DNA Sequencing of Bovine Leukemia Proviruses with Determination of the Genetic Status of Isolated Isolates

N G Kozyreva¹, L A Ivanova¹, T V Stepanova¹

¹Laboratory of Leukemia, Federal Scientific Center - All-Russian Scientific Research Institute of Experimental Veterinary Medicine named after K.I. Scriabin and Ya.R. Kovalenko of the Russian Academy of Sciences (FGBNU FSC VIEW RAS), Russia, 109428, Moscow, Russia, Ryazansky pr., 24, building 1

E-mail: nk07-73@mail.ru

Abstract. This paper presents data on the study of bovine leukemia virus (BLV) polymorphism in the Ulyanovsk region. A phylogenetic analysis of 33 field BLV isolates isolated from seropositive animals was performed. The nucleotide sequence of env and pol gene regions was determined.

The study revealed heterogeneity of the studied BLV population by the env gene, which corresponds to 2 genotypes of the virus: IV (97%), VII/VIIA (3%) with confirmation of the dominance of the IV genetic variant of the pathogen in Russia, and homogeneity by the pol gene, expressed in clustering of all analyzed isolates with those representatives of the RRJEM database and international isolates assigned to clade 1 (100%). The revealed nature and degree of genetic differences/divergences maintain the general trend in the RRJEM database and do not exceed the existing range.

1. Introduction

Enzootic bovine leukemia is a malignant lymphoproliferative disease. The etiological agent of this disease is bovine leukemia virus (Bovine leukemia virus, BLV), belonging to the genus Deltaretrovirus from the family of RNA-containing viruses Retroviridae, which also includes T-lymphotropic viruses of primates (Primate T-lymphotropic virus, PTLV), including monkeys (Simian T-lymphotropic virus, STLV), humans (Human T-lymphotropic virus, HTLV).

Retroviruses are currently important pathogens for both humans and many animal species. The family of these viruses is distinguished by the presence of a unique stage in the replication cycle – reverse transcription, mediated by the enzyme reverse transcriptase (RT) – RNA-dependent DNA polymerase, which is encoded by the pol gene in the genome of retroviruses [1-2]. It is known that in vivo, the comparable rate of mutations in the env gene region in BLV and lentiviruses is 0.009 and 0.0085 nucleotide changes per year, respectively. At the same time, lentiviral reverse transcriptase (OT) is wrong in 81% of cases with the production of non-synonymous nucleotide substitutions against that of the BLV enzyme, which is wrong in 50% of cases with an amino acid change [3-5].

Molecular diagnostics uses the most conservative regions of the BLV genome located in the gag, pol, and env genes. The variability of the pol gene used as a standard for phylogenetic comparison of the nucleotide sequence of different BLV and PTLV strains is 42% [6]. Within the retrovirus family, BLV forms a separate phylogenetic branch, where the differences in nucleotide sequences are less than 6% for the pol and env genes, which is presumably due to the higher accuracy of the OT enzyme
Due to the important biological functions and localization on the viral surface of the gp51 protein encoded by the env gene, this site is the main subject of phylogenetic analysis [8]. A significant conservativeness of the gp51 sequence was revealed in strains from several geographical locations [9-10], and the small differences found are mainly represented by point mutations [6;11], some of which lead to differences in restriction sites, which allows to classify BLV by different genotypes [12-13]. Phylogenetic studies based on partial or complete sequences of the env gene show clustering of the pathogen both into separate genotypes, as a result of which at least 11 genetic variants of BLV (G1-G11) widely distributed around the world are currently classified [12;14-21], and into subgroups/subgenotypes (G4, G6, G7) [17-18;22-23]. Conformational changes in this part of the virus genome or the absence of glycoprotein expression on the surface of infected host cells may provide new opportunities for elucidating mechanisms of the phenomenon of "escape" from the host's immune system during the retroviral strategy [24].

Genetic differences between viruses circulating in different livestock populations, which undoubtedly reflects its evolution, can be used as markers in the study of this infection in order to develop or improve means of specific prevention of the disease (diagnosticums, vaccines), determine the source of the pathogen in export-import operations, as well as establish the relationship between infection with individual genetic variants of BLV and the severity of the infectious process in the host body [25-26].

To study the genetic diversity of BLV isolates and identify their relationship to one of the known genotypes, as well as with a high probability of finding out the origin of the pathogen, it is necessary to determine the primary nucleotide sequence of the pol and env genes. Such topical aspects of studying BLV have scientific and practical significance.

The aim of our work was to determine the primary sequence of pol and env gene loci with further phylogenetic analysis and assessment of the genetic status of isolated BLV isolates circulating in the Ulyanovsk region.

2. Research materials and methods

The following research methods were used: molecular-biological – PCR, DNA sequencing; bioinformatic – phylogenetic analysis. Methodological procedures were performed according to the sources [27] [28, with our modifications], as well as the manufacturer's instructions when using commercial diagnostic kits.

Blood samples stabilized with EDTA-K3 were examined from 33 seropositive animals from an animal husbandry in the Ulyanovsk region of the Volga Federal district of the Russian Federation.

Proviral DNA was isolated from 100 µl of whole blood by nucleic acid precipitation with isopropanol and elution in 50 µl of buffer using the "Ribo-Prep" kit (Federal Budget Institution of Science Central Research Institute of Epidemiology of Rospotrebnadzor, Russia).

Identification of target fragments was performed using PCR with electrophoretic detection in the following variants: classical single-round PCR with the production of amplicons of 438 p. o. (pol); "nested" PCR with the production of amplicons of 341 p. o. (env).

Sequencing was performed directly on an automatic Beckman Coulter analyzer in accordance with the manufacturer's recommendations.

The primary nucleotide sequences obtained as a result of sequencing were identified in the GenBank data Bank using the blast service of the NCBI resource [29]. For comparison, we used reference nucleotide sequences of pol and env gene loci of international and Russian isolates of bovine leukemia virus, presented in GenBank and RRIEVM databases.

The evolutionary analysis of gene loci (pol – approximately 400 nucleotides; env – approximately 300 nucleotides) was performed using the Mega V. 6 program. Dendrograms are constructed using distance methods of evolution minimum (ME) [30], neighbor joining (NJ) [31] with determination of p-distances. The statistical reliability of the tree topology was evaluated using the bootstrap analysis method at 1000 iterations. Evolutionary distances were calculated using the models of Kimura [32], Tajima, and Nei [33].
3. The results of the study and discussion
In the course of the work, sequences of pol and env gene fragments of BLV isolates circulating in the Ulyanovsk region and isolated from 33 seropositive animals were obtained. Based on the results of phylogenetic analysis of field samples of proviral DNA of pathogen isolates, dendrograms were constructed, and the genetic status for each of them was established (Figures 1, 2).

When analyzing the pol gene locus, it was found that the General trend for the RRIEVM database persists – all analyzed Russian isolates (33/33, 100%) were grouped with representatives of cluster 1 (Figure 1).

During the determination of the structure of stable nucleotide substitutions in the sequences of the proviral pol gene locus of the studied BLV isolates, an average of 7 transitions (from 2 to 9) and 1 transversion were detected depending on the reference sequence. Single nucleotide substitutions of various types with a predominance of transitions were observed: A→G, G→A, C→T, T→C, which are evenly distributed and occur mainly in positions 1-3 in 11-33 isolates (33-100%); stable transversions were also detected: T→A, found when compared with two (K02120, M16017) of the seven reference sequences and, which is typical, mainly for all (33, 100%) or most of the studied isolates (31, 94%). The structure of the identified substitutions for these isolates corresponds to those of the representatives of the BLV RRIEVM database.

The degree of divergence/discrepancy between the analyzed isolates averaged 0.019±0.004 with differences of 6-8 mismatched nucleotides depending on the reference sequence, while within clade 1, the average was 0.021±0.006, which corresponds to the results of international studies [16].

The study of polymorphism of the env gene locus allowed us to estimate the heterogeneity of the studied group of BLV isolates (Figure 2). Thus, 32 (32/33, 97%) pathogen isolates were assigned to GIV BLV: 17 of these studied isolates were grouped with the M35238 France strain; a group of 5 isolates was clustered with the international M35240 Belgium strain. 1 (1/33, 3%) isolate 19/12 ULJAN was grouped with GVII BLV and assigned to the GVIIA subgroup, showing that they are 100% identical to the Russian isolates HM563748 Russia, HM563750 Russia, and HM563749 Russia, which confirms the role of geographical clustering in the spread of BLV [18].
**Figure 1.** Phylogenetic comparison of the pol gene regions of BLV proviruses. A – the main tree is constructed using the remote method of joining neighbors, using bootstrap analysis with 1000 random samples, the data array is 91 sequences; B — a branch with representatives of clade 1. Symbols indicate: ▲ – studied isolates; O – previously characterized isolates of the Russian Research Institute of Experimental Veterinary Medicine (RRIEVM) database; ■ – KK FLK-BLV (K+) isolate.
Figure 2. Phylogenetic comparison of the env gene regions of BLV proviruses. The tree was constructed using the remote method of minimal evolution, using p-distances, bootstrap analysis with 1000 random samples, and the data array consists of 103 sequences. Symbols indicate: ▲ – the studied isolates. ▼ – isolate KK FLK-BLV (K+).

The degree of divergence between the studied isolates and reference strains averaged 0.013 (1.3%) for GIV and 0.015 (1.5%) for GVII (0.004 (0.4%) for the GVIIA subgroup), which corresponds to the results of international studies [16;18].

In our research, we have shown the dominance of the IV genotype of the virus, which corresponds to the general trend of BLV distribution in the Russian Federation when we studied the polymorphism of this retrovirus ([34], according to the information from our own RRIEVM database (Russian Research Institute of Experimental Veterinary Medicine) and confirms the long-term circulation in Russia of the BLV variant imported from Europe with infected cattle after World War II. Today, BLV, on the one hand, has its distribution far beyond the territories of Memel (Klaipeda, from 1945 to 1991 as part of the Lithuanian SSR, now – Lithuania, authors’ note) and Konigsberg (since 1946 Kaliningrad in the USSR, now Russia, authors’ note) East Prussia since 1916 [35]. On the other hand, such extensive dissemination of infection occurred due to the import of livestock from European countries to other countries (for example, the countries participating in the COMECON (Council for mutual economic assistance (1970-80)), free of leukemia, through international trade operations conducted with various breeds of livestock, which may explain the topology of the most common variant of the pathogen.

When analyzing the topology of the GVII genotype, the existence of well-defined subgroups was shown, each of which contained exclusively isolates from Russia (GVIIA), Ukraine (GVIIIB) and Poland (GVIIIC) [18], from Moldova GVIID, GVIIIE [17]. These isolates were identified as belonging to one specific region of each country, i.e. clustering occurred in accordance with the place of origin and unique and specific amino acid mutations were identified for some subgroups of the GVII genotype: for GVIIC - H142R [18]; for GVIIIE – Y229C; for GVIID – N141H. The tendency to diversify the virus into homogeneous genetic groups is common. After dispersal, the virus then assimilates into the herd, and gradually its populations become homogeneous [17]. It is possible that the presence of such genetic groups of BLV occurs as a result of the redistribution of part of the virus
population, which presumably confirms the role of geographical clustering in the spread of BLV due to genetic drift [18]. The large-scale distribution of virus genotypes within and between remote geographical regions may be initiated by the movement of animal populations associated with the migration of humans and livestock, as well as with viral transmission between individual animals over a long period of close contact [36]. The presence of more than one BLV genotype in certain geographical areas is also reflected [10;22;37].

In connection with health measures, there is a constant culling of livestock for hematological and serological signs of infection, which leads to the appearance of animals with an alaikemic state in the absence of an immune response, accompanied by the production of antibodies, and, at the same time, it is quite possible that BLV variants with mutations in epitopes that affect the structure and further function of the corresponding protein (artificial selection), which, ultimately, can lead to the manifestation of the phenomenon of "withdrawal" from the immune response and have a potential impact on the results of diagnostic tests, which must be taken into account when developing diagnostic and preventive drugs for the purpose of effective anti-leukemia measures.

4. Conclusions
Nucleotide sequences of BLV isolates circulating among animals in the Ulyanovsk region were obtained.

This pathogen population is dominated by the GIV genetic variant BLV (32/33, 97%).

The study revealed heterogeneity of the pathogen population by the env gene: GIV (32/33, 97%) and GVII/subgroup GVIIA (1/33, 3%) genotypes.

All BLV isolates characterized by the pol gene belong to cluster 1 (33/33, 100%), where all the previously studied BLV representatives are distributed.

The detected genetic variability among the analyzed isolates for the env GIV – 1.3%, GVII – 1.5% and pol – 1.9% genes does not exceed the average intra-group divergences and does not exceed the range within the RRIEVM own database for the nucleotide sequences of BLV isolates for the env GIV – 1.5%, GVII – 1.6% and pol – 2.1% genes.

5. References
[1] Gerald G and Grangenet D 1980 Retrovirus reverse transcriptase Molecular Biology of RNA tumor viruses ed J Stephenson (New York: Academic Press) pp 346-94
[2] Kettmann R and Burny A 1994 Bovine Leukemia Virus The Retroviridae vol 3 ed J A Levy (New York: Plenum Press) pp 39-81
[3] Willems L, Thiempont E, Kerkhofs P, Burny A, Mammerickx M and Kettmann R 1993 Bovine leukemia virus an animal model for the study of intrastrain variability. J Virol. 67 1086-9
[4] Burns D P, Desrosiers R C 1991 Selection of genetic variants of simian immunodeficiency virus in persistently infected rhesus monkeys J Virol. 65(4) 1843-54
[5] Overbaugh J, Rudensey L M, Papenhausen M D, Benveniste R E and Morton W R 1991 Variation in simian immunodeficiency virus env is confined to V1 and V4 during progression to simian AIDS J Virol. 65(12) 7025-31
[6] Mamoun R Z et al., 1990 Sequence variability of bovine leukemia virus env gene and its relevance to the structure and antigenicity of the glycoproteins J Virol. 64(9) 4180-8
[7] Gillet N et al, 2007 Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human Retrovirology 4 18
[8] Polat M, Takeshima S and Aida Y 2017 Epidemiology and genetic diversity of bovine leukemia virus J Virol. 14 209
[9] Camargos M F, Stancek D, Rocha M A, Lessa L M, Reis J K and Leite R C 2002 Partial sequencing of env gene of bovine leukaemia virus from Brazilian samples and phylogenetic analysis J. Vet. Med. B Infect. Dis. Vet. Public Health 49 325-31
[10] Camargos M F, Pereda A, Stancek D, Rocha M A, dos Reis J K, Greiser-Wilke I and Leite R C 2007 Molecular characterization of the env gene from Brazilian field isolates of bovine
leukemia virus *Virus Genes* **34** 343-50

[11] Beier D, Blankenstein P, Marquardt O and Kuzmak J 2001 Identification of different BLV provirus isolates by PCR, RFLPA and DNA sequencing *Berliner und Munchener Tierarztliche Wochenschrift* **114**(7–8) 252-6

[12] Licursi M, Inoshima Y, Wu D, Yokoyama T, González E T and Sentsui H. 2003 Provirus variants of bovine leukemia virus in naturally infected cattle from Argentina and Japan *Vet Microbiol.* **96**(1) 17-23

[13] Fechner H et al, 1997 Provirus variants of the bovine leukemia virus and their relation to the serological status of naturally infected cattle *J. Virol* **237**(2) 261-9

[14] Lee E, Kim E J, Joung H K, Kim B H, Song J Y, Cho I S, Lee K K and Shin Y K 2015 Sequencing and phylogenetic analysis of the gp51 gene from Korean bovine leukemia virus isolates *Virol. J.* **12**(64)

[15] Lee E, Kim E J, Ratthanophart J, Vitoonpong R, Kim B H, Cho I S, Song J Y, Lee K K and Shin Y K 2016 Molecular epidemiological and serological studies of bovine leukemia virus (BLV) infection in Thailand cattle *Infect Genet. Evol.* **41** 245-54

[16] Polat M et al, 2016 A new genotype of bovine leukemia virus in South America identified by NGS-based whole genome sequencing and molecular evolutionary genetic analysis *Retrovirology* **13**(4) 23

[17] Pluta A, Rola-Luszczak M, Kubis P, Balov S, Moskalik R, Choudhury B and Kuz´mak J 2017 Molecular characterization of bovine leukemia virus from Moldovan dairy cattle *Arch. Virol.* **162**(6) 1563-76

[18] Rola-Luszczak M, Pluta A, Olech M, Donnik I, Petropavlovskiy M, Gerilovych A, Vinogradova I, Choudhury B and Kuzmak J 2013 The molecular characterization of bovine leukemia virus isolates from Eastern Europe and Siberia and its impact on phylogeny *PLoS One* **8**(3)

[19] Donnik I M et al, 2016 Revisiting the issue of the molecular–genetic structure of the causative agent of the bovine leukemia virus in the Russian Federation *Indian Journal of Science and Technology* **9**(42)

[20] Shaeva A Y, Garaeva Z R, Vafin R R, Khazipov N Z, and Alimov A M 2012 Identification of a new BLV genotype *Scientific notes of the Kazan State Academy of Veterinary Medicine named after N.E. Bauman* **211** 192-7 In Russian

[21] Yu C, Wang X, Zhou Y, Wang Yu, Zhang X and Zheng Y 2019 Genotyping bovine leukemia virus in dairy cattle of Heilongjiang, northeastern China *BMC Vet Res* **15** 179

[22] Polat M, Ohno A, Toshimasa S N, Kim J, Kikuya M, Matsumoto Y, Mingala C N, Onuma M and Aida Y 2015 Detection and molecular characterization of bovine leukemia virus in Philippine cattle *Arch Virol.* **160** 285-96

[23] Gautam S, Mishra N, Kalaiyarasu S, Jhade S K and Sood R 2018 Molecular characterization of bovine leukaemia virus (BLV) strains reveals existence of genotype 6 in cattle in India with evidence of a new subtype *Transbound Emerging Diseases* **65**(6) 1968-78

[24] Pluta A, Albritton L M, Rola-Luszczak M and Kuzmak J 2018 Computational analysis of envelope glycoproteins from diverse geographical isolates of bovine leukemia virus identifies highly conserved peptide motifs *Retrovirology* **15**(1) 2

[25] Inoue E, Matsumura K, Maekawa K, Nagatsu K, Nobuta M, Hirata M, Minagawa A, Osawa Y and Okazaki K 2011 Genetic heterogeneity among bovine leukemia viruses in Japan and their relationship to leukemogenicity *Arch. Virol.* **156** 1137-41

[26] Aida Y, Murakami H, Takahashi M and Takeshima S N 2013 Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus *Front. Microbiol.* **4** 328

[27] Methodical guidelines for the detection of DNA of bovine leukemia provirus by PCR method and sample preparation of the obtained genetic material for the purpose of further determination of the nucleotide sequence by sequencing Gulyukin M I, Kozyreva N G, Ivanova L A, Kolbasov D V, Tsybanov S Zh, Kalabekov I M and Malogolovkin A S 2011 *Modern means and methods*
of ensuring veterinary welfare in infectious and protozoal pathology of animals, fish and bees ed Gulyukin M. I. et al (Moscow: Agency for Creative Technologies Publ.) 356 p In Russian

[28] Belak S and Ballagi–Pordány A 1993 Experiences on the application of the polymerase chain reaction in a diagnostic laboratory Mol. Cell. Probes 7 241-8

[29] National Center for Biotechnology Information (NCBI) [Electronic resource] U.S. National Library of Medicine 8600 Rockville Pike Bethesda MD 20894 1998

[30] Rzhetsky A and Nei M 1993 Theoretical foundation of the minimum-evolution method of phylogenetic inference Mol. Biol. Evol. 10(5) 1073-95

[31] Saitou N and Nei M 1987 The neighbor-joining method: a new method for reconstruction of phylogenetic trees Mol. Biol. Evol. 4(4) 406-25

[32] Kimura M. 1980 A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences Journal of Molecular Evolution 16 111-20

[33] Tajima F and Nei M. 1984 Estimation of evolutionary distance between nucleotide sequences Molecular Biology and Evolution 1(3) 269-85

[34] Kozyreva N G and Gulyukin M I 2017 Distribution of cattle leukemia and genetic variants of the pathogen on the territory of livestock farms in the Central Federal District of the Russian Federation Veterinariya Kubani 6 4-9

[35] Burba L G et al, 1977 Leukemias and malignant tumors of animals ed Shishkov V P, Burba L G (Moscow: Kolos) 375 p.

[36] Rodríguez S M, Golemba M D, Campos R H, Trono K and Jones L R 2009 Bovine leukemia virus can be classified into seven genotypes: evidence for the existence of two novel clades Journal of General Virology 90(11) 2788-97

[37] Zhao X and Buehring G C 2007 Natural genetic variations in bovine leukemia virus envelope gene: possible effects of selection and escape Virology 366 150-65

Acknowledgements
The work was performed as part of the state task in accordance with the approved research plan of the Science and Research Work Department of the Federal Science Center — Russian Science and Research Institute of Experimental Veterinary Medicine named after K. Skrybain and Ya. Kovalenko of the Russian Academy of Sciences for 2020.