Scientific Annals of the Polish Society of Animal Production - Vol. 16 (2020), No 1, 9-16

DOI: 10.5604/01.3001.0014.0506

Effect of polymorphism in the 5' promoter region of the *TG* gene on the milk production traits of Polish Holstein-Friesian cattle

Ewa Czerniawska-Piątkowska[1], Inga Kowalewska-Łuczak[2#]

West Pomeranian University of Technology in Szczecin,
Faculty of Biotechnology and Animal Husbandry,
[1]Department of Ruminant Science,
ul. Klemensa Janickiego 29, 71-270 Szczecin;
[2]Department of Genetics,
Aleja Piastów 45, 70-311 Szczecin, Poland; e-mail: inga.kowalewska-luczak@zut.edu.pl

The aim of the study was to determine the relationship between the occurrence of three different point mutations (257C>T, 335A>G and 422 C>T) in the promoter region of the gene encoding thyroglobulin (causing polymorphism) and the milk production traits of cattle. The research was conducted on a herd of Polish Holstein-Friesian cows. Genotypes were determined by PCR and its ACRS modification. The frequency of more common alleles for the point mutations was determined: *C* – 0.830 (257C>T), *A* – 0.765 (335A>G), and *C* – 0.635 (422C>T). The statistical analysis showed that all three point mutations in the promoter region of the *TG* gene significantly ($P \leq 0.05$; $P \leq 0.01$) affect the analysed milk production traits, i.e. daily milk yield, content of milk fat and protein, and somatic cell count.

KEY WORDS: thyroglobulin, cattle, SNP, milk production

Thyroglobulin (TG) is a glycoprotein homodimer that is mainly produced by thyroid follicular cells. Its primary role is as a substrate in the synthesis of thyroxine (T4) and triiodothyronine (T3). Thyroglobulin is secreted from the endoplasmic reticulum to a site where it undergoes iodination (coupling of iodine to tyrosine residues of thyroglobulin). A major portion of it is stored inside the thyroid follicles, but is also located within the colloid, i.e. fluid that fills the thyroid follicles (van der Spek et al., 2017). When thyroid cells are stimulated by TSH (thyroid-stimulating hormone), thyroid hormones ready to act are released. Thyroid hormones affect many processes in the body, including control of metabolic processes. They stimulate gluconeogenesis and lipogenesis in the liver, as well as glycogenolysis. Thyroxine is involved in various metabolic transformations, but its most important role is in the conversion of triiodothyronine. This in turn has a major impact on

the functioning of the nervous system, as well as on the development and growth of the body (Mullur et al., 2014).

The thyroglobulin coding gene (*TG*) is located in the centromeric region of chromosome 14 (BTA14). It has 37 exons separated by introns and at least 300 kb of genomic DNA (Khatib et al., 2007). The *TG* gene is located in a QTL region, and because the product of this gene is a precursor of a hormone that indirectly affects lipid metabolism, Thaller et al. (2003) considered it a candidate gene, due to the function of its product. Various studies have shown that point mutations located in the 5' promoter region of the *TG* gene are associated with meat production traits in cattle, such as fat content in the longissimus dorsi muscle (Anton et al., 2008, 2011) and the level of marbling in meat (Shin and Chung, 2007; Gan et al., 2008). Because a relationship had been demonstrated between meat quality traits (mainly fat content) and the location of the *TG* gene in the BTA14 QTL region associated with milk yield and fat content, Khatkar et al. (2004) and Khatib et al. (2007) concluded that it was worth examining the relationship between point mutations in the 5' promoter region and milk production traits.

The aim of the study was to estimate the frequency of genotypes and alleles of three substitutions – transitions (257 C>T, 335 A>G and 422 C>T) located in the 5' promoter region of the *TG* gene and to investigate the possibility of a link between individual genotypes and selected milk production traits of Polish Holstein-Friesian cattle.

Material and methods

The study was conducted in a herd of 100 Polish Holstein-Friesian cows on a farm in the West Pomeranian Voivodeship. The cows were kept in a free-stall system and fed TMR (total mixed rations). The beds were covered with straw. The feed ration was prepared using a Santrak 4.0 Selfline self-propelled feed mixer. The cows were milked twice a day using a mechanical milking machine. The milk yield of the herd was assessed by the A4 method in accordance with the recommendations of the International Committee for Animal Recording (ICAR). Daily milk yield, milk fat and protein content, and somatic cell count were analysed. As the somatic cell count is highly variable and does not have a normal distribution, the actual somatic cell count was transformed to a natural logarithm to meet the conditions of normal distribution.

The biological material was whole peripheral blood (5 ml) collected in vacuum tubes containing K_3EDTA as an anticoagulant. DNA was isolated from the blood using the MasterPure™ DNA isolation kit (Epicenter®) according to the included isolation protocol. The genotype of each cow was determined using PCR-RFLP. Three single nucleotide variants (SNVs) in the 5' promoter region of the *TG* gene, described by Shin and Chung (2007), were analysed. Three primer sequences (two forward

primers and one reverse primer) were used in specific combinations to amplify selected *TG* fragments. The primer sequences used to determine the genotypes were designed using Primer 3 software (Untergasser et al., 2012) based on gene sequences available in the Ensembl database (https://www.ensembl.org). In two cases (both forward primers) the ACRS (artificially created restriction site) method was used to create an enzyme cleavage site. Detailed information on the SNPs, i.e. location, primer sequences, elongation temperature, restriction enzymes, PCR product sizes, and the size of fragments after enzyme digestion are given in Table 1.

Table 1
PCR-RFLP conditions for the analysed polymorphisms in the *TG* gene

SNP	Primer sequence (5′ – 3′)	AT	AS (bp)	RE	RFLP (bp)
257 C>T *rs207642293*	F: GAGGGAGCATTGTGTTT**G**T R: CTGTTTTCTCTGCTGGTCAC	50°C	440	*Alw*26I	*C*: 256, 158, 26 *T*: 282,158
335A>G *rs523209078*	F: TGAAGATGAATTATGAAGCC**G**C R: CTGTTTTCTCTGCTGGTCAC	50°C	363	*Hha*I	*A*: 363 *G*: 339, 24
422C>T *rs135751032*	F: GAGGGAGCATTGTGTTT**G**T R: CTGTTTTCTCTGCTGGTCAC	50°C	440	*Psu*I	*C*: 256, 184 *T*: 440

AT – annealing temperature; AS – amplicon size, RE – restriction enzyme, mismatched nucleotides are underlined; RFLP – digestion product size

PCR was carried out in mixtures with a final volume of 25 µl containing the forward primer (1.0 µl), reverse primer (1.0 µl), standard ready-to-use 2xPCR Mix (12.5 µl), DNA (2 µl) and nuclease-free water (8.5 µl). The PCR steps were carried out under the following temperature conditions: initial denaturation at 94°C for 5 minutes, denaturation at 94°C for 45 seconds, hybridization at 50°C for 45 seconds, elongation at 72°C for 45 seconds (repeated in 30 cycles) and final elongation at 72°C for 5 minutes.

The time and temperature of restriction enzyme digestion of the PCR product was in accordance with the manufacturer's instructions. The resulting fragments were separated by horizontal electrophoresis on 2% agarose gels stained with ethidium bromide and visualized under a UV transilluminator.

Genotype and allele frequencies were calculated, and polymorphism information content (PIC) was estimated according to Nei and Roychoudhury (1974).

A statistical analysis was performed of the relationship between individual genotypes and selected milk production traits: daily milk yield, milk fat and protein content, and somatic cell count. The statistical analysis between genetic variants of the

gene encoding thyroglobulin and production traits was carried out using STATISTICA®12.0 software (Statsoft Inc., 2014) using the following model of one-way analysis of variance:

$$Y_{ij} = \mu + t_i + e_{ij}$$

where:

μ – expected value
t_i – analysed variable
e_{ij} – sampling error

Results and discussion

Table 2 shows the frequency of genotypes and alleles for individual transitions in the 5' region of the *TG* gene. The presence of all three possible genotypes was identified for all three point mutations analysed. Analysis of the results indicates that for the 257 C>T transition the highest frequency was shown for the *CC* genotype and for the *C* allele, while for the 335 A>G transition the homozygous *AA* genotype and the *A* allele were most common. In the case of the 422 C>T transition, the heterozygous genotype and the *C* allele were most common.

Table 2
Genotype and allele frequencies (n – number of cows)

Polymorphism	n	Genotype frequencies		Allele frequencies		PIC
257C>T	70	*CC*	0.70	*C*	0.830	0.242
	26	*CT*	0.26	*T*	0.170	
	4	*TT*	0.04			
335A>G	59	*AA*	0.59	*A*	0.765	0.295
	35	*AG*	0.35	*G*	0.235	
	6	*GG*	0.06			
422C>T	40	*CC*	0.40	*C*	0.635	0.356
	47	*CT*	0.47	*T*	0.365	
	13	*TT*	0.13			

The 422 C>T transition is the most commonly studied single nucleotide variant in the gene encoding thyroglobulin in cattle. The frequency of the *C* allele reported by other authors is highly varied: from 0.51 in Angus cattle to 0.98 in Hereford cattle (Pannier et al., 2010), most often ranging from 0.65 to 0.87 (Khatib et al., 2007; Anton et al., 2008, 2012). On the other hand, the 257 C>T and 335 A>G transitions have previously been described only by Shin and Chung (2007), but the authors did not specify the allele or genotype frequency for these mutations.

The PIC assessment according to Botstein et al. (1980), where polymorphism is considered low when the PIC value is less than 0.25, intermediate when PIC is 0.25-0.5, and high when it is greater than 0.5, showed intermediate polymorphism for all three polymorphic sites.

Table 3 presents the results of the statistical analysis of individual genotypes in relation to selected milk production traits: daily milk yield (kg), milk fat content (%), milk protein content (%), and somatic cell count.

Table 3
Mean values and standard deviations for milk production traits in relation to *TG* genotypes

Genotype	Milk yield (kg)	Fat content (%)	Protein content (%)	LnSCC*
		257C>T		
CC	18.68 ±6.17^{a}	3.95 ±0.59^{A}	3.36 ±0.39^{a}	4.12 ±0.79^{Ac}
CT	16.56 ±6.08^{b}	4.71 ±0.68^{B}	3.75 ±0.46^{b}	4.36 ±0.84^{Bd}
TT	18.61 ±5.62^{c}	5.25 ±0.82^{C}	3.66 ±0.30^{c}	3.67 ±0.74^{AB}
		335A>G		
AA	17.40 ±6.05^{A}	4.46 ±0.70^{A}	3.57 ±0.47^{Ac}	4.13 ±0.85^{a}
AG	19.41 ±6.17^{Bc}	3.87 ±0.60^{B}	3.35 ±0.35^{Bd}	4.19 ±0.77
GG	17.94 ±6.45^{d}	3.48 ±0.59^{C}	3.23 ±0.39^{B}	4.37 ±0.75^{b}
		422C>T		
CC	19.59 ±6.22^{A}	3.78 ±0.57^{A}	3.30 ±0.36^{A}	4.21 ±0.78
CT	17.03 ±5.96^{Bc}	4.33 ±0.60^{B}	3.53 ±0.43^{B}	4.13 ±0.82
TT	17.64 ±6.02^{Bd}	4.98 ±0.82^{C}	3.80 ±0.47^{C}	4.14 ±0.91

*LnSCC – natural logarithm of somatic cell count
a, b – values in columns with different lowercase letters differ significantly (P≤0.05)
A, B – values in columns with different capital letters differ highly significantly (P≤0.01)

Analysis of the results for the 257 C>T transition shows that cows with the heterozygous genotype had the lowest daily milk yield, the highest milk protein content, and the highest somatic cell count in the milk. Animals with the homozygous *CC* genotype had the lowest milk protein and fat content, while cows with the homozygous *TT* genotype had the highest milk fat content and the lowest somatic cell count. All these relationships were confirmed statistically ($P \leq 0.05$; $P \leq 0.01$).

Statistically significant relationships ($P \leq 0.05$; $P \leq 0.01$) between the genotypes and milk production traits were also found in the case of the 335 A>G transition. Milk obtained from individuals with the *AA* homozygous genotype was shown to have the highest fat and protein content and the lowest somatic cell count. Cows with heterozygous genotypes

produced the most milk, while those with homozygous *GG* genotypes produced milk with the lowest fat content and the highest somatic cell count.

For the 422 C>T transitions, cows with a homozygous CC genotype were observed to have the highest daily milk yield, the lowest milk fat and protein content, and the highest somatic cell count in the milk. Cows with a heterozygous genotype produced the least milk, while milk obtained from cows with homozygous *TT* genotypes had the highest fat and protein content. These relationships were confirmed statistically ($P \leq 0.05$; $P \leq 0.01$) except for somatic cell count.

There have been few studies analysing the relationship between polymorphism in the 5' promoter region and milk production traits, and these concern only the 422 C>T transition. Khatib et al. (2007) in a study on Holstein cattle and Kowalewska-Łuczak et al. (2010) in a study on Jersey cattle found no correlation between the 422 C>T transition and milk production traits. In contrast, Anton et al. (2012) studied three breeds of cattle, Holstein-Friesian, Jersey, and Hungarian Simmental, and showed that this point mutation affected milk yield in 305-day lactation, as well as milk fat content in the case of the Jersey breed.

Much more scientific research concerns point mutations in the *TG* gene in relation to meat production traits in cattle. The studies have largely focused on the 5' promoter region, particularly the 422 C> T transition. The authors showed that the mutation affects the level of meat marbling (Casas et al., 2007; Shin and Chung, 2007) and the fat content in the longissimus dorsi muscle (Thaller et al., 2003; Anton et al., 2008). Within the *TG* gene, the 3' flanking region (Gan et al., 2008; Hou et al., 2011) has also been studied in various breeds of cattle, and the substitutions analysed have been shown to significantly affect the level of meat marbling.

Livestock improvement focuses primarily on selective breeding of individuals with the desired phenotypes. The increasing amount of information available on the organization and functioning of the genome can be used in breeding programmes to improve a number of traits. Determination of the location of QTL genetic markers associated with the gene responsible for a given trait is often used in breeding programmes, but identifying the genes underlying control of variation provides stronger markers (Williams, 2005). The gene encoding thyroglobulin appears to be one that can be used in breeding programmes. It was indicated by Barendse (1999) as a positional candidate gene because its locus was mapped in the chromosome region associated with cattle performance traits and by Thaller et al. (2003) as a functional candidate gene due to the effect of the gene product on lipid metabolism. Numerous studies, mainly regarding meat traits, seem to confirm this.

Conclusions

Our research indicates that all three analysed transitions in the *TG* promoter region affect milk production traits, i.e. daily milk yield, milk fat and protein content, and somatic cell

count in the milk of Holstein-Friesian Polish cows, and these results were statistically confirmed ($P \leq 0.05$; $P \leq 0.01$). In the case of the 237 C>T transition, genotypes with an adverse effect on milk performance were identified, as cows with the heterozygous genotype had the lowest daily milk yield, while milk obtained from cows with the *CC* genotype had the lowest fat and protein content. For the other two transitions, the genotypes were shown to have a positive effect on milk production traits, i.e. daily milk yield (genotypes *AG* 335 A>G and *CC* 422 C>T), milk protein content (*AA* 335 A>G genotype), and milk fat content (genotype *TT* 422 C>T). Therefore, selection of dairy cattle should avoid animals with the *CT* genotype (237 C>T) and give preference to individuals with the *CC* (422 C>T) and *AG* (335 A>G) genotypes.

To sum up, the results of the research can be used in selection of dairy cattle and translate into economic benefits for milk producers and dairy farmers. Of course, these results should be regarded as preliminary due to the small number of animals included in the experiment. The next step should be to expand the research to include other cattle breeds and larger herds.

REFERENCES

Anton I., Kovács K., Fésüs L., Várhegyi J., Lehel L., Hajda Z., Polgár J.P., Szabó F., Zsolnai A. (2008). Effect of DGAT1 and TG gene polymorphisms on intramuscular fat and on milk production traits in different cattle breeds in Hungary. Acta Veterinaria Hungarica, 56 (2): 181–186 (DOI: 10.1556/AVet.56.2008.2.5).

Anton I., Kovács K., Holló G., Farkas V., Lehel L., Hajda Z., Zsolnai A. (2011). Effect of leptin, DGAT1 and TG gene polymorphisms on the intramuscular fat of Angus cattle in Hungary. Livestock Science, 135: 300–303 (DOI:10.1016/j.livsci.2010.07.012).

Anton I., Kovács K., Holló G., Farkas V., Szabó F., Egerszegi I., Rátky J., Zsolnai A., Brüssow K.P. (2012). Effect of DGAT1, leptin and TG gene polymorphisms on some milk production traits in different dairy cattle breeds in Hungary. Archiv Tierzucht, 55 (4): 307–314.

Barendse W.J. (1999). Assessing lipid metabolism. Patent, International Publication Number: WO 99/23248. World International Property Organization.

Botstein D., White R.L., Skolnik M., Davis R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. American Journal of Human Genetics, 32: 314–331.

Casas E., White S.N., Shackelford S.D., Wheeler T.L., Koohmaraie M., Bennett G.L. (2007). Assessing the association of single nucleotide polymorphisms at the thyroglobulin gene with carcass traits in beef cattle. Journal of Animal Science, 85: 2807–2814.

Gan Q.F., Zhang L.P., Li J.Y., Hou G.Y., Li D.H., Gao X. (2008). Association analysis of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. Journal of Applied Genetics, 49, 251–255.

Hou G.Y., Yuan Z.R., Zhou H.L., Zhang L.P., Li J.Y., Gao X., Wang D.J., Gao H.J., Xu S.Z. (2011). Association of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. Molecular Biology Reports, 38: 4705–4708 (DOI:10.1007/s11033-010-0605-1).

Khatib H., Zaitoun I., Chang Y.M., Maltecca C., Boettcher P. (2007). Evaluation of association between polymorphism within the thyroglobulin gene and milk production traits in dairy cattle. Journal of Animal Breeding and Genetics, 124: 26–28.

Khatkar M.S., Thomson P.C., Tammen I., Raadsma H.W. (2004). Quantitative trait loci mapping in dairy cattle: review and meta-analysis. Genetics Selection Evolution, 36: 163–190.

Kowalewska-Łuczak I., Kulig H., Szewczyk K. (2010). Polimorfizm w genie tyreoglobuliny u bydła rasy jersey. Acta Scientiarum Polonorum Zootechnica, 9 (4): 129–134.

Mullur R., Liu Y.Y., Brent G.A., (2014). Thyroid hormone regulation of metabolism. Physiological Reviews, 94: 355–382 (DOI:10.1152/physrev.00030.2013)

Nei M., Roychoudhury A.K. (1974). Sampling variances of heterozygosity and genetic distance. Genetics, 76: 379–390.

Pannier L., Mullen A.M., Hamill R.M., Stapleton P.C., Sweeney T. (2010). Association analysis of single nucleotide polymorphisms in DGAT1, TG and FABP4 genes and intramuscular fat in crossbred Bos taurus cattle. Meat Science, 85: 515–518.

Shin S.C., Chung E.R. (2007). Association of SNP Marker in the thyroglobulin gene with carcass and meat quality traits in Korean cattle. Asian-Australasian Journal of Animal Sciences, 20: 172-177.

Thaller G., Kuhn C., Winter A., Ewald G., Bellmann O., Wegner J. (2003). DGAT1, a new positional and functional candidate gene for intramuscular fat deposition in cattle. Animal Gene-tics, 34: 354–357.

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M., Rozen S.G. (2012). Primer3 – new capabilities and interfaces. Nucleic Acids Research 40(15):e115.

Van der Spek A.H., Fliers E., Boelen A. (2017). The classic pathways of thyroid hormone meta-bolism. Molecular and Cellular Endocrinology, 458: 29–38 (DOI.org/10.1016/j.mce.2017.01.025).

Williams J.L. (2005). The use of marker-assisted selection in animal breeding and biotechnology. Revue scientifique et technique – Office international des epizooties, 24 (1): 379–391.