.H., Shin S.S., Ku J., Seo S., Kim I.H., Kim S.W., Kim K.S. Genome-wide association study of carcass weight in commercial Hanwoo cattle. *Asian-Australasian Journal of Animal Sciences (AJAS)*, 2018, 31(3): 327-334 (doi: 10.5713/ajas.17.0276).
41. Utsunomiya Y.T., Milanesi M., Utsunomiya A.T.H., Torrecilha R.B.P., Kim E.S., Costa M.S., Aguiar T.S., Schroeder S., do Carmo A.S., Carvalheiro R., Neves H.H.R., Padula R.C.M., Sussai T.S., Zavarez L.B., Cipriano R.S., Caminhas M.M.T., Hambrecht G., Colli L., Eufemi E., Ajmone-Marsan P., Cesana D., Sannazaro M., Buora M., Morgante M., Liu G., Bickhart D., Van Tassell C.P., Sölkner J., Sonstegard T.S., Garcia J.F. A *PLAG1* mutation contributed to stature recovery in modern cattle. *Scientific Reports*, 2017, 7(1): 17140 (doi: 10.1038/s41598-017-17127-1).
42. Li C., Sun D., Zhang S., Liu L., Alim M.A., Zhang Q. A post-GWAS confirming the *SCD* gene associated with milk medium- and long-chain unsaturated fatty acids in Chinese Holstein population. *Animal Genetics*, 2016, 47(4): 483-490 (doi: 10.1111/age.12432).
43. Hering D.M., Olenski K., Kaminski S. Genome-wide association study for poor sperm motility in Holstein-Friesian bulls. *Animal Reproduction Science*, 2014, 146(3-4): 89-97 (doi: 10.1016/j.anireprosci.2014.01.012).
44. Santos D.J.A., Cole J.B., Null D.J., Byrem T.M., Ma L. Genetic and nongenetic profiling of milk pregnancy-associated glycoproteins in Holstein cattle. *Journal of Dairy Science*, 2018, 101(11): 9987-10000 (doi: 10.3168/jds.2018-14682).
45. Ibeagha-Awemu E.M., Akwanji K.A., Beaudoin F., Zhao X. Associations between variants of *FADS* genes and omega-3 and omega-6 milk fatty acids of Canadian Holstein cows. *BMC Genetics*, 2014, 15: 25 (doi: 10.1186/1471-2156-15-25).