Daria K. SZALAST[1], Inga KOWALEWSKA[1] ⓘD, Ewa CZERNIAWSKA-PIĄTKOWSKA[2] ⓘD

ANALYSIS OF POLYMORPHISM IN THE *NUCB2* GENE AND THE MILK PRODUCTION TRAITS OF POLISH HOLSTEIN-FRIESIAN CATTLE

[1] Department of Genetics, West Pomeranian University of Technology, Szczecin, Poland
[2] Department of Ruminant Science, West Pomeranian University of Technology, Szczecin, Poland

Abstract. The aim of the study was to analyze the single nucleotide substitution polymorphism (SNP) located in exon 9 (1079 *G>T, rs471862956*) of the gene encoding nucleobindin-2 (NUCB2) in relation to milk production traits such as milk, protein and fat yield and protein and fat content in milk. The research was carried out in a herd of Polish Holstein-Friesian black and white cattle, and genotyping of individual animals was carried out using PCR-RFLP. The most common genotype is the heterozygous genotype (0.53), and the more frequent allele is the *T* allele (0.52). The analysis of the obtained results showed that there are statistically significant relationships at the level of P ≤ 0.05 and P ≤ 0.01 between individual genotypes of the polymorphism studied and the considered milk performance characteristics. It was observed that individuals with the *TT* homozygous genotype were characterized by the highest values of such traits as: milk yield as well as fat and protein yield. However, in the case of animals of the homozygous *GG* genotype, it was observed that the milk obtained from them was characterized by the highest fat and protein content.

Key words: *NUCB2*, polymorphism, dairy cattle, milk production characteristics.

INTRODUCTION

Nucleobindins are a class of calcium ion binding multidomain proteins that interact with nucleic acids and a variety of regulatory proteins and thus play an important role in the functioning of various signaling pathways. Originally, nucleobindin was classified as a transcription factor because it showed the ability to bind to DNA fragments in vitro. It is now known that these proteins have a broader spectrum of action and are distinguished into two nucleobindins: nucleobindin-1 (NUCB1) and nucleobindin-2 (NUCB2). These proteins are encoded by two independent genes, although the level of homology with the human amino acid sequence for *NUCB1* and *NUCB2* is 62% (Gonzalez et al. 2012). Furthermore, in 2006 (Oh-I et al. 2006) the nesfatin-1 protein was identified as the 82 amino acid nucleobindin 2 cleavage product in the *NUCB2* sequence. Nesfatin-1 is the *N*-terminal fragment of NUCB2. Its biological activity is related with inhibition of nocturnal food intake with free access to food, after peripheral and central administration into the ventricles of the brain (Oh-I et al. 2006). Due to the observed biological activity, the name of this protein was created, because the word nesfatin-1 is an acronym for the phrase "NUCB2 – encoded satiety and fat

influencing protein" (Shimizu et al. 2009). In another study, Li et al. (2010) indicate that nesfatin-1 is considered to be a factor in the regulation of appetite and is considered to be one of the influencing factors on nutritional behavior through its effects on melanocortin in the hypothalamus. It has been observed that the expression of nesfatin-1 is decreased in the fasting period. Additionally, central treatment with nesfatin-1 can suppress food intake, and this protein can block increased food intake. Based on these findings, NUCB2 was found to be associated with eating habits and effects on the appetite process (Li et al. 2010).

The gene encoding NUCB2 in cattle has been mapped on chromosome 15 (chr. 15: 43.5418), consists of 12 exons, including 11 coding ones, and 11 introns; it is 1466 base pairs long (https://www.ensembl.org). In human organisms, NUCB2 has been shown to be involved in the proliferation and migration of mammary epithelial cells, and its expression is reduced with decreased milk production, meaning that *NUCB2* expression may be beneficial for increasing milk production. Moreover, in terms of mere localization of the gene in cattle, it is 0.86 and 2.54 cM from the top of the reported regions of the quantitative trait (QTL) locus, which has a significant impact on the percentage of protein content and its yield in the milk produced. *NUCB2* is also in close proximity (0.93–5.99 Mb) to three important SNPs for dairy performance traits (BFGL-NGS-116109, BTB-00590405 and BTB-00590603), which indicates the possible involvement of the gene in milk production (Han et al. 2019).

The purpose of the study was to analyze the polymorphism mapped within the exon 9 – 1079 *G>T* (*rs471862956*), which is a missense mutation causing the replacement of tryptophan with leucin (W347L) and, as a result, to estimate the frequency of genotypes and alleles of the polymorphism and to determine the possible association of individual genotypes with production traits of Holstein-Friesian dairy cattle.

MATERIALS AND METHODS

For the study there were used 279 individuals of the Holstein-Friesian black and white variant kept in Western Pomeranian region. All of them were kept in a tether system under similar environmental conditions. They were fed as standard and in the spring and summer seasons they were given free access to the pasture. The milk yield in the herd was assessed using the A4 method. Peripheral whole blood was collected from each subject into vacuum tubes.

DNA isolation was performed using a DNA isolation kit (MasterPure DNA Purification Kit for Blood, Lucigen) according to the isolation protocol provided with the kit. Individual genotypes were determined using the PCR-RFLP method. The SNP polymorphism located in exon 9 1079 *G>T* (*rs471862956*) was analyzed. The PCR reaction was carried out using primers prepared on the basis of the sequences obtained from the Ensembl database. The starter sequences were as follows: forward 5'-TCAGTTTCTATGGATGTTAGGCT-3' and reverse: 3'-CCA-CAGCTTTTCAAGAAATCGAG-5'. The amplification reaction was conducted following to a standard thermal profile and the primer annealing temperature was 50°C. A ready-to-use Mix Plus Green PCR mixture (A&A Biotechnology) was used to perform the PCR. The 390 bp PCR products were assessed electrophoretically using an agarose electrophoresis kit. The obtained amplification products were then digested with the restriction enzyme *Bse*YI.

The obtained genotyping results were analyzed statistically. In the first stage, the frequency of the occurrence of particular genotypes and alleles in the studied cattle herd was estimated; then the correlations between genotypes and the following milk performance characteristic were analyzed: milk yield (kg), protein yield (kg), fat yield (kg), protein content (%) and fat content (%). The data on the performance of cows was obtained according to the basis of the breeding documentation of the farm kept by the Polish Federation of Cattle Breeders and Milk Producers (Polska Federacja Hodowców Bydła i Producentów, PFHBiPM). The statistical analysis of the

relationship between the 1079 *G>T* polymorphism and the milk production traits was analyzed using the Statistica 13.3 PL program. Mean values and standard deviation were calculated and a one-way analysis of variance was achieved using Duncan's multiple range test.

RESULTS

Restriction analysis of the 390 bp *NUCB2* gene fragment treated with the restriction enzyme *Bse*YI showed the presence of three genotypes existing in the studied population: *TT* (390 bp, no restriction site), *GG* (218, 172 bp) and *GT* (390, 218, 172 bp); these genotypes are determined by two alleles: *G* and *T*. The analysis of the genetic structure has shown that the *GT* genotype is the most common genotype, the *GG* and *TT* genotypes were at a fairly similar level. The frequency of the genotypes and alleles is presented in Table 1.

Table 1. Frequency of genotypes and alleles of the 1079 *G>T* polymorphism under study in the *NUCB2* gene

Parameter	Genotype			Alleles	
	GG	*GT*	*TT*	*G*	*T*
Frequency	0.21	0.53	0.26	0.48	0.52
Number	59	147	73	–	–

The analysis of the obtained results has shown that there are statistically significant associations at the level of P ≤ 0.05 and P ≤ 0.01 between individual genotypes of the studied polymorphism and the considered milk performance characteristics. Table 2 shows the mean values and standard deviation for the examined traits.

Table 2. Average values and standard deviations in the studied milk production traits

Lactation	Genotype	Milk (kg)	Fat (kg)	Protein (kg)	Fat (%)	Protein (%)
I	*GG*	8433.69 ± 1326.56[a]	336.76 ± 52.28[Aa]	288.59 ± 44.51[A]	4.01 ± 0.46[a]	3.44 ± 0.20[a]
	GT	9430.32 ± 1018.18[b]	356.55 ± 52.64[b]	315.53 ± 40.51[B]	3.85 ± 0.40[b]	3.40 ± 0.18[a]
	TT	10122.7 ± 1005.1[c]	374.45 ± 47.16[Bc]	337.29 ± 38.57[C]	3.73 ± 0.33[b]	3.32 ± 0.18[b]
II	*GG*	9356.52 ± 1207.53[a]	368.33 ± 56.82[a]	315.08 ± 43.63[a]	4.06 ± 0.52[a]	3.49 ± 0.22[a]
	GT	10764.41 ± 1615.09[b]	407.54 ± 67.79[b]	359.77 ± 40.74[b]	3.86 ± 0.46[b]	3.41 ± 0.20[a]
	TT	12131.55 ± 1502.07[c]	445.25 ± 71.72[c]	391.10 ± 60.57[c]	3.74 ± 0.45[b]	3.32 ± 0.22[b]
III	*GG*	9069.51 ± 694.88[A]	372.51 ± 53.92[A]	317.69 ± 34.16[A]	4.07 ± 0.47[a]	3.46 ± 0.23[a]
	GT	11019.22 ± 589.80[B]	423.47 ± 58.93[B]	374.58 ± 30.22[B]	3.91 ± 0.45[a]	3.42 ± 0.18[a]
	TT	13031.12 ± 864.02[C]	494.88 ± 65.81[C]	429.97 ± 31.95[C]	3.80 ± 0.49[b]	3.31 ± 0.19[b]

[A, B, C] P ≤ 0.01; [a, b, c] P ≤ 0.05.

The analysis of the obtained results shows that the *TT* homozygous cows were determined to have the highest average milk yield in all three tested lactations. These differences were confirmed statistically – for lactation I and II the significance level was P ≤ 0.05, and for lactation III this level was P ≤ 0.01. The next analyzed milk performance trait was the efficiency of the produced milk protein – the highest average value of this trait was present in animals with the *TT* genotype also in all three consecutive lactations.

The observed differences in protein efficiency between cows of particular genotypes turned out to be statistically significant for lactation II at the level of P ≤ 0.05, and for lactation I and III at the level of P ≤ 0.01. As for the fat yield in the produced milk, the tendency was similar and cows with the *TT* homozygous genotype were determined by the highest value of this parameter in all

three lactations. In the case of this considered feature, the differences between the mean values for cows of particular genotypes also turned out to be statistically significant – for lactation II at the level of P ≤ 0.05, and for lactation III at the level of P ≤ 0.01. However, for lactation I, it was observed that the fat yield in milk produced by *TT* cows differed significantly (P ≤ 0.01) from the fat yield obtained from animals with the homozygous *GG* genotype. Statistical analysis of the protein content in milk showed that cows with the *TT* homozygous genotype produced milk with a significantly (P ≤ 0.05) lower protein content than the cows with the other genotypes. Furthermore it can be concluded that cows with the homozygous *GG* genotype produced milk with the highest fat content compared to those of the other genotypes. In the case of lactation I and II, this difference turned out to be statistically significant (P ≤ 0.05).

DISCUSSION

Milk and its products are one of the most important foods in the human diet, especially when it comes to functional foods. Milk contains nearly 1,000 ingredients, including over 400 fatty acids, many important proteins, sugars, salts and minerals, and contributes to the creation of nearly 2,000 bioactive peptides (Szulc and Barłowska 2020). Therefore, it is important to constantly gather information which holds the potential to have a significant impact on the further improvement of the dairy industry, and furthermore it is worth relying on genetic factors and extending research on the occurrence of differences in the genetic material and their connection with breeding performance, in order to be able to consolidate favorable polymorphisms in populations, and, consequently, increase milk yield even more, while maintaining animal welfare.

The individual traits of dairy performance are controlled by multiple genes. By identifying candidate genes related to production traits, it is possible to obtain information that can be used to increase the accuracy of selecting animals in accordance to moderately heritable traits such as milk production. Genomic selection can help increase the accuracy and intensity of selection and shorten the generation gap (Ma et al. 2021).

Polymorphism in the *NUCB2* gene in cattle was analyzed in relation to the milk performance traits and growth parameters (Li et al. 2010; Han et al. 2019). As for the work focusing on the growth parameters (Li et al. 2010), the research was carried out in herds of native Chinese cattle (Qinchuan, Jiaxian Red, Nanyang) and it analyzed two polymorphisms mapped in exon 9, which were synonymous mutations. It was shown that the tested SNPs influenced all analyzed parameters (body length, weight, average daily weight gain). On the other hand, Han et al. (2019) focused on the analysis of polymorphisms mapped mainly in the 3' flanking region with regard to dairy performance traits such as: milk, protein and fat yield, and milk protein and fat content in Chinese Holstein cattle herds. All polymorphisms considered in the study were associated with at least one milk yield trait during the first or second lactation.

Lents et al. (2013), in a study conducted on pigs, has shown that the *NUCB2* gene participates in the hypothalamic pathways regulating feed intake and energy homeostasis, acting on leptin. As has long been known, leptin is one of the factors influencing both nutritional behavior and fat accumulation, and plays a key role in regulating the energy homeostasis of the whole organism (Komisarek et al. 2005). Therefore, it seems justified to study mapped polymorphisms or search for new SNPs in the *NUCB2* gene in the context of improving the dairy performance characteristics of cattle. Our own study has proven statistically significant relationships between particular genotypes and milk performance traits in Holstein-Friesian black and white cattle. Subjects with the *TT* genotype were defined by the highest milk, protein and fat yield as well as milk protein content, and the observed differences were statistically significant, as it has been statistically confirmed in the previously mentioned study by Han et al. (2019).

CONCLUSIONS

The results of the analysis of the genetic structure in the studied herd of Polish Holstein-Friesian black and white cattle indicate that the most common were the heterozygous genotype and the *T* allele. Summarizing the obtained results of the statistical analysis, statistically significant (P ≤ 0.01, P ≤ 0.05) correlations between individual genotypes and the tested milk performance characteristics were found. It was proven that individuals with the *TT* homozygous genotype were defined by a tendency to achieve the highest values of the tested traits, such as milk yield, fat and protein yield. However, as for the animals of the homozygous *GG* genotype, it was observed that the milk obtained from them has the highest fat and protein content. In order for the results obtained in the study above to be used in the breeding of dairy cows, further research on the influence of individual genotypes of the 1079 *G>T* polymorphism (*rs471862956*) in the *NUCB2* gene on the characteristics of dairy performance in various cattle breeds and on more numerous herds should be carried out.

REFERENCES

Ensembl, https://www.ensembl.org/Bos_taurus/Transcript/Summary?db=core;g=ENSBTAG00000017468;r=15:35156660-35179353;t=ENSBTAT00000023221, access: 20.06.2022.

Gonzalez R., Mohan H., Unniappan S. 2012. Nucleobindins: Bioactive precursor proteins encoding putative endocrine factors? Gen. Comp. Endocrinol. 176, 341–346.

Han B., Yuan Y., Li Y., Liu L., Sun D. 2019. Single nucleotide polymorphism of NUCB2 and their genetic associations with milk production traits in dairy cows. Genes 10, 449.

Komisarek J., Szyda J., Michalak A., Dorynek Z. 2005. Impact of leptin gene polymorphisms on breeding value for milk production traits in cattle. J. Anim. Feed Sci. 14, 491–500.

Lents C.A., Barb C.R., Hausman G.J., Nonneman D., Heidorn N.L., Cisse R.S., Azain M.J. 2013. Effects of nesfatin-1 on food intake and LH secretion in prepubertal gilts and genomic association of the porcine NUCB2 gene with growth traits. Domest. Anim. Endocrinol. 35, 89–97.

Li F., Chen H., Lei C.Z., Ren G., Wang J., Li Z.J., Wang J.Q. 2010. Novel SNPs of the bovine NUCB2 gene and their association with growth traits in three native Chinese cattle breeds. Mol. Biol. Rep. 37, 541–546.

Ma Y., Khan M.Z., Xiao J., Alugongo G.M., Chen X., Chen T., Liu S., He Z., Wang J., Shah M.K., Cao Z. 2021. Genetic markers associated with milk production traits in dairy cattle. Agriculture 11, 1018. DOI: 10.3390/ agriculture11101018.

Oh-I S., Shimizu H., Satoh T., Okada S., Adachi S., Inoue K., Eguchi H., Yamamoto M., Imaki T., Hashimoto K., Tsuchiya T., Monden T., Horiguchi K., Yamada M., Mori M. 2006. Identification of nesfatin 1 as a satiety molecule in the hypothalamus. Nature 443, 709–712.

Shimizu H., Oh-I S., Okada S., Mori M. 2009. Nesfatin-1: An overview and future clinical application. Endocr. J. 56, 4, 537–554.

Szulc T., Barłowska J. 2020. Tajemnice mleka [Secrets of milk]. Wrocław, Wydawnictwo Uniwersytetu Przyrodniczego we Wrocławiu. [in Polish]

ANALIZA POLIMORFIZMU W GENIE *NUCB2* A CECHY PRODUKCJI MLECZNEJ BYDŁA RASY POLSKIEJ HOLSZTYŃSKO-FRYZYJSKIEJ

Streszczenie. Celem pracy była analiza polimorfizmu typu pojedyncze podstawienie nukleotydowe (SNP) zlokalizowanego w eksonie 9 (1079 *G>T*, *rs471862956*) genu kodującego nukleobindynę-2 (NUCB2) w odniesieniu do cech użytkowości mlecznej takich jak: wydajność mleka, białka i tłusz-

czu oraz zawartość białka i tłuszczu w mleku. Badania prowadzono w stadzie bydła rasy polskiej holsztyńsko-fryzyjskiej odmiany czarno-białej, a genotypowanie poszczególnych osobników prowadzono przy użyciu PCR-RFLP. Najczęściej występującym genotypem jest genotyp heterozygotyczny (0,53), a najczęściej występującym allelem jest allel *T* (0,52). Analiza otrzymanych wyników wykazała, że istnieją statystycznie istotne zależności na poziomie $P \leq 0,05$ oraz $P \leq 0,01$ między poszczególnymi genotypami badanego polimorfizmu a rozpatrywanymi cechami użytkowości mlecznej. Zaobserwowano, że osobniki o genotypie homozygotycznym *TT* charakteryzowały się najwyższymi wartościami takich cech jak wydajność mleka oraz wydajność tłuszczu i białka. Natomiast w przypadku zwierząt o genotypie homozygotycznym *GG* zaobserwowano, że otrzymane od nich mleko cechowało się najwyższą zawartością tłuszczu i białka.

Słowa kluczowe: *NUCB2*, polimorfizm, bydło mleczne, cechy produkcji mleka.