MULTI-LOCUS GENOTYPING OF CATTLE GENOMES ON THE BASES OF THE REGION HOMOLOGY TO RETROTRANSPOSONS

V.I. GLAZKO[1, 2], G.Yu. KOSOVSKII[2], S.N. KOVAL'CHUK[2], T.T. GLAZKO[1, 2]

[1]K.A. Timiryazev Russian State Agrarian University—Moscow Agrarian Academy, 49, ul. Timiryazevskaya, Moscow, 127550 Russia;
[2]Center for Experimental Embryology and Reproductive Biotechnology, Federal Agency of Scientific Organizations, 12/4, ul. Kostyakova, Moscow, 127422 Russia, e-mail vglazko@yahoo.com, gkosovsky@mail.ru, s.n.kovalchuk@mail.ru, tglazko@rambler.ru

A b s t r a c t

Intensification in livestock has led to necessity of developing methods which allows to estimate animal genomic breeding values (GBV). In this, an approach was based on identifying association of animal production traits' variability to polymorphism in different genome elements, from microsatellite loci to millions of single nucleotide polymorphisms (SNPs), in genome-wide sequencing. Despite numerous accumulated data, the development of relatively simple, fast and affordable to interpretation methods of polyloci genotyping (genome scan) for assessing breed «gene pool standard» or genomic selection is still actual. The mobile genetic elements, being high polymorphic genomic sites, can serve as anchors in a multi-locus genotyping. The species-specific retrotransposons such as L1_BT LINE which present in the bovine genome with high frequency and the endogenous retrovirus ENV1_BT are of particular interest. In this regard, the aim of this work was the comparative analysis of polymorphism in genomic DNA fragments flanked by inverted repeats of nucleotide sequences homologous to fragments of these retrotransposons in genomes of specialized dairy Ayrshire and Black and White Holstein breeds and the Kalmyk beef breed. A recombination between L1_BT LINE and ENV1_BT we reported earlier turned out to be conservative in studied breeds. This apparently indicated it to be old in appearance. Since the genetic relationships identified from distribution patterns and length polymorphisms of DNA flanked by ENV1_BT inverted repeats differed significantly between specialized dairy breeds and local meat breed but were low variable within the breeds, this allowed to conclude about its relatively low transposition activity. The PIC values (polymorphic information content) for spectra produced in PCR with the endogenous retrovirus ERV1_BT fragment as a primer ranged from 0.075 to 0.089 for studied breeds. Distribution of inverted repeats homologous to L1_BT was different in character and varied among breeds and within breeds, the PIC values were substantially higher and ranged from 0.062 to 0.260. Thus, the use of multi-locus genotyping of transposable elements is effective for the purposes of animal genetics and breeding but their application depends on transposition activity. Our data showed that the ENV1_BT inverted repeats seem to be convenient marker in genetic differentiation among breeds and revealing «gene pool standard», whereas more transposable L1_BT elements are more relevant in indicating individual variability within a breed.

Keywords: multi-locus genotyping, retrotransposons, inverted repeats, genomic scanning, breed, cattle.

In recent years, a possible inclusion of genomic selection in traditional cattle breeding programs attracts the increasing attention [1-4]. To identify the genomic targets of artificial selection associated with desirable manifestation of productivity traits in beef and dairy cattle breeds, the genomic maps of single number polymorphisms (SNP) including 705243 such sites distributed over all autosomes have been created [5-7]. However, these associations appeared to be breed specific and characterized by interbreed variability.

To increase the efficiency of marking genomic domains the polymorphism of which could be used to solve problems of genomic selection, a new generation of markers resulting from the study on polymorphisms based on the short (less than 400 base pairs, bp) DNA fragments copy number variability

(CNV) and their genomic distribution [8-10] is involved in SNPs. Nevertheless, these two approaches (SNP and CNV) to bovine genome multi-locus genotyping turned out not to provide unequivocal results. Thus, the chromosomal domains of quantitative trait loci (QTL) have been identified in Holstein breed using the estimates of microsatellite loci polymorphism and SNPs. Of 82 such domains in which CNVs were detected, 17 only overlapped with SNPs. In 20 chromosome domains of QTL localization, CNVs, but not SNPs were present. Based on their results, the authors conclude on the need of the joint use of SNPs and CNVs to include genomic selection in traditional breeding methods.

At the same time, such multi-locus genotyping (genomic scan) is an expensive and quite complicated procedure, and its results are difficult to interpret. Attention should be drawn to the fact that, as a rule, the structural and functional features of CNV and SNP localization sites and their potential predisposition to polymorphism remain unknown. In this regard, genomic elements with high polymorphism which include microsatellite loci are of particular interest. Based on the estimates of the polymorphism of the microsatellites which form inverted repeats on the short segments of genomic DNA, the ISSR-PCR (inter-simple sequence repeats) method for multi-locus genomic genotyping has been developed [11]. Due to the relative cheapness of obtaining amplification product multi-locus spectra and the simplicity of their interpretation, this generation of markers has been widely used and is used in a large number of plant and animal species in the solution of population genetics problems.

However, this type of markers has several disadvantages as well (in addition to the dominant character of manifestation). They are mainly related to the «anonymity» of DNA sites flanked by an inverted microsatellite repeat. This makes it not possible to distinguish between such fragments, if they have the same or close in the nucleotide number, but are located in different genome regions. Obviously, such a source of potential errors can significantly distort the interpretation of results. To overcome this «anonymity», we have performed sequencing of the fragments of genomic DNA flanked by inverted repeat of the $(AGC)_6G$ sequence. As a result, we obtained the evidence that the sites homologous to mobile genetic elements such as L1_2_BT, BovA, ERV1_2_1_BT, LTR Gypsy, and for the products of their recombination predominate in these fragments [12].

After the complete sequencing of bovine genome in 2009 [13], special attention was drawn to AGC microsatellite, as a few facts have been found: firstly, it is found in bovine genome 90 and 142 times more often than in the genomes of humans and dogs, respectively; secondly, 39 % of AGC microsatellites in bovine genome are associated with retroposon Bov-A2 SINE localization. This element has AGC in its tail portion and is an evolutionarily more advanced derivative of an ancient Bov-B (LINE) repeat in cattle. It should be noted that 65 taxa have sequences homologous to this long dispersed nuclear element (Bov-B) that are presumably involved in the horizontal transfer of genetic material between taxa [14].

There are a lot of published experimental data showing that the distribution of microsatellites (and of their inverted repeats in particular) is closely related to the transpositions of retrotransposons [7-10]. The areas of localization of retrotransposons such as LINE were also found to be frequently involved in recombination and polymorphisms associated for CNV markers [15-18]. LINE RTE (BovB) and BovB SINE related repeats are most common among dispersed repeats in bovine genome. L1 LINEs are the next ones in their frequency. L1_BTs are the most common representatives of this subfamily. L1_BTs are the younger elements to BovB; L1_BTs actively participate in the genomic reorganization. In bovine genome, 811 L1 copies have been identified, 73 of which are

active. A relatively reduced nucleotide substitution rate has been detected in these copies compared to BovB (0.031 ± 0.001 vs 0.065 ± 0.002, respectively). Statistically significant correlations between L1 localization and the areas of copy number variability (CNV markers) have been described [19, 20]. The accumulated data allow us to think that the use of L1_BT sequences as primers in PCR for multi-locus genotyping may become the basis for a new generation of genome scanning methods in cattle as it potentially combines SNP and CNV marker associations.

To test this hypothesis, in this study we first compared amplification products in three bovine breeds, i.e. two highly specialized dairy (Ayrshire and Black and White Holstein) breeds and a local (Kalmyk) breed, using as primers the L1_BT sequences and also ERV1_2_1_BT with which, based on the results of our sequencing [12], L1_BT forms a recombinant product. These data suggest that the distribution of inverted repeats of endogenous retrovirus fragment ENV1_BT in bovine genomes differentiates dairy and beef cattle, while L1_BT differentiates individual animals within breeds.

The purpose of this study was to estimate the possibility of multi-locus genotyping in different bovine breeds by the regions homologous to retrotransposons to identify the inter- and crossbreed differentiation.

Technique. Genomic scanning was performed in 45 blood samples from three cattle breeds: Kalmyk (15 animals, Republic of Kalmykia), Ayrshire (Moscow stud № 1, Moscow, 15 animals) breeds and Black and White Holstein breeds («Mozhaiskii» farm, Moscow Province, 5 animals; «Kuibyshevo» enterprise, Moscow Province, 10 animals).

Total DNA was isolated from whole blood samples using BioSilica kit (LLC Biosilika, Russia). PCR primers were the fragments of L1-2_BT retransposon and ERV1-2C-LTR bovine endogenous retrovirus. Primers were estimated using GeneRunner (Hastings Software, Inc., Hastings, USA; http://www.generunner.net/) and PerlPrimer (http://perlprimer.sourcefo-rge.net/) software on the basis of L1-2_BT and ERV1-2C-LTR sequences which we have determined earlier [12]. PCR was performed using an Amplitronyx A6 amplificator (Nyx Technik Inc., USA). PCR was performed with HS Taq DNA polymerase (Evrogen, Russia) according to manufacturer's protocol. From the stock 2 μM solution of one of the primers an aliquot of 2.5 μl (final primer concentration of 0.5 μM) were added to the reaction mixture. The amount of template DNA in the reaction mixture was 4 ng. Amplification regime was as follows: 2 min at 95 °C; 40 cycles of 15 sec at 94 °C, 15 sec at 56 °C, 2 min at 72 °C; 2 min at 72 °C.

The amplification products were separated in agarose gel (1.5 %), stained with ethidium bromide (0.5 μg/ml) for 80 min at 120 V (power supply Voltronyx Reactor-330, Nyx Technik Inc., USA). Visualization of gel electrophoresis results was performed using ECX-F20.M transilluminator and Doc-Print DP-C-P-20.M gel documentation systems (Vilber Lourmat, Germany). Fragment length was estimated using the O'GeneRuler DNA Ladder mix (100-10000 bp) (Thermo Fisher Scientific, USA) DNA marker. Based on the presence or absence of a certain length DNA fragments, binary matrices were constructed.

Mathematical data processing was performed using the Dendro UPGMA service [21] with the similarity coefficient according to Dice (L.R. Dice), identical genetic distances were estimated according to M. Nei and W.-H. Li [22]. The Newick tree format was used to construct dendrograms with the TreeView software [23]. Statistical processing was performed using the GenAlEx v. 6.5 software [24].

Results. L1-2_BT retransposon (referred to as L1dir1 and Rec-rev2) and ERV1-2C-LTR bovine endogenous retrovirus (referred to as Rec-dir2) frag-

ments were used as primers. As a result of sequencing [12], L1_BT produces a recombinant product with ERV1_2_1_BT. Primers were selected for the product of L1_BT and ERV1_BT recombination so that the distance between them was at least 100 bp.

1. Retransposon fragments used as primers in bovine genome multi-locus genotyping PCR

Primer	Retrotransposon	Primer nucleotide sequence
L1dir1	L1-2_BT retrotransposon	ACTCCCAGCAGAGAAGTAGC
Rec-rev2	L1-2_BT retrotransposon	CATGTACACACTGCTATATTT
Rec-dir2	ERV1-2C-LTR	TCCTCACTCCGGTCCTC

To start with, we performed a multi-locus genotyping in the breeds included in the study, using as primers the regions homologous to L1_BT and EVR1_BT sequences involved in recombination and identified in the relevant product under sequencing genomic DNA fragments flanked by inverted microsatellite $(AGC)_6G$ repeats [12]. It was found that the spectra obtained were conservative and it was not possible to identify neither inter- nor crossbreed differences with their help. We can expect that this recombination product was formed long ago, in the early distribution for the ancestral L1 and EVR1 genome, and is protected from aberrations. Protection against reorganization was described in the domains where the ancient retrotransposons are localized, preserving a high degree of similarity between the cattle, humans and mice. These include, in particular, LINE RTE/SINE ART2A and LINE L2/SINE MIR, some DNA transposons and tandem repeats [17, 20, 25-27].
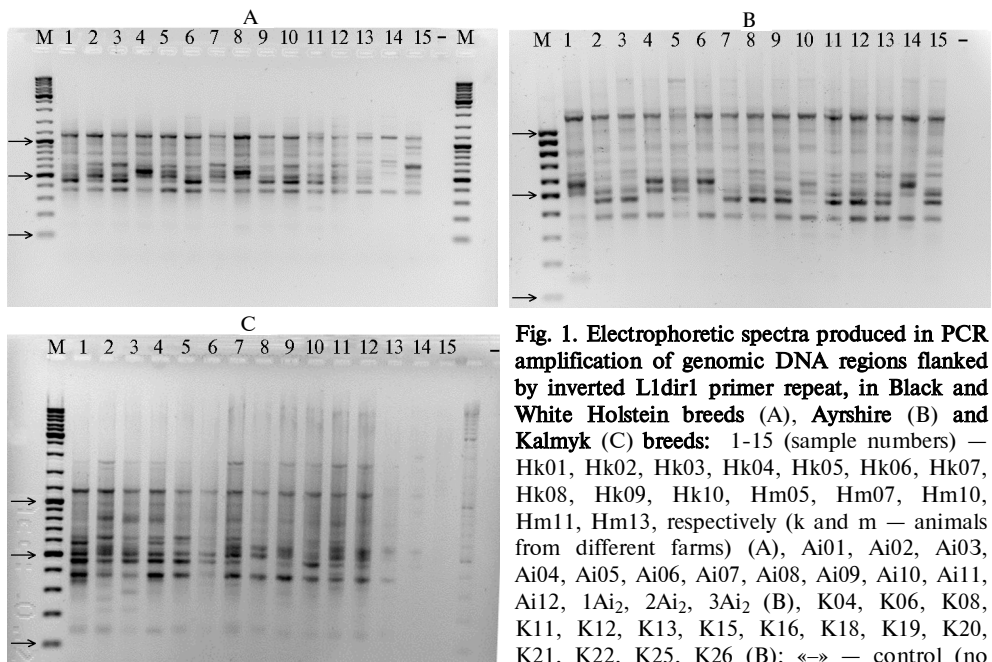

Fig. 1. Electrophoretic spectra produced in PCR amplification of genomic DNA regions flanked by inverted L1dir1 primer repeat, in Black and White Holstein breeds (A), Ayrshire (B) and Kalmyk (C) breeds: 1-15 (sample numbers) — Hk01, Hk02, Hk03, Hk04, Hk05, Hk06, Hk07, Hk08, Hk09, Hk10, Hm05, Hm07, Hm10, Hm11, Hm13, respectively (k and m — animals from different farms) (A), Ai01, Ai02, Ai03, Ai04, Ai05, Ai06, Ai07, Ai08, Ai09, Ai10, Ai11, Ai12, $1Ai_2$, $2Ai_2$, $3Ai_2$ (B), K04, K06, K08, K11, K12, K13, K15, K16, K18, K19, K20, K21, K22, K25, K26 (B); «–» — control (no genomic DNA), M — molecular weight marker O'GeneRuler DNA Ladder mix (100-10000 bp) (Thermo Fisher Scientific, USA). Arrows mark the fragments of 1000, 500, and 100 bp. Characteristics of amplification products spectra in the species studied are provided in Table 2, for primer description please refer to Table 1.

Due to the conservative fragments of DNA in which the product of L1-BT and EVR1_BT recombination is localized, we compared the spectra of genomic DNA regions flanked by inverted repeats of sequencing of each recombinant product member. In multi-locus genotyping with L1dir1 primer in PCR,

the spectra of amplification products of DNA fragments flanked by inverted repeat of selected L1dir1 sequences were polymorphic. Each DNA fragment was considered as a single locus and the analysis included only those fragments that were reproduced on DNA template in 3 replications.

Figure 1 shows the resulting spectra of amplification products. It turned out that the main interbreed polymorphism was observed among the fragments from 450 to 650 bp in length, in all three breeds.

Based on PIC (polymorphic information content) values, the greatest interbereed variability for amplicon spectra was characteristic of Ayrshire breed, the least variability was observed in Black and White Holstein breeds (Table 2). To test this hypothesis, we calculated genetic distances for these spectra for individual animals, which served as the basis for the dendrogram demonstrating their genetic relationship (Fig. 2).

2. The main characteristics of spectra produced in PCR amplification of DNA fragments flanked by inverted L1dir1 primer repeat, in three breeds studied

Parameter	Ayrshire breed ($n = 15$)	Kalmyk breed ($n = 15$)	Black and White Holstein breeds ($n = 15$)
L1dir1 p r i m e r			
Total number of amplicons	6	9	7
Number of unique amplicons	0	1	1
PIC	0.260	0.193	0.189
Rec-rev2 p r i m e r			
Total number of amplicons	6	10	7
Number of unique amplicons	1	2	0
PIC	0.159	0.163	0.062
Rec-dir2 p r i m e r			
Total number of amplicons	8	11	7
Number of unique amplicons	1	4	0
PIC	0.089	0.075	0.077

N o t e. PIC — polymorphic information content. For primer description please refer to Table 1.
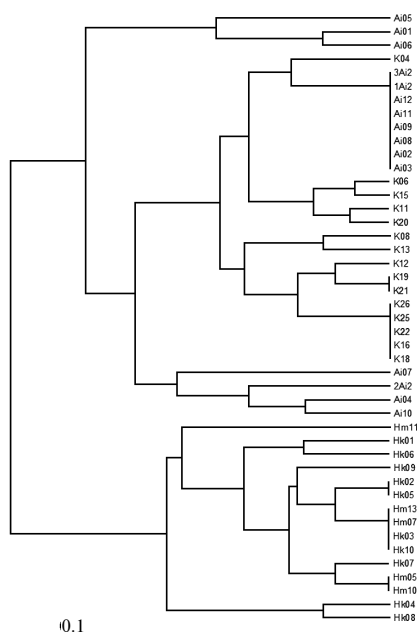

Fig. 2. Dendrogram of genetic relationship between Ayrshire, Kalmyk breeds and Black and White Holstein breeds (marked as A, K and N) for the spectra produced in PCR amplification with L1dir1 primer. Constructed based on calculation of genetic distances according to M. Nei and W.-H. Li (1979). Scale bar represents a percentage of the genetic distance unit. For the primer description please refer to Table 1.

Two major clusters are distinguished in the dendrogram (Fig. 2). One of them combines all the Black and White Holstein breeds representatives, the other one combines the groups of Ayrshire and Kalmyk breeds, which, in turn, are divided into small subclusters, while the representatives of different breeds are not intermixed in subclusters (except in one case for the animal referred to as K04). The findings suggest pronounced distribution heterogeneity for the regions homologous to direct and inverted L1dir1 sequence over short distances (450-650 bp) in Ayrshire and Kalmyk breeds compared to Holstein cattle.

When the L1_2_BT alternative DNA chain region was used as primer (Rec-rev2 primer), PCR spectra had a certain similarity to the previous ones: basic polymorphism was identified among the 450-700 bp amplicons (Fig. 3),

but polymorphic information content (PIC) was lower compared to the spectra with L1dir1 primer (Table 2), and, as in the previous case, the minimum PIC value was found in Black and White Holstein breeds.
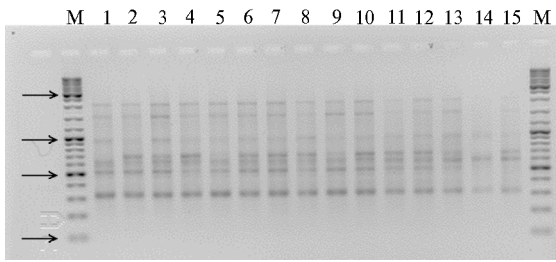

Fig. 3. Electrophoretic spectra produced in PCR amplification of genomic DNA regions flanked by Rec-rev2 primer inverted repeat, in Black and White Holstein breeds: 1-15 (sample numbers) — Hk01, Hk02, Hk03, Hk04, Hk05, Hk06, Hk07, Hk08, Hk09, Hk10, Hm05, Hm07, Hm10, Hm11, Hm13, respectively (k and m — animals from different farms); M — molecular weight marker O'GeneRuler DNA Ladder mix (100-10000 bp) (Thermo Fisher Scientific, USA). Arrows mark the fragments of 3000, 1000, 500, and 100 bp. Characteristics of amplification products spectra are provided in Table 2, for primer description please refer to Table 1.

Despite the overall similarity of the spectra, the dendrogram constructed based on the estimation of genetic distances between the individual animals for the Rec-rev2 primer spectra (Fig. 4) differed from the dendrogram for the L1dir1 spectrum (Fig. 2) so that the total cluster included Ayrshire breed animals, Black and White Holstein breeds appeared to be divided into three small subclusters, and subdivision was the greatest in Kalmyk breed (Fig. 4). In both cases (L1dir1 and Rec-rev2 primers, see Fig. 2, Fig. 4), maximum internal subdivision in dendrograms was found in Kalmyk breed.
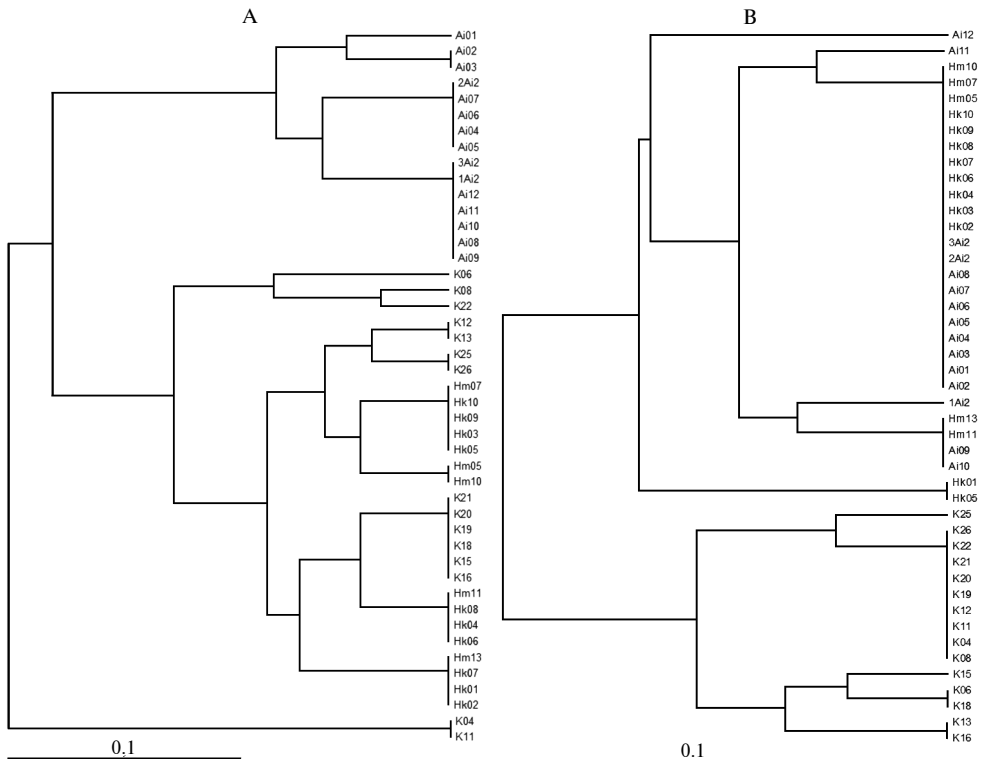

Fig. 4. Dendrogram of genetic relationship between Ayrshire, Kalmyk breeds and Black and White Holstein breeds (marked as A, K and N) for the spectra produced in PCR amplification with Rec-rev2 (A) and Rec-dir2 (B) primers. Constructed based on calculation of genetic distances according to M. Nei and W.-H. Li (1979). Scale bar represents a percentage of the genetic distance unit. For primer description please refer to Table 1.

Our results are consistent with the data of high L1_BT transposition ac-

tivity [20]. The closeness of inverted repeats homologous to L1_BT in the forward and alternative DNA strands as well as the localization variability, as evidenced by the polymorphism of amplification products detected using L1dir1 and Rec-rev2 primers in the regions of 450-700 bp DNA fragments, appears to reflect the presence of preferred transpositions sites in bovine genome.

The next spectrum of amplification products was produced with the endogenous retrovirus ENV1 fragment species specific for *Bos taurus*. This retrovirus belongs to the family of class I retroviruses; according to publications, the regions homologous to it are presented in bovine genome in multiple copies [25]. As follows from the characteristics of the spectra of the amplicons produced with Rec-dir2 primer (see Table 2), PIC value was significantly lower in this case than in the spectra with L1dir1 and Rec-rev2 primers. This indicates the less Rec-dir2 involvement in transposition compared with L1_BT at least for small distances from each other (not more than 2000 bp) in alternate DNA strands. Our findings support the assumption that the product of recombination between L1_BT and ENV1_BT is of a relatively more ancient origin compared to polymorphic transpositions of these retrotransposons in alternative DNA strands in bovine genome which were identified in this study.
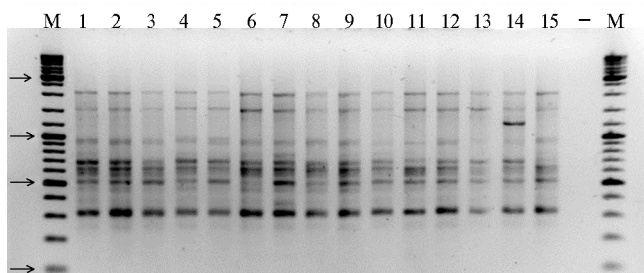

Fig. 5. Electrophoretic spectra produced in PCR amplification of genomic DNA regions flanked by Rec-dir2 primer inverted repeat, in Kalmyk breed: 1-15 (sample numbers) — K04, K06, K08, K11, K12, K13, K15, K16, K18, K19, K20, K21, K22, K25, K26, respectively; «–» — control (no genomic DNA); M — molecular weight marker O'GeneRuler DNA Ladder mix (100-10000 bp) (Thermo Fisher Scientific, USA). Arrows mark the fragments of 3000, 1000, 500, and 100 bp. Characteristics of amplification products spectra are provided in Table 2, for primer description please refer to Table 1.

With Rec-dir2 as the primer, most complex spectra of DNA sequences unique in their length were detected in Kalmyk breed (Fig. 5).

Based on the polymorphism of DNA fragments flanked by inverted ERV1_BT (primer Rec-dir2) repeat, we estimated genetic distances between specimens and constructed a dendrogram (Fig. 4). Two clusters are distinguished in it, one of them combines specialized dairy Ayrshire breed and Black and White Holstein breeds, the second one combines the Kalmyk breed animals.

These findings suggest that distribution of the ENV1_BT endogenous retrovirus fragment inverted repeats in bovine genomes differentiates dairy and beef cattle. It should be noted that the dairy cattle breeding was originated several thousand years ago [28]. The results of comparative analysis of L1_BT ENV1_BT inverted repeats distribution suggest that L1_BT transposition activity is substantially higher than that of ENV1_BT. This is consistent with the available reports of L1_BT wide representation in bovine genomes and of high transposition activity of this mobile element [20].

Thus, from our study, the product of recombination between species-specific for bovine genome mobile genetic elements, the long interspersed nuclear element LINE L1 Bos taurus (L1_BT) and endogenous retrovirus ENV1 Bos taurus (ENV1_BT) discovered earlier, turned out to be conservative in the studied breeds which apparently indicates the antiquity of its formation. Genetic relationship identified by analyzing the distribution and length polymorphism of DNA regions flanked by inverted ENV1_BT repeats, differ significantly in specialized dairy breeds and in the local beef breed, but at the same time they are

characterized by low interbreed variability. A relatively low interbreed variability makes it possible to recommend multi-locus genotyping using ENV1_BT sites as primers to assess the genetic differentiation of breeds, diverged from the common root long ago. Distribution of inverted repeats of the regions homologous to L1_BT has a relatively higher interbreed variability. This may be due to increased L1_BT transposition activity and can be used in the study of individual interbreed variability for the purposes of genomic selection.

R E F E R E N C E S

1. Daetwyler H.D., Capitan A., Pausch H., Stothard P., van Binsbergen R., Brøndum R.F., Liao X., Djari A., Rodriguez S.C., Grohs C., Esquerré D., Bouchez O., Rossignol M.N., Klopp C., Rocha D., Fritz S., Eggen A., Bowman P.J., Coote D., Chamberlain A.J., Anderson C., VanTassell C.P., Hulsegge I., Goddard M.E., Guldbrandtsen B., Lund M.S., Veerkamp R.F., Boichard D.A., Fries R., Hayes B.J. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.*, 2014, 46(8): 858-865 (doi: 10.1038/ng.3034).
2. Hozé C., Fritz S., Phocas F., Boichard D., Ducrocq V., Croiseau P. Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. *J. Dairy Sci.*, 2014, 97(6): 3918-3929 (doi: 10.3168/jds.2013-7761).
3. Stothard P., Liao X., Arantes A.S., De Pauw M., Coros C., Plastow G.S., Sargolzaei M., Crowley J.J., Basarab J.A., Schenkel F., Moore S., Miller S.P. A large and diverse collection of bovine genome sequences from the Canadian Cattle Genome Project. *Gigascience*, 2015, 4: 49 (doi: 10.1186/s13742-015-0090-5).
4. Xu L., Bickhart D.M., Cole J.B., Schroeder S.G., Song J., VanTassell C.P., Sonstegard T.S., Liu G.E. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. *Mol. Biol. Evol.*, 2015, 32(3): 711-725 (doi: 10.1093/molbev/msu333).
5. Patry C., Jorjani H., Ducrocq V. Effects of a national genomic preselection on the international genetic evaluations. *J. Dairy Sci.*, 2013, 96(5): 3272-3284 (doi: 10.3168/jds.2011-4987).
6. Zhao F., McParland S., Kearney F., Du L., Berry D.P. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Select. Evol.*, 2015, 47: 49 (doi: 10.1186/s12711-015-0127-3).
7. Sempéré G., Moazami-Goudarzi K., Eggen A., Laloë D., Gautier M., Flori L. WIDDE: a Web-Interfaced next generation database for genetic diversity exploration, with a first application in cattle. *BMC Genomics*, 2015, 16(1): 940 (doi: 10.1186/s12864-015-2181-1).
8. Cicconardi F., Chillemi G., Tramontano A., Marchitelli C., Valentini A., Ajmone-Marsan P., Nardone A. Massive screening of copy number population-scale variation in *Bos taurus* genome. *BMC Genomics*, 2013, 14: 124 (doi: 10.1186/1471-2164-14-124).
9. Xu L., Cole J.B., Bickhart D.M., Hou Y., Song J., VanRaden P.M., Sonstegard T.S., VanTassell C.P., Liu G.E. Genome wide CNV analysis reveals additional variants associated with milk production traits in Holsteins. *BMC Genomics*, 2014, 15: 683 (doi: 10.1186/1471-2164-15-683).
10. Bickhart D.M., Liu G.E. The challenges and importance of structural variation detection in livestock. *Front. Genet.*, 2014, 5: 37 (doi: 10.3389/fgene.2014.00037).
11. Kalendar' R.V., Glazko V.I. *Fiziologiya i biokhimiya kul'turnykh rastenii*, 2002, 34(4): 279-296.
12. Glazko V.I., Kosovskii G.Yu., Koval'chuk S.N., Arkhipov A.V., Petrova I.O., Dedovich G.O., Glazko T.T. *Innovatsionnye tekhnologii v meditsine*, 2014, 2(03): 63-79.
13. Elsik C.G., Tellam R.L., Worley K.C. et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 2009, 324: 522-528 (doi: 10.1126/science.1169588).
14. Walsh A.M., Kortschak R.D., Gardnerb M.G., Bertozzia T., Adelsona D.L. Widespread horizontal transfer of retrotransposons. *PNAS USA*, 2013, 110(3): 1012-10160 (doi: 10.1073/PNAS.1205856110).
15. Smyka P., Kalendar R., Ford R., Macas J., Griga M. Evolutionary conserved lineage of *Angela*-family retrotransposons as a genome-wide microsatellite repeat dispersal agent. *Heredity*, 2009, 103: 157-167 (doi: 10.1038/hdy.2009.45).
16. McInerney C.E., Allcock A.L., Johnson M.P., Bailie D.A., Prodohl P.A. Comparative genomic analysis reveals species dependent complexities that explain difficulties with microsatellite marker development in mollusks. *Heredity*, 2011, 106: 78-87 (doi: 10.1038/hdy.2010.36).
17. Adelson D.L., Raison J.M., Garber M., Edgar R.C. Interspersed repeats in the horse (*Equus caballus*); spatial correlations highlight conserved chromosomal domains. *Anim. Genet.*, 2010, 41 (Suppl. 2): 91-99 (doi: 10.1111/j.1365-2052.2010.02115.x).
18. Grandi F.C., An W. Non-LTR retrotransposons and microsatellites. Partners in genomic variation. *Mobile Genetic Elements*, 2013, 3: e25674 (doi: 10.4161/mge.25674).
19. Startek M., Szafranski P., Gambin T., Campbell I.M., Hixson P., Shaw C.A.,

Stankiewicz P., Gambin A. Genome-wide analyses of LINE-LINE-mediated nonallelic homologous recombination. *Nucl. Acids Res.*, 2015, 43(4): 2188-2198 (doi: 10.1093/nar/gku1394).

20. Adelson D.L., Raison J.M., Edgar R.C. Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *PNAS USA*, 2009, 106(31): 12855-12860 (doi: 10.1073/pnas.0901282106).

21. Garcia-Vallve S., Puigbo P. *DendroUPGMA: A dendrogram construction utility*. 2002 (http://genomes.urv.cat/UPGMA/).

22. Nei M., Li W.-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *PNAS USA*, 1979, 76: 5269-5273.

23. Page R. TreeView: An application to display phylogenetic trees on personal computers. *CABIOS applications note*, 1996, 12(4): 357-358.

24. Peakall R., Smouse P.E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research — an update. *Bioinformatics*, 2012, 28: 2537-2539 (doi: 10.1093/bioinformatics/bts460).

25. Xiao R., Park K., Oh Y., Kim J., Park C. Structural characterization of the g-genome of BERV γ4, the most abundant endogenous retrovirus family in cattle. *Mol. Cells*, 2008, 26: 404-408.

26. Garcia-Etxebarria K., Jugo B.M. Genome-wide detection and characterization of endogenous retroviruses in *Bos taurus*. *J. Virol.*, 2010 Oct, 84(20): 10852-10862 (doi: 10.1128/JVI.00106-10).

27. Saylor B., Elliott T.A., Linquist S., Kremer S.C., Gregory T.R., Cottenie K. A novel application of ecological analyses to assess transposable element distributions in the genome of the domestic cow, *Bos taurus*. *Genome*, 2013, 56: 521-533 (doi: 10.1139/gen-2012-0162).

28. Beja-Pereira A., Luikart G., Bradley D., Bradley D.G., Jann O.C., Bertorelle G., Chamberlain A.T., Nunes T.P., Metodiev S., Ferrand N., Erhardt G. Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nature Genet.*, 2003, 35(4): 311-313 (doi: 10.1038/ng1263).