

## Designing of a DNA matrix for transgene integration into the bovine beta-lactoglobulin gene locus using CRISPR/Cas9 technology

E.M. Koloskova\*, V.A. Ezerskiy, K.S. Ostrenko

*All-Russia Research Institute of Animal Physiology, Biochemistry, and Nutrition,  
Branch of Ernst Russian Institute of Animal Breeding Federal Science Center for Animal Husbandry,  
Federal State Budgetary Scientific Institution, Borovsk, Russian Federation*

*\*Corresponding author E-mail: [heleko3@yandex.ru](mailto:heleko3@yandex.ru)*

*Received: 03.09.2020. Accepted: 26.10.2020*

Beta-lactoglobulin (BLG) is the main protein in milk serum in almost all mammals, with the exception of rodents and primates. Regulatory regions of the beta-lactoglobulin gene in ruminants (sheep, goats, and cattle) as part of genetic constructs provide tissue - specific expression of recombinant protein in the mammary gland and have been actively used in genetic engineering since the beginning of the era of creating transgenic animals. To work effectively with the CRISPR/Cas9 genomic editing method, it is necessary to know the exact DNA sequence of the target gene: this is necessary both for creating a DNA matrix for homologous recombination and for the targeted accuracy of guide RNAs. A polymorphic variant of the bovine BLG gene was identified, whose sperm was used to fertilize cow oocytes in vitro. The aim of this work was to create a plasmid containing 5' - and 3' - arms of homology (ha) to the bovine BLG gene. Based on *pTZ57R/T*, the pTZhaBLG plasmid was obtained, which has a unique site for *EagI* restriction at the junction of the homology arms. A fragment containing a biologically active protein gene can be embedded in the resulting plasmid at this restriction site. We created the pBLGcmvEGFP plasmid containing the green fluorescent protein (EGFP) gene under the cytomegalovirus (cmv) promoter: protein expression can serve as a reliable indicator of successful integration of the transgene into the genome. The resulting plasmids in circular or linearized form are intended for site-specific integration by homologous recombination repair into the BLG gene using CRISPR/Cas9 components.

**Key words:** genetic constructs,  $\beta$ -lactoglobulin, BLG gene, cattle, CRISPR/Cas9, microinjection, arms of homology, green fluorescent protein.

### Introduction

For the constantly growing population of the world, issues of medical, biological and food security will inevitably become more acute in the near future, which makes it urgent to find and develop effective ways to solve them. One of the most important areas in the development of such methods is research in the field of genetic engineering. The possibility of performing high-precision molecular genetic manipulations in the near future should accelerate the development of therapeutic biomaterials for the prevention and treatment of human diseases. The second promising area is the introduction of modifications to the genome of farm animals to increase the production and production of human food with specified properties (Tan et al., 2012; Bishop and Van Eenennaam, 2020).

Historically, the application of genetic engineering in animal husbandry has proved challenging, especially for large farm animals with a small number of offspring in the litter. To obtain genetically modified (GM) animals, the somatic cell nuclear transfer (SCNT) method was used in most studies, but its effectiveness was insufficient. Currently, new genome editing tools based on the action of genetically engineered endonucleases are used to produce animal GM. One of them, CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) has found the widest application for creating GM organisms. The method of microinjection of zygotes, previously used only for small laboratory animals with a large number of offspring (mice, rats, rabbits), has become available for obtaining transgenic (or other genetically modified) farm animals – goats, pigs, and cattle.

The application of endonuclease technologies using the homology directed repair (HDR) mechanism is an effective method for changing the composition of milk by replacing the genes of endogenous milk proteins in farm animals in order to improve the consumer qualities of milk and use the animal's mammary glands as a bioreactor for the production of heterologous proteins (Shepelev et al., 2018). It's possible to exclude any protein from the milk composition, replace it with a protein of pharmacological or diagnostic purpose, and modify milk proteins in order to increase their nutritional value and digestibility for the processing into different dairy products or to obtain biologically active peptides from them as vaccines (Whitelaw et al., 2016).

Regulatory regions of the genes of tissue – specific milk proteins - Caseins and whey proteins ( $\alpha$ -lactalbumin,  $\beta$ -lactoglobulin) - are used as part of genetic constructs (GC) to create transgenic animals-producers of biologically active proteins with milk. The

concentration of  $\beta$ -lactoglobulin in cattle milk is on average 4.6 mg/ml. Regulatory regions of the BLG gene in sheep, goats, and cattle are used as part of GC for obtaining transgenic animals using the classical transgenesis method (Trubitsina et al., 2018). For example, human alpha-1-antitrypsin expression in the milk of transgenic sheep reached 5 mg/ml (Wright et al., 1991) when using the sheep BLG gene promoter as part of the gene construct.

Cow's milk is a complete food product containing proteins, fats, carbohydrates, and minerals, but it often causes allergic reactions and reduces the effectiveness of using the nutrients of dairy products. BLG is the main dairy allergen. Milk Allergy is detected in 17.5% of young children, 13.5% of children aged 5 to 16 years, and 4% of the adult population [Fiocchi et al., 2010]. Sensitization to milk proteins is the beginning of the formation of a number of pathological processes, such as bronchial asthma, urticaria, atopic dermatitis, and gastrointestinal tract dysfunction. Food Allergy in some cases precedes the development of severe pollen allergy (Restani et al., 2009).

To reduce the allergenicity of milk, various methods are used: heating, high pressure, enzymatic hydrolysis, glycation (non-enzymatic glycosylation). These methods reduce the allergenicity of BLG to varying degrees, but the structure and functions of other milk proteins are disrupted, which significantly affects the nutritional value of milk. Knockout of the BLG gene using gene editing technology completely solves a problem that is of great importance for creating hypoallergenic dairy products. The BLG gene is a promising target both for its knockout to produce hypoallergenic milk, and for replacing it with a heterologous protein gene when using endonuclease editors. Precise embedding of the transgene as part of a gene construct containing arms of homology to the BLG gene will ensure its expression under the control of full-length endogenous regulatory sequences.

The aim of this work is to create a plasmid containing homology arms to the regulatory regions of the BLG gene in cattle so that the DNA sequence of the target protein can be integrated into it. To evaluate the effectiveness of integrating the HDR DNA-matrix into the genome of a microinjected zygote, a gene construct containing the green fluorescent protein (EGFP) gene under the cytomegalovirus (CMV) promoter was required. Integration of such a gene construct by homologous recombination into the bovine BLG locus during joint microinjection with site-specific components of the CRISPR/Cas9 system will allow *in vitro* visual evaluation of the effectiveness of selected guide RNAs using a luminescent microscope.

## Material and methods

All the work was performed in All-Russian Research Institute of Physiology, Biochemistry and Animal Nutrition in the framework of the state order 2019 on the theme 0445-2019-0030.

The sequence of the BLG gene of cattle (*Bos taurus*) was taken from the GenBank database, entry X14710 (<https://www.ncbi.nlm.nih.gov/nuccore/48,127>). Genomic DNA was extracted from the semen of a black-and-white bull named Moroz (Bykovo breeding farm) using the phenol-chloroform method. Selection of primers, construction of recombinant DNA, and restriction analysis were performed using the Vector NTI program.

For intermediate cloning of PCR products, the *pTZ57R/T* plasmid and T4 DNA ligase of the InsTAclone PCR Cloning Kit (Thermo Scientific) was used. All primers, oligonucleotides, and plasmid DNA sequences were ordered from "Syntol" company (<http://www.syntol.ru>).

The following enzymes and reagents were used: FastAP Thermosensitive Alkaline Phosphatase, *EagI* restrictases (analog – BseX31, "SibEnzyme"), *BglIII*, *BamHI*, *BsmBI* with corresponding buffers. A mixture of dNTP (2 mM, "Fermentas"), Taq polymerase (5 u/ml), 10×Taq buffer (with 25 mM  $MgCl_2$ ), and Pfu DNA polymerase (5 u/ml, "Silex") was used for polymerase chain reaction (PCR). The time and temperature parameters of PCR were selected depending on the structure of the primers. The sequences and descriptions of the primers used to create the genetic construction are shown in table 1. The PCR procedure was performed on a "Tertsik" DNA amplifier ("DNA technology LLC", Moscow).

**Table 1.** Primers used to create the gene construct

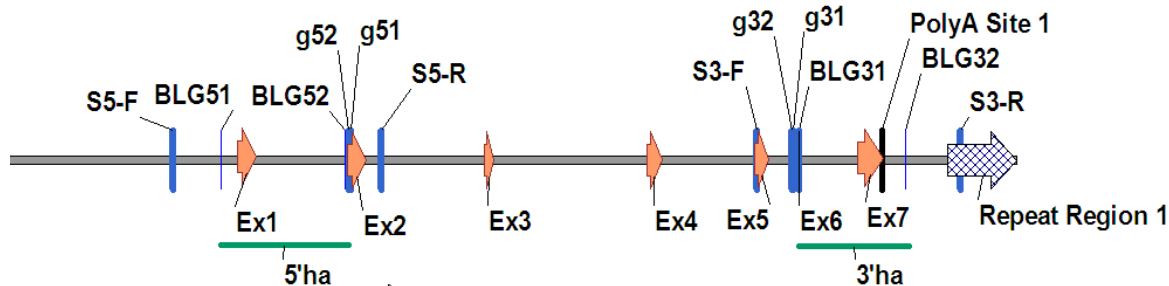
| Designation | 5'-3' sequence                          | Introduced site of the restriction |
|-------------|---|------------------------------------|
| S5-F        | gagctcgacctaaccat                       | -                                  |
| S5-R        | cctccaggggtaccagaaa                     | -                                  |
| S3-F        | ggacgacgaggccctggagaaatt                | -                                  |
| S3-R        | gctctccgcatcatgtggccta                  | -                                  |
| BLG51       | gccttagaaacccagggccaaggata              | <i>XbaI</i>                        |
| BLG52       | aaagtcgac cggccgacacctggggaggaccttgagct | <i>Sall, EagI</i>                  |
| BLG31       | ta cggccgagcagtgccacatctaggtgag         | <i>EagI</i>                        |
| BLG32       | gaagtcgac agatctgcccactccctacctggtt     | <i>Sall-BglIII</i>                 |

Transformation of competent cells *E. coli* Dh5a and TG1 was performed using the method and reagents of the Transform-Aid Bacterial Transformation Kit. The transformed cells were seeded on an agarized Luria-Bertani medium with ampicillin. The grown clones were sown again. DNA from grown clones was isolated for PCR analysis using a lysing mixture containing proteinase K. A suitable clone was grown in LB-Am+ medium. Plasmid DNA was isolated using the GeneJET Plasmid Miniprep Kit with our modifications (Koloskova and Ezerskiy, 2019) and the classical method of alkaline lysis. The quality and quantity of isolated DNA were evaluated visually in UV light after agarose gel (AG) electrophoresis. DNA from the gel was isolated using the GeneJET Gel Extraction Kit. The concentration of plasmid DNA was determined after its restriction (x1) using the method of sequential aliquot dilution. A linearized plasmid of known concentration in similar dilutions was used as a standard. Electrophoresis was performed in horizontal AG in a x0.5 TBE buffer, pH 8.0, with the ethidium bromide (EBr), using a set of equipment from Hoeffer (USA). Preparative electrophoresis was performed in x1 TAE buffer, pH 8.0, EBr. The size of DNA fragments in AG was estimated using the DNA Ladder Mix ("Fermentas") as the standard.

## Results and Discussion

*Genotyping of the BLG gene in cattle.* The BLG gene of cattle is located on the 10th chromosome, contains 4 exons and encodes a protein of 127 amino acids (of which 19 is a signal peptide). We presented the scheme for selecting homology arms and fragments for sequencing (Fig. 1). Since the efficiency of the CRISPR/Cas9 components depends on the quality of the sgRNA target sequences, the possible polymorphism of the targeted gene in the region of potential double-stranded sections (DSS) is of great importance when selecting them.

Polymorphism of the BLG gene can cause low efficiency of the CRISPR/ Cas9 components: the ovaries of cows are selected at meat processing plants from cows of different breeds. In vitro experiments, oocytes fertilization is usually performed with the sperm of a single donor bull. PCR amplifications with genomic DNA of a bull named "Moroz" of 5' and 3' regions of the BLG gene (about 1.5 kb each) were obtained using pairs of primers S5-F/ S5-R and S3-F/ S3-R. Amplifications were cloned into a *pTZ57R/T* plasmid vector to obtain sequences of selected fragments for sequencing (Fig.1).



**Fig. 1.** Scheme for selecting homology arms to the BLG gene of cattle for creating a DNA matrix of homologous recombination. The 5'-arm of homology (5'ha) with a size of 929 bp includes the 1st exon and 8 bp of the second exon. The 805-bp 3'-arm of homology (3'ha) contains a fragment of the 6th exon and the 7th exon. The binding sites of the primers used, localization of protospacer sequences (g52, etc.) are indicated.

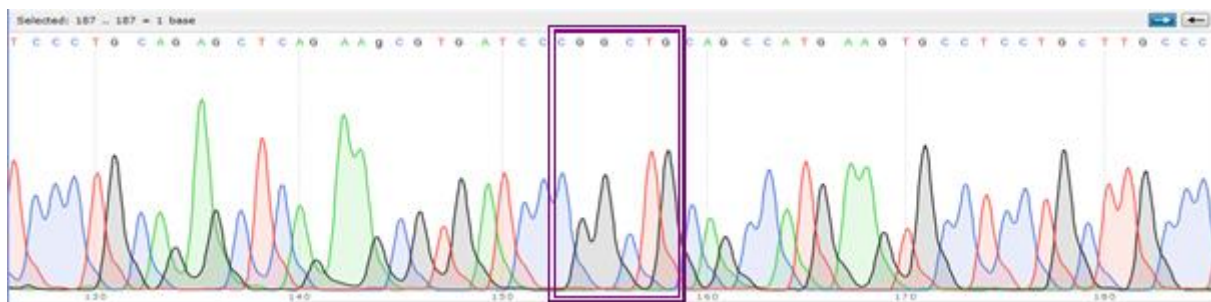
The classic method for genotyping the BLG gene in cattle is using the Medrano technique (Medrano and Aguilar–Cordova, 1990), which allows finding polymorphism in the 4th exon. Using the Gladyr method (Gladyr E.A., 2001), the region of the 1st exon and intron is analyzed, which allows to determine the polymorphism of the BLG gene from the "Moroz" bull DNA using one of the pairs of our primers and *PvuII* restrictase (table 2).

The genomic DNA extracted from the sperm was denatured at 94°C for 4 min., 35 amplification cycles were performed in the following mode: 94°C – 1 min., primer annealing at 62°C - 1 min., elongation at 72°C – 1.5 min. The final stage of synthesis was performed at 72°C for 4 minutes. PCR mixture (10 µl) was treated with restriction endonuclease *PvuII* (5 units) in the appropriate buffer at 37°C for 3 hours. After processing PCR products with *PvuII*, allele-specific fragments appear: 774 and 474 bp correspond to the A-allele and 774, 297 and 177 bp correspond to the B variant.

**Table 2.** Determination of gene polymorphism in the BLG bull "Moroz".

| Primers                       | According to the Gladyr method<br>Primer 1/Primer 2 | Our primers<br>BLG51/S5-R |
|-------------------------------|---|---------------------------|
| PCR amplifier, bp.            | 1248  | 1195                      |
| <i>PvuII</i> for the A-allele | 774, 474  | 743, 452                  |
| <i>PvuII</i> for the B-allele | 774, 297, 177                                       | 155, 297, 743             |

Sequencing and restriction analysis of cloned sequences of the BLG gene of the bull "Moroz" black-and-white breed showed an A-allele (table 2, Fig. 3), while the records X14710 and Z48305 (*B. taurus* gene for beta-lactoglobulin variant B) published in the GenBank database correspond to the B-allele.

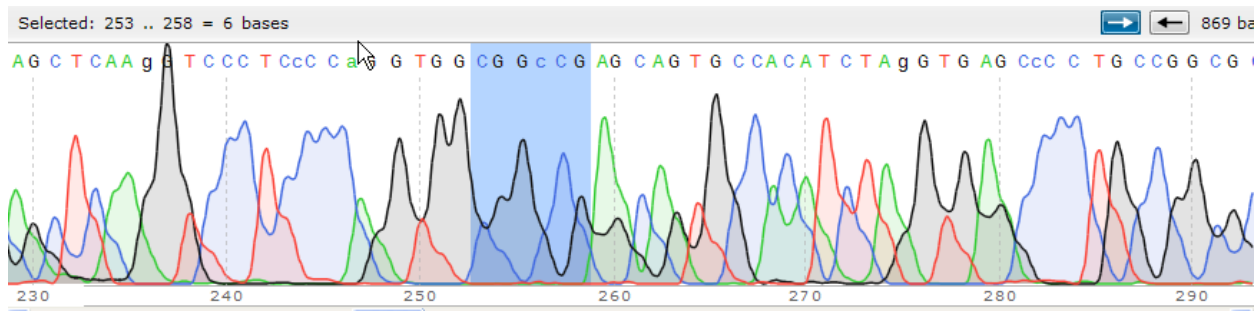


**Fig. 2.** Fragment of the sequence of the pTZ5'haBLG plasmid containing a section of polymorphism in the region of the 1st exon of the BLG gene (highlighted by a frame).

*Creating a pTZHABLG plasmid matrix.* The primers for PCR amplification of 5' and 3' homology arms (5'HA and 3'HA) included sequences for *XbaI*, *EagI*, and *EagI* and *BglIII* restrictases, respectively (table 1). The 5'HA fragment contained the 1st exon the

BLG gene with the sequence of the signal peptide BLG, the 1st intron. PCR amplifications were cloned into the *pTZ57R/T* vector to produce intermediate plasmids pTZ 5'HaBLG and pTZ3'HaBLG.

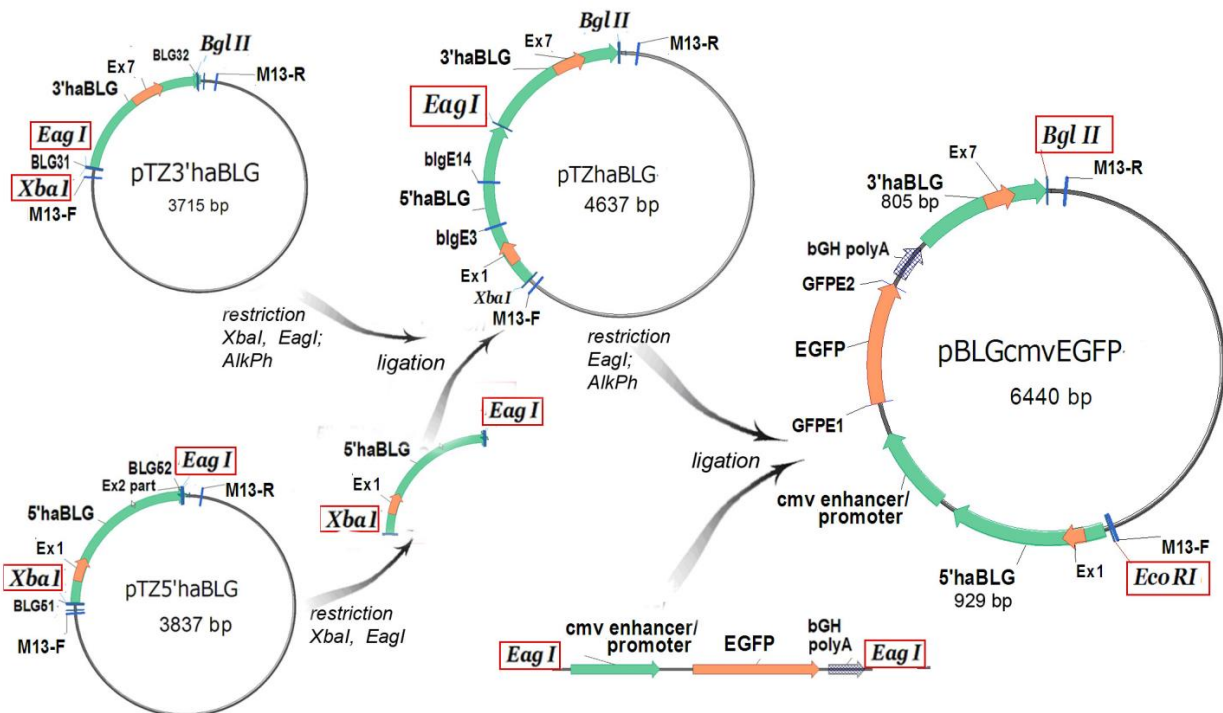
From the pTZ5'HaBLG plasmid with *Xba*I and *Eag*I restrictases, the *Xba*I-5'HaBLG-*Eag*I insertion fragment was cut out and, after purification, cloned into the *Xba*I-pTZ3'HaBLG-*Eag*I acceptor plasmid pretreated with the same restrictases and alkaline phosphatase. After ligation of the fragments, the pTZHaBLG plasmid was obtained, which has a unique site for *Eag*I restriction at the junction of the homology arms, which can be used for cloning the target protein's DNA (Fig.3).



**Fig. 3.** A fragment of a sequenced pTZhaBLG plasmid containing a junctionsite of 5'- and 3'- homology arms to the BLG gene. The sequence of the introduced restriction site *Eag*I is highlighted

**Creating a pBLGcmvEGFP plasmid.** The genes of fluorescent colored proteins are often used as reporters when creating genetic constructs: protein expression in the case of integration of a transgene into the genome can serve as a reliable indicator of successful transgenesis (Kato et al., 1999). It was assumed that the *cmvEGFP* sequence embedded in the gene structure between the 5' and 3' arms of homology to the BLG gene, in the case of homologous recombination using CRISPR/Cas9 technology, will provide in vitro expression of the green fluorescent protein. This will be an effective marker for evaluating the HDR event.

From the previously created plasmid pBLGfcmvEGFP (Ezerskiy and Shevchenko, 2008), the *cmvEGFP*-bGHpolyA fragment was cut by *Not*I restriction with the formation of *Eag*I sticky ends. The purified DNA insert was cloned into a prepared pTZHaBLG plasmid using the *Eag*I site. As a result, the pBLGcmvEGFP plasmid was obtained, which can be used as a DNA matrix for homologous recombination with the BLG gene using CRISPR/Cas9 components in circular or linear form (by cutting out the gene structure with *Eco*RI and *Bgl*II restrictases) (Fig. 4).



**Fig. 4.** Scheme for obtaining a pTZHaBLG matrix plasmid and a pBLGcmvEGFP model plasmid containing 5' and 3' homology arms to the BLG gene of cattle. The main restriction sites (within the framework) and binding sites of the primers used are indicated.

## Conclusions

The somatic cell nuclear transfer (SCNT) method is still the main method for obtaining genetically edited small-scale farm animals, while the method of microinjecting zygotes with recombinant DNA solution is considered simpler and more reliable, especially in combination with CRISPR/Cas9 technology. We created a recombinant pTZhaBLG plasmid containing 5' and 3' homology arms (929bp and 805 bp, respectively) to the BLG gene of cattle. The pTZhaBLG plasmid was created taking into account the BLG gene polymorphism of the bull, whose sperm will be used for fertilization of cow oocytes in vitro. The created



plasmid can be used for cloning the target biologically active protein gene using the *EagI* DNA site with the prospect of obtaining a transgenic cattle using CRISPR/Cas9 technology that produces recombinant proteins with milk instead of endogenous beta-lactoglobulin. In addition, a pBLGcmvEGFP plasmid was created containing a fragment of *cmvEGFP-bGHpolyA* (a reporter gene of a green fluorescent protein under a cytomegalovirus promoter) flanked by homology arms to BLG gene sequences to evaluate the effectiveness of transgene integration under *in vitro* culture conditions.

## Acknowledgements

The work was carried out within the framework of the state task No. 0445-2019-0030: "To create genetic constructs of the GFP gene under the cytomegalovirus (CMV) promoter for site-specific integration into the genome of rabbits and cattle using CRISPR/Cas9 technology. Obtain components of the CRISPR/Cas9 system for the WAP genes of rabbit and bovine BLG" approved and supported by the Ministry of Science and Higher Education of the Russian Federation. It was published with the support of the Federal State Budgetary Scientific Institution "Federal scientific center of animal husbandry – Russian Institute of animal husbandry named after academician L. K. Ernst", as part of the program "Research of molecular-biological and physiological-embryological aspects of bioengineering technologies for improving genetic resources and creating new breeding forms of farm animals and poultry" for the period 2015-2020. State registration number of research, development and technological works AAAA-A18-118021590132-9.

## Authors' contribution

KE developed a research program, conducted a comprehensive study, conducted work on the search for genetic sequences, selection of primers and restriction analysis in the VECTOR NTI program, analyzed the results of sequencing, participated in the implementation of genetic engineering methods, and compiled a manuscript. EV participated in a comprehensive study, performed genetic engineering and molecular biology methods, analyzed the data obtained, and participated in the drafting of the manuscript. All authors approved the final version of the manuscript.

## Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

## References

- Bishop, T.F. and Van Eenennaam, A.L. (2020). Genome editing approaches to augment livestock breeding programs. Review. *Journal of Experimental Biology*, 223, jeb207159. Doi:10.1242/jeb.207159
- Fiocchi, A., Brozek, J., Schunemann, H., Bahna, S.L., Von Berg, A., ..., Vieths, S. (2010). World Allergy Organization (WAO) diagnosis and rationale for action against cow's milk allergy (DRACMA) guidelines. *Pediatric Allergy and Immunology, Suppl.* 21, 57–161. Doi: 10.1097/WOX.0b013e3181defeb9
- Ezerskiy, V.A., Shevchenko, V.G. (2008). Creation of a genetic engineering design containing the structural gene of human lactoferrin under the control of regulatory elements of the bovine  $\beta$ -lactoglobulin gene and the GFP reporter gene. *Problems of Productive Animal Biology*, 2, 3-12 (In Russian)
- Gladyr', E.A. (2001). DNA diagnostics of kappa-casein and beta-lactoglobulin gene variants in cattle. Extended Abstract of the dissertation for the degree of candidate of Sciences biological, Dubrovitsy, VIZh (In Russian)
- Kato, M., Yamanouchi, K., Ikawa, M., Okabe, M. (1999). Efficient selection of transgenic mouse embryos using EGFP as a marker gene. *Molecular Reproduction and Development*, 54, 43-48. Doi: 10.1002/(SICI)1098-2795(199909)54:1<43::AID-MRD6>3.0.CO;2-N
- Koloskova, E.M., Ezerskiy, V.A. (2019). Method of an isolation and purification of plasmid DNA from transformed strains of *E. coli* DH 5 $\alpha$  and TG1. *Problems of Productive Animal Biology*, 4, 105-112. (In Russian)
- Medrano, J.F., and Aguilar-Cordova, E. (1990). Polymerase chain reaction amplification of bovine  $\beta$ -lactoglobulin genomic sequences and identification of genetic variants by RFLP analysis. *Animal Biotechnology*, 1, 73-74. Doi: 10.1080/10495399009525730.
- Restani, P., Ballabio, C., Di Lorenzo, C., Tripodi, S., Fiocchi, A. (2009). Molecular aspects of milk allergens and their role in clinical events. *Analytical and Bioanalytical Chemistry*, 395, 47–56. DOI: 10.1007/s00216-009-2909-3
- Shepelev, M.V., Kalinichenko, S.V., Deikin, A.V., Korobko, I.V. (2018). Production of recombinant proteins from milk of transgenic animals: current state and prospects. *Acta Naturae*, 10(3), 42-50. (In Russian)
- Tan, W.S., Carlson, D.F., Walton, M.W., Fahrenkrug, S.C., Hackett, P.B. (2012). Precision editing of large animal genomes. *Advances in Genetics*, 80, 37-97. Doi: 10.1016/B978-0-12-404742-6.00002-8
- Trubitsina, T.P., Ryabykh, V.P., Koloskova, E.M., Ezerskiy, V.A., Maksimenko, S.V. (2018). Using the  $\beta$ -lactoglobulin gene for recombinant proteins from old transgenesis technologies to new genome editing methods (review). *Problems of Productive Animal Biology*, 3, 15-34. (In Russian)
- Whitelaw, C.B.A., Sheets, T.P., Lillico, S.G., Telugu, B.P. (2016). Engineering large animal models of human disease. *The Journal of Pathology*, 238, 247-256. Doi:10.1002/path.4648
- Wright, G., Carver, A., Cottom, D., Reeves, D., Scott, A., Simons, P., Wilmut, I., Garner, I., Colman, A. (1991). High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Biotechnology (NY)*, 9, 830-834. Doi: 10.1073/pnas.87.13.5178

## Citation:

Koloskova, E.M., Ezerskiy, V.A., Ostrenko, K.S. (2020). Designing of a DNA matrix for integration of a transgene into the bovine beta-lactoglobulin gene locus using CRISPR/Cas9 technology. *Ukrainian Journal of Ecology*, 10(5), 206-210.



This work is licensed under a Creative Commons Attribution 4.0. License