POTENTIAL CONSEQUENCES OF *DGAT2* AND *BTN* GENES POLYMORPHISM IN IRAQI HOLSTEIN CATTLE

Mohammed Baqur S. Al-Shuhaib[1], Rawan A. Al-Fihan[1], Ali A. Al-Qutbi[2], Tahreer M. Al-Thuwaini[1]

[1]Al-Qasim Green University, College of Agriculture, Department of Animal Production, Babil, Iraq
[2]Al-Qasim Green University, College of Food Science, Department of Food Science and Technology, Babil, Iraq

This study was conducted to describe the genetic polymorphism within the bovine *diacylglycerol acyltransferase 2* (*DGAT2*) gene and 5′-UTR of *butyrophilin* (*BTN*) gene by several *in vitro* tools, as well as to predict their consequences using several *in silico* tools. After withdrawing blood, extracting DNA, and designing specific polymerase chain reaction (PCR) primers, single-stranded conformation polymorphism (SSCP) and PCR experiments were performed. Each different SSCP set was sequenced and analyzed. Two genotypes were detected with two alleles. Concerning *DGAT*2, the *AA* genotype was the most common, and seven single nucleotide polymorphisms (SNPs) were found in the *AB* genotype while four SNPs were found in the *AA* genotype. The only coding SNP (p.Ala7Val) was found in the *AB* genotype. Concerning *BTN*, the *AB* genotype was the most common, and twenty-two novel SNPs were detected in the *AB* genotype, while six SNPs were observed in the *AA* genotype. The effect of the observed coding and noncoding SNPs was computationally analyzed. It was shown that the coding SNP, p.Ala7Val, was predicted to have functional consequences on the DGAT2 protein structure and function. Several outcomes were deduced from several noncoding SNPs in both genetic fragments studied. This study determines the pattern of polymorphism within the *DGAT*2 and *BTN* genes to conceive their consequences within Holstein breed.

genotyping, milk synthesis, *in vitro*, *in silico*, SSCP

DE GRUYTER OPEN

doi: 10.1515/sab-2017-0020
Received for publication on December 30, 2016
Accepted for publication on April 4, 2017

INTRODUCTION

The proportion of milk constituents of livestock usually differs from one species to another, reflective of the needs of their respective offspring. Usually, these variations originate from the DNA level, which may contribute to the genetic characterization of livestock population. The genetic marker based molecular studies of Holstein breed have shown that these breeds are significantly differentiated at the molecular level (Felius et al., 2015). The implications of genetic tools in the Holstein cattle have been found to effectively produce a rapid genetic gain for the intended industrial goals (Zenger et al., 2006). Although the milk production traits are polygenic and all the genes affecting them are difficult to know, a few potential candidate genes have been recognized (Gupta et al., 2009). Several candidate markers were targeted in this regard, two of these candidate genes were *diacylglycerol acyltransferase* (*DGAT*) gene (Cases et al., 2001) and *butyrophilin* (*BTN*) gene (Mather, Jack, 1993). The *DGAT* gene is involved in triglyceride synthesis that influences dairy production in bovines. Two genes are known as *DGAT* (*DGAT*1 and *DGAT*2). The *DGAT*1 gene was the first identified gene encoding a protein with DGAT activity. The DGAT1 protein is an enzyme catalyzing the final step of triglyceride synthesis (Coleman et al., 2000). It was found that

a missense mutation (p. Lys232Ala) in *DGAT*1 gene has been significantly associated with variation in milk fat percentage in cattle (Winter et al., 2002). The *DGAT*1 gene became a promising candidate gene for milk fat percentage in cattle (Hradecká et al., 2008). Actually, the genetic variations in *DGAT*1 locus were found to be associated with a major effect on milk yield and composition (Grisart et al., 2004). However, DGAT-like activity has also been shown in other enzymes encoded by other genes and led to the detection of *diacylglycerol transferase* 2 (DGAT2) proteins. It is well demonstrated that DGAT2 protein plays an essential role in catalyzing the final step of the triacylglycerol (TAG) biosynthesis of the Kennedy pathway (Wakimoto et al., 2003). The *DGAT*2 gene is a closely related candidate for quantitative traits, and it is associated with lipid synthesis and storage in eukaryotes. Despite the fact that the *DGAT*2 gene may be of the same importance as *DGAT*1 gene in relation with milk production traits, most of polymorphism studies were focused on *DGAT*1 genetic polymorphism in correlation with milk production traits, such as (Spelman et al., 2002; Winter et al., 2002; Thaller et al., 2003; Grisart et al., 2004; Kuhn et al., 2004). On the other hand, the bovine *BTN* gene encodes an acidic transmembrane glycoprotein with a cytoplasmic C-terminal tail that is concentrated in the apical plasma membranes of mammary epithelial cells (Ogg et al., 2004). The BTN protein is secreted in association with the milk fat globule membrane from mammary epithelial cells. Through its location and specifications, it plays a crucial role in the regulation of milk fat globule secretion (Robenek et al., 2006). Since the BTN proteins are one of the several proteins that are involved in the milk fat globule formation, they indispensably participate in milk production (Heid, Keenan, 2005). Hence, both *DGAT*2 gene and *BTN* gene may be strong candidates for evaluation of genetic polymorphism. To unmask their potential polymorphism, several post-polymerase chain reaction (PCR) techniques were commonly employed, such as single-stranded conformation polymorphism (SSCP) (Orita et al., 1989). Its power potentially lies behind its likelihood of the detection of unknown mutation(s) (Gasser et al., 2006). Studies in cattle with respect to *DGAT2* gene are scarce, as the variations of *DGAT2* gene have been studied in very few domestic animals (Winter et al., 2003). Very little genetic data on this genetic fragment is currently available, at least in terms of its genetic variation in Holstein cattle. While, though several manuscripts have described the genetic variation of *BTN* gene in several cattle breeds, such as in American Jersey cattle (Komisarek et al., 2006), Indian crossbred cattle (Bhattacharya et al., 2006), or in Thailand Holstein cattle (Cheunsuk et al., 2010), the *BTN* genetic polymorphism has not been described yet in the middle east Holstein cattle. On the other hand, it was obviously demonstrated that many single nucleotide polymorphisms (SNPs) have functional effects on their corresponding protein structure either by a single change in the amino acid (Bahrami et al., 2013) or by modulating the transcription regulation rate (Liao, Lee, 2010), or by other reasons. By using bioinformatics prediction tools, it would be possible to discriminate neutral SNPs from SNPs of likely functional importance and it could also be useful to reveal the structural basis of the desired and undesired variations in milk characteristics. Thus, in addition to the common *in vitro* description, the employing *DGAT*2 and *BTN* genes in this population could be considered as interesting genetic markers to give a perspective *in silico* insight in the nature of the genotypes within this population. So, the main objective of this study is to detect the polymorphism of the *DGAT*2 and *BTN* genes using the *in vitro* genotyping PCR-SSCP-sequencing tools, as well as to predict the potential functional effect of both synonymous SNPs (sSNPs) and non-synonymous SNPs (nsSNPs) and to speculate on their possible influence on milk synthesis in the productionally important Holstein cattle breed using several common *in silico* computational tools.

MATERIAL AND METHODS

Blood sampling and DNA extraction

The study was conducted on a total of 60 individuals of Holstein cattle in Iraq. All of the studied population was reared in a private station in Al-Qadisya governorate. Both maintenance and feeding were similar for all animals and remained in accordance with the obligatory standards. About 3–4 ml of blood samples were collected from the jugular veins and placed in EDTA-anticoagulation tubes. Genomic DNA was isolated using a manual salting out method (Al-Shuhaib, 2017). The extracted DNA was evaluated by 0.8% agarose gel electrophoresis in 1X TAE (40 mM Tris acetate, 2 mM EDTA; pH 8.3), and quantified using a Nanodrop spectrophotometer (BioDrop μLITE; Biodrop, UK).

PCR design

Concerning both *DGAT*2 and *BTN* loci, three pairs of primers were designed using the NCBI primer BLAST online software (https://www.ncbi.nlm.nih.gov/tools/primer-blast), only the first and the third pairs were highly specific and reliable in our experiments (Table 1), while the PCR reaction of the second primer pairs turned out to be problematic because of the appearance of some PCR artifacts bands. The *DGAT*2 designed *BTN* primer pair in this study was also designed by Kale et al., 2013 using Primer3 software.

Table 1. Designed primers for *DGAT*2 and *BTN* genetic loci

GenBank Acc. No. and position	Sequence (5′-3′)	Length (bp)	Specificity	Amplified segment
AC_000172.1 (55940770–55940976)	CCAAGGCTCCGCTTTCCG TAGGCGGCTATGAGGGTCTT	207	highly specific	partially covering exon 1 and its 5′ UTR of *DGAT*2 gene
AF037402.1 (369–614)	AAAGCGAAGGACTGGTTGGT ACGTCTCCTGGAAGATCGG	246	non-specific (PCR artifacts)	partially covering 5′ untranslated region of *BTN* gene
AC_000180.1 (31362007–31362260)	CCTGCTTATTTCCCTAGTCTC CCACCCTAAGGTTAGTCAATC	254	highly specific	

PCR analysis

The PCR reaction was performed using the AccuPower PCR PreMix (Bioneer, South Korea). Each 20 µl of PCR premix contained 1 U of *Top* DNA polymerase, 250 µM of dNTPs, 10 mM of Tris-HCl (pH 9.0), 30 mM of KCl, and 1.5 mM of $MgCl_2$. The reaction mixture was completed with 10 pmol of each primer and 50 ng of genomic DNA. The optimum annealing temperatures were determined empirically in our extracted genomic DNA template using a gradient PCR thermocycler (version Mastercycler-nexus; Eppendorf, Germany). The PCR program was: initial denaturation at 94°C for 5 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 54,4°C for 30 s (for *DGAT*2 fragment) and 66°C (for *BTN* fragment), elongation at 72°C for 30 s, and final extension at 72°C for 5 min. After performing the PCR thermocycling, the PCR products were verified by electrophoresis on a 1.5% (w/v) agarose gel in 1X TBE buffer (2 mM of EDTA, 89 mM of Tris-Borate, pH 8.3), using a 100-bp ladder (Bioneer) as a molecular weight marker. Gels were stained with ethidium bromide (0.5 mg/ml) and visualized with a gel image documentation system (ChemiDoc; Bio-Rad, USA). All SSCP non-suitable amplicon bands were eliminated.

SSCP analysis

SSCP experiments were performed according to Al-Shuhaib's protocol with some optimizations (Al-Shuhaib, 2017). Several parameters have been optimized to avoid false positive results as well as to enhance sensitivity. Briefly, 7 µl of each amplification product was mixed with 7 µl of SSCP denaturing loading buffer (95% formamide, 20 mM EDTA pH 8, 0.05% xylene cyanol, and 0.05% bromophenol blue). The samples were heat-denatured at 95°C for 10 min and chilled on ice for at least 5 min. Then, PCR amplicons were separated in a vertical mini-gel format, gel size (W×L) 10 × 10 cm, and gel thickness 1 mm (OmniPAGE; Cleaver Scientific, UK). Denatured PCR products were loaded into the wells of 8% acrylamide/*bis* (37.5 : 1), containing 7% glycerol, and 1× TBE buffer. The gel was run at constant conditions (200 V/100 mA/140 min) at room temperature. Gels were stained by Red-PAGE gel dye (Cat # 41014; Biotium, Hayward, USA), and visualized with a gel image documentation system.

Statistical analysis

The genetic diversity statistical analysis was performed for each genetic fragment individually to estimate allele and genotype frequencies, as well as Nei's heterozygosity. These criteria were calculated with PopGene32 software, Version 1.31 (Yeh et al., 1999). A χ^2 test was performed to verify possible deviations from Hardy-Weinberg equilibrium (HW) expectations for the distribution of genotypes. Average heterozygosity was employed to estimate genetic diversity within the population.

DNA sequencing and sequencing analysis

Each of the varying patterns of the SSCP samples for both amplified *DGAT*2 and *BTN* gene fragments were purified and sequenced from both ends (Macrogen Inc. Geumchen, Seoul, South Korea). Only clear chromatographs obtained from ABI sequence files were further analyzed, ensuring that the annotation and variations are not because of the PCR or sequencing artifacts. The cattle reference sequences of both bovine *DGAT*2 and *BTN* loci, as well as other related details, were retrieved from their NCBI websites (https://www.ncbi.nlm.nih.gov/nuccore/aj534368 and https://www.ncbi.nlm.nih.gov/nuccore/af037402.1, respectively). Then, the sequenced genotypes were edited, aligned, and compared with their reference sequences using the *BioEdit* Sequence Alignment Editor Software, Version 7.1 (DNASTAR; Madison, USA). The sequencing results of the PCR products of different SSCP patterns of this study for both *BTN* and *DGAT*2 studied loci were submitted into the NCBI-BankIt server (https://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank) and got unique GenBank accession numbers (KY202553

and KY202554 for *BTN AB* and *AA* genotypes, and KY245897 and KY245898 for *DGAT*2 *AB* and *AA* genotypes, respectively).

Effect of *BTN* and *DGAT*2 SNPs within 5′UTR on transcription factor binding sites by PROMO

PROMO, Version 3.0.2, is a virtual laboratory for the identification of putative transcription factor binding sites (TFBS) in DNA sequences of a species or groups of species of interest (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). TFBS defined in the TRANSFAC database are used to construct specific binding site weight matrices for TFBS prediction. The result of this search can be inspected through a graphical interface and downloadable text files (Farré et al., 2003). Input data in both *DGAT*2 and *BTN* cases were the whole reference sequences as long as their *AB* and *AA* genotypes were submitted into PROMO server in FASTA format.

Prediction of noncoding SNPs effects in regulatory regions by UTR scan

UTRscan was used to predict the functional significance of noncoding SNPs in the UTR regions (Grillo et al., 2010). The program compares the query sequences (without and with nucleotide change) with the functional sequence patterns collection (located in 5′UTR and 3′UTR sequences) of the UTR site (http://itbtools.ba.itb.cnr.it/utrscan). If the UTR SNP has different functional patterns, this polymorphism is predicted to have functional significance.

Prediction of the nuclear receptor binding site for *BTN* and *DGAT*2 loci by NHR scan

NHR scan is a computational predictor of the nuclear hormone receptor binding sites web server (http://cisreg.cmmt.ubc.ca/cgi-bin/NHR-scan/nhr_scan.cgi). It can predict and classify potential nuclear hormone receptor binding sites in genomic sequences using Hidden Markov Models (HMM) derived from a large collection of verified binding sites (Sandelin, Wasserman, 2005). Essentially the NHR scan model consists of three 'match state chains' corresponding to each type of the site configuration (direct repeat (DR), inverted repeats (IR), and everted repeats (ER)), and one 'background state' corresponding to no prediction. Input data was submitted as FASTA format.

Three-dimensional structure construction of DGAT2 protein by PhyRe2 and PyMol

The two observed genotypes were translated into amino acids in a reading frame corresponding to the reference DGAT2 amino acid sequences using the Expasy online program (http://web.expasy.org/translate/). The primary structure designing of each SSCP genotype started with mutating the available reference NCBI DNA sequence of exon 1/*DGAT2* gene, by substituting each observed SNP into its reference sequences to represent *AA* and *AB* genotypes. The whole amino acid sequences of the bovine DGAT2 protein (Entry: Q70VZ8, EC: 2.3.1.20) were retrieved online from the protein data bank (http://www.uniprot.org/uniprot/Q70VZ8). Multiple amino acid sequences alignment was made between the reference exon 1/*DGAT2* amino acid sequence and its two genotypes observed using the Clustal Omega program from the UniProt website (http://www.uniprot.org/align/). The three-dimensional structure of the *DGAT2* gene was constructed from the online three-dimensional model prediction software Protein Homology/analogY Recognition Engine (Phyre2), Version 2.0 (Kelley et al., 2015). The virtually proposed changes within its corresponding mutants were performed by using PyMOL Version 1.7.0.1 software (https://www.schrodinger.com).

Finding the deleterious effect of DGAT2 protein nsSNP by SIFT

We submitted the amino acid sequence of DGAT2 protein along with non-synonymous SNPs (nsSNPs) with corresponding amino acid positions to SIFT (Sorting Intolerant From Tolerant) program (Pauline, Steven, 2003). The SIFT prediction was given as a tolerance index (TI) score ranging from 0.0 to 1.0, which was the normalized probability that the amino acid change was tolerated. Substitutions at each position with normalized probabilities less than a tolerance index of 0.05 are predicted to be intolerant or deleterious; those greater than or equal to 0.05 are predicted to be tolerated. The amino acid sequence of the native protein of interest along with nsSNPs with corresponding amino acid positions was submitted to SIFT program (http://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html).

Validation of the functional characterization of the DGAT2 protein nsSNP by PANTHER

The SIFT predicted nsSNPs were validated by the PANTHER (Protein Analysis Through Evolutionary Relationship) program (Tang, Thomas, 2016), which is an open access server that estimates the likelihood of a particular nsSNP to cause a functional impact on the protein (http:// www.pantherdb.org/tools/csnp). Data were submitted as the FASTA format protein sequence together with the mutational position of particular amino acid variant(s). PANTHER calculates the substitution position-specific evolutionary conservation (subPSEC) score based on an alignment of evolutionarily related proteins. The position-specific evolutionary preservation (PSEP) measures the length of time (in millions of years, my) a position in cur-

Table 2. Observed and expected heterozygosities, χ^2 test for Hardy-Weinberg equilibrium, and allelic frequencies of (A) *DGAT*2 5′-UTR/exon 1 fragment, and (B) *BTN* 5′-UTR fragment for the Holstein cattle breed

	Obs-Het	Exp-Het	Avr-Het	Nei's Exp-Het	χ^2	Allele *A* frequency	Allele *B* frequency
(A) *DGAT*2	0.3	0.257	0.255	0.255	1.752	0.85	0.15
(B) *BTN*	0.7333	0.4683	0.4644	0.4644	19.58	0.633	0.366

Obs-Het = observed heterozygosity, Exp-Het = expected heterozygosity, Avr-Het = average heterozygosity, Nei's Exp-Het = Nei's expected heterozygosity

rent protein has been preserved by tracing back to its reconstructed direct ancestors. The longer a position has been preserved, the higher is the functional impact a particular amino acid substitution is likely to have. The thresholds we chose were: 'probably damaging' (time > 450 my), 'possibly damaging' (450 my > time > 200 my, corresponding to a false positive rate of ~0.4), and 'probably benign' (time < 200 my).

Investigation of mutant protein stability of the DGAT2 protein nsSNP by I-Mutant 2.0

To have a better insight to the stability of the protein caused by mutation, the I-Mutant 2.0 tool (C a p r i o t t i et al., 2005) was used to analyze whether a single-point mutation may lead to several drastic changes. The effect of mutation may alter the stability of the protein and may lead to a change in the main characteristics of the milk components. The I-Mutant 2.0 (http://folding.biofold.org/cgi-bin/i-mutant2.0) is a Support Vector Machine-based web server for the automatic prediction of protein stability changes upon single-site mutations. The input FASTA sequence of protein along with the residues changes were provided for the analysis of the relative stability changes upon protein mutation (or DDG) value (kcal/mol). The reliability index (RI) value was computed as well.

Prediction of severity effect of the DGAT2 protein nsSNP by SNAP2

The SNAP2 tool is based on a machine learning device called 'neural network'. It distinguishes between effect and neutral variants/nonsynonymous SNPs by taking a variety of sequence and variant features into account (S m i g i e l s k i et al., 2000). The most important input signal for the prediction is the evolutionary information taken from an automatically generated multiple sequence alignment. SNAP2 revealed that if the nsSNP has scored more than zero (> 0) it has a damaging effect and vice versa. Predicting a score (ranges from −100 strong neutral prediction to +100 strong effect prediction) (https://www.predictprotein.org).

Prediction of functional impact of the DGAT2 protein nsSNP by PROVEAN

Further confirmation of the effect of nsSNPs on protein was done using the PROVEAN (Protein Variation Effect Analyzer) tool (C h o i et al., 2012) which can predict the impact of an amino acid substitution on the biological function of a protein (http://provean.jcvi.org/index.php). A query peptide sequence of the protein was provided in FASTA format to the PROVEAN server for predicting the functional impact of the SNPs. This algorithm allows for the best-balanced separation between the deleterious and neutral amino acids, based on a threshold. The default threshold is –2.5, i.e. variants with a score equal to or below –2.5 are considered 'deleterious', while variants with a score above –2.5 are considered 'neutral'.

RESULTS

The polymorphisms of both *DGAT2* and *BTN* genes were detected by PCR-SSCP and DNA sequencing methods, and the observed variations of SSCP gels indicate the detection of two different genotypes with two alleles (Table 2). The χ^2 test showed that the polymorphism of both *DGAT*2 and *BTN* genes in Holstein cattle was at Hardy-Weinberg equilibrium for these loci in the studied population. Nei's expected heterozygosity for both genes in this study has the same value of average heterozygosity. This indicates a high domination of allele *A* in both cases. Add to that, values of the observed heterozygosity for both *DGAT2* and *BTN* genotypes were higher than their expected values. This refers to the high level of genetic variability in the studied population.

***DGAT*2 genotyping**

Two genotypes (*AA* and *AB*) and two alleles (*A* and *B*) were observed (Fig. 1A). POPGENE 32 software showed that the observed frequencies of *DGAT*2 genotypes were 0.3 (n = 18) and 0.7 (n = 42) for *AB* and *AA* genotypes, respectively. The observed frequency of al-

Table 3. List of all single nucleotide polymorphisms observed in *DGAT*2 (A) and *BTN* (B) genetic fragments

Locus	Position in PCR amplicon	Position in reference gene	Nucleotide substitution	Type of SNP	Amino acid change
(A) *DGAT*2	129	55940898	C>T	synonymous	none
	130	55940899	A>C	synonymous	none
	135	55940904	C>T	synonymous	none
	145	55940914	G>C	synonymous	none
	168	55940937	C>T	synonymous	none
	171	55940940	G>C	synonymous	none
	179	55940948	T>C	synonymous	none
	204	55940973	C>T	non-synonymous	Ala>Val
(B) *BTN*	21	31362027	C>T	synonymous	none
	22	31362028	T>C	synonymous	none
	35	31362041	A>T	synonymous	none
	54	31362060	C>T	synonymous	none
	55	31362061	A>C	synonymous	none
	57	31362063	T>A	synonymous	none
	65	31362071	G>T	synonymous	none
	71	31362077	G>C	synonymous	none
	78	31362084	T>A	synonymous	none
	94	31362100	A>G	synonymous	none
	110	31362116	C>G	synonymous	none
	128	31362134	T>A	synonymous	none
	129	31362135	C>A	synonymous	none
	132	31362138	T>C	synonymous	none
	142	31362148	A>G	synonymous	none
	146	31362152	A>G	synonymous	none
	148	31362154	C>A	synonymous	none
	151	31362157	T>A	synonymous	none
	163	31362169	A>G	synonymous	none
	169	31362175	C>A	synonymous	none
	175	31362181	T>C	synonymous	none
	184	31362190	C>G	synonymous	none

PCR = polymerase chain reaction

lele *A* was 0.85, and of allele *B* it was 0.15 (Table 2, A). Sequencing results confirmed these two different electrophoretic SSCP patterns since several SNPs were detected among the two resolved genotypes and NCBI reference sequences (Fig. 1B). The pattern and nature of each SNP that was detected by sequencing indicated that the genotype *AB* has seven novel SNPs (C 129 T, A 130 C, C 135 T, G 145 C, G 171 C, T 179 C, C 204 T), while the genotype *AA* has four novel SNPs (C 129 T, A 130 C, G 145 C, C 168 T) (Table 3, A). Thus, both genotypes share three SNPs (C 129 T, A 130 C, G 145 C). The variations in the binding with the putative transcription factors were recorded for all 5′-UTR regions of the *DGAT*2 submitted sequences using the Promo server. The variability rate of the binding with the putative transcription factors found in the *AB* genotype is higher than in the *AA* genotype in comparison with their reference sequences (Table 4, A). This is usually attributed to the high number of the observed SNPs found in the *AB* genotype in this 5′-UTR region. However, the Promo results were not confirmed by the UTRscan server. The later program has discovered the presence of two regulatory elements (IRES and uORF), and their positions have never been exposed to any change and retained as they are in all submitted sequences (reference, *AB* genotype, and *AA* genotype sequences) (Table 5, A). Conversely, when we analyzed the effect of the observed SNPs on the

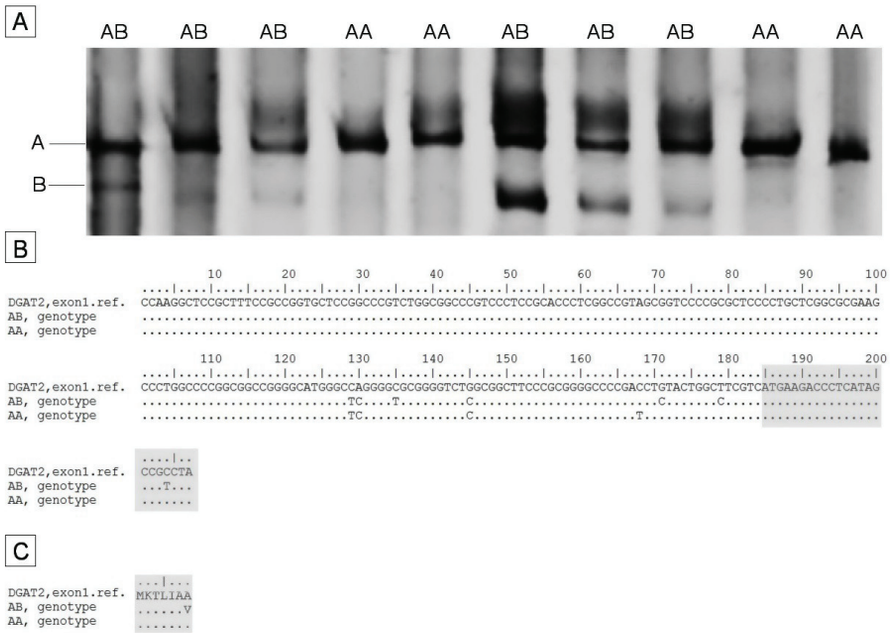

Fig. 1. SSCP-sequencing results for Holstein cattle partial exon 1 of the *DGAT2* genetic locus. **A**) PCR-SSCP of the two genotypes (*AB* and *AA*) and the two alleles (*A* and *B*). **B**) Multiple nucleic acid sequences alignment of two SSCP genotypes with their reference sequence. The shaded sequences refer to the nucleic acid sequences of the exon 1. **C**) Multiple sequence alignment of amino acid sequences. The shaded sequences refer to the coding region of the exon 1.

binding site motif using the NHR scan program, we discovered that DR1, which is the most represented motif in the reference DGAT2 enzyme, is converted into ID site, which is the most represented motif in AB genotype DGAT2 enzymes. Meanwhile, the original DR1 motif was lost in the *AA* genotype (Table 6, A).

The only observed exonic SNP was found in the *AB* genotype (C 204 T). It was found that this SNP constitutes non-synonymous mutation (Fig. 1C), in which the amino acid alanine (Ala) (A) (Fig. 2A) is changed into the amino acid valine (Val) (V) (Fig. 2B). The consequences of this nsSNP were evaluated using different publicly available computational algorithms, namely SIFT, PANTHER, I-Mutant 2, SNAP2, and PROVEAN bioinformatics tools. By comparing the prediction of these methods, the deleterious consequence of this nsSNP is confirmed in four out of the five utilized tools. Usually, the SIFT tool is utilized first, which predicts whether an amino acid substitution has affected the protein function among the related genes and domains over the evolutionary time (Ng, Henikoff, 2006). SIFT predicted the p.Ala7Val nsSNP to be deleterious, with a highly deleterious tolerance index score 0.00, which could affect the protein function in this gene. The SIFT predicted deleterious characterization of p.Ala7Val nsSNP was further validated by PANTHER through investigating the effect of this nsSNP on DGAT2 protein function

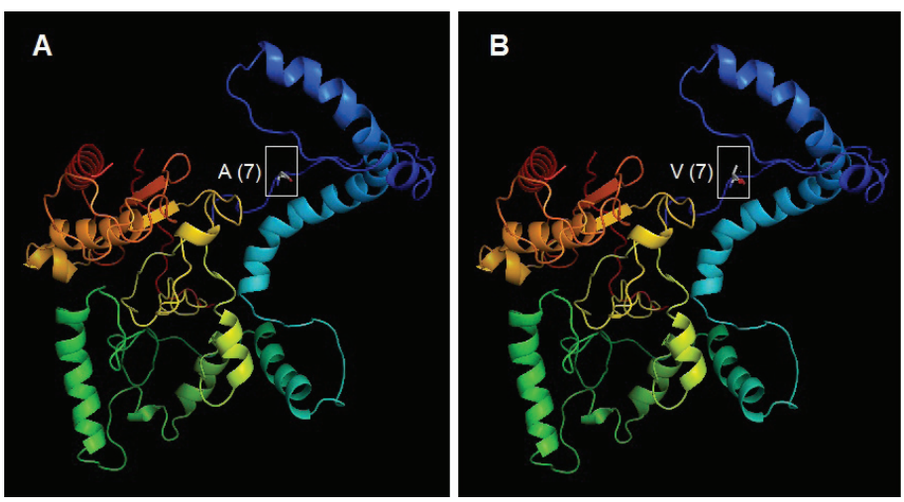

Fig. 2. Postulated three-dimensional structure of the bovine DGAT2 protein showing the change of amino acid 'Ala' (in the seventh amino acid position) of the reference protein (A) into amino acid 'Val' of the *AB* genotype DGAT2 protein of Holstein cattle (B).

Table 4. Differences in the pattern observed in the reference (Ref), *AB*, and *AA* genotypes of (A) *DGAT2* 5′-UTR fragment and (B) *BTN* 5′-UTR fragment with their putative transcription factors as predicted by PROMO program. Each transcription factor was presented by a unique colour/number

(A) ***DGAT2***						
	130 bp	140 bp	150 bp	160 bp	170 bp	180 bp
Ref	GR-beta, NF-1(37), NF-1(39), ENKTF-1, NF-1(98), XPF-1, EBF, Sp3, ETF	p53, GCF, AP-2beta, Sp3, BTEB4, NF-X3	NF-1(37), NF-1(39), Sp1(54), Sp1(59), ENKTF-1, MF3, NF-1(98)	ETF	GR-beta, Pax-6, T3R-beta, RC2, NHP-1	
AB	ETF	Alfin1, WT1-I		ETF	GR-beta, Pax-6, T3R-beta, RC2, NHP-1, Nkx2-1	GR-beta, Nkx2-1
AA		p53, GCF, AP-2beta, Sp3, BTEB4			NF-X3	
(B) ***BTN***						
	20 bp	30 bp	40 bp	50 bp	60 bp	70 bp
Ref	GATA-2(302), GATA-2(308), GATA-3	GATA-1, NIT2, GA-BF	MF3, ABF1, YY1(65), YY1(15)W	HOXA5, GATA-2(302), GT-1, PF1, GATA-1(67), GATA-1(305)	GATA-1, ABF1, C/EBPgamma, SQUA, DBP, XBP-1, C/EBPalpha, C/EBPbeta, Cut1, HOXA5, POUF1a, C/EBPdelta, unc-86, POU2F1, GATA-1(705)	C/EBPbeta, GR, DBP, c-Jun(31), JunB, c-Fus, c-Jun(33), NHP-1, GR-alpha
AB	XBP-1, RXR-alpha			SPF1, HOXD9, HOXD10, c-Ets-1, GA-BF	Cut1, HOXA5, POUF1a, C/EBPdelta, unc-86, Prd	HOXA5, POUF1a, C/EBPalpha, C/EBPdelta, unc-86, 56, POU2F1, HNF-3alpha
AA	GATA-2(302), GATA-2(308), GATA-3	GATA-1, NIT2, GA-BF	MF3, ABF1, YY1(65), YY1(15)	GATA-2(302), GT-1, PF1, GATA-1(67), GATA-1(305)	GATA-1, ABF1, SQUA, XBP-1	C/EBPbeta, GR, DBP, c-Jun(31), JunB, c-Fus, c-Jun(33), NHP-1, GR-alpha, HNF-3alpha
	80 bp	90 bp	100 bp	110 bp	130 bp	140 bp
Ref	JunD, Pax-6, Gt			Pax-5, Nrf2:MafK,	C/EBPbeta, GR, HNF-3, HNF-3beta, HOXD8(26), HOXD8(54), GR-alpha, TFIID, Nkx6-2, PF1, HNF-3alpha	C/EBPbeta, GR, HNF-3, HNF-3beta, Pax-6, HOXD8(26), HOXD8(54), GR-alpha, TFIID, Nkx6-2, PF1, HNF-3alpha
AB	SPF1, Nkx6-2	Nkx2-1	FACB, GR-beta	FACB, Mad	HOXA5, RXR-alpha	HOXA5, POUF1a, C/EBPdelta, RXR-alpha
AA	JunD, Pax-6, SPF1, Nkx6-2	Nkx2-1	FACB, GR-beta	Pax-5, Nrf2:MafK	HNF-3, HNF-3beta, HOXD8(26), HOXD8(54), GR-alpha, TFIID, Nkx6-2, PF1, HNF-3alpha	C/EBPbeta, GR, HNF-3, HNF-3beta, HOXD8(26), HOXD8(54), GR-alpha, TFIID, Nkx6-2, PF1, HNF-3alpha
	150 bp	170 bp	180 bp	190 bp		
Ref	NF-X3	Alfin1, ABF1, Fra-1, POU2F1, GATA-1(705)	LFA1, C/EBPbeta, GR, HOXA5, POUF1a, POU2F1, Nkx2-1, GR-alpha, SF-1	FACB, DSXF, DSXM, GR-beta		
AB	FACB, GATA-2(302), GATA-2(308), GATA-1, Alfin1, DBP, Cut1, GATA-1(67), GATA-1(305)	Pax-5, MYB2, HSF1, GAMYB, PR B, PR A, Prd, HNF3-beta	GR, GATA-1(705), GR-alpha, SF-1			
AA	NF-X3	Alfin1, ABF1, Fra-1, Prd	GATA-1(705)	FACB, DSXF, DSXM, GR-beta		

using HMM-based tool (Tang, Thomas, 2016). The later tool confirmed the SIFT prediction tool. The analysis of PANTHER revealed that p.Ala7Val nsSNP reflected a probably damaging preservation time score (time > 450 my), thus PANTHER classified them as deleterious. To add another layer of confirmation, the effect of this nsSNP was also analyzed using I-Mutant 2.0. It gave result in the form of the effect of mutants on the stability of protein with the reliability index at pH 7.0 and temperature 25°C (Capriotti et al., 2005). The p.Ala7Val nsSNP showed a decrease in stability of the DGAT2 protein. The SNAP2 server was also utilized in this regard (Smigielski et al., 2000), and it also predicted that p.Ala7Val had a damaging effect (score > 0). However, the prediction of the functional impact of p.Ala7Val nsSNP on the biological function of the DGAT2 protein was further investigated using PROVEAN tool (Choi et al., 2012). In contrast to the other four previously mentioned bioinformatics tools, PROVEAN revealed that this missense mutation was not deleterious (Table 7). It predicts that p.Ala7Val is neutral (below –2.5).

BTN genotyping

Undoubtedly, the small size of the studied population is still a shortcoming of our research. Nevertheless, two distinct genotypes were observed in this population. As in the *DGAT*2 case, two genotypes and two alleles were detected in the *BTN* fragment (Fig. 3A). The POPGENE 32 software showed that the observed frequencies of genotypes were about 0.73 ($n = 44$) and 0.27 ($n = 16$) for *AB* and *AA* genotypes, respectively. The observed frequency of allele *A* was 0.633, and of allele *B* 0.366 (Table 2, B). On the other hand, sequencing results confirmed these two different electrophoretic SSCP patterns with a very high SNPs variability, especially in the *AB* genotype as many SNPs were detected between the two resolved genotypes and the reference *BTN* sequence. The pattern and nature of each SNP that was detected by sequencing indicated that both genotypes have many novel SNPs. The genotype *AB* has twenty-two novel SNPs (C 21 T, T 22 C, A 35 T, C 54 T, A 55 C, T 57 A, G 65 T, G 71 C, T 78 A, A 94 G, C 110 G, T 128 A, C 129 A, T 132 C, A 142 G, A 146 G, C 148 A, T 151 A, A 163 G, C 169 A, T 175 C, C 184 G), while the genotype *AA* has six novel SNPs (T 57 A, T 78 A, A 94 G, C 129 A, T 175 C, G 181 C) (Table 3, B). In both genotypes, five in common SNPs were observed (T 57 A, T 78 A, A 94 G, C 129 A, T 175 C) (Fig. 3B).

The variations in the binding with the putative transcription factors were recorded for all 5′-UTR regions of the *BTN* submitted sequences using the PROMO computational tool (Farré et al., 2003). Through the later tool, the multitude of the binding with putative transcription factors was obviously observed in both genotypes. This high binding ability with putative transcription factors was highly reduced by representing only the differences among the submitted sequences, while the other similarities were omitted (Table 4, B). Nevertheless, the wide spectrum of the binding with putative transcription factors could not be eliminated in this program. Thus, another UTR prediction tool was utilized to sum up this wide variability spectrum, namely the UTRscan program (Sandelin, Wasserman, 2005). Indeed, the PROMO results were confirmed by the UTRscan server (Grillo et al., 2010), which showed a dramatic change in the positions of two regulatory elements (IRES and uORF) out of three UTRscan observed elements (Table 5, B). In the *AB* genotype, the magnitude of these changes

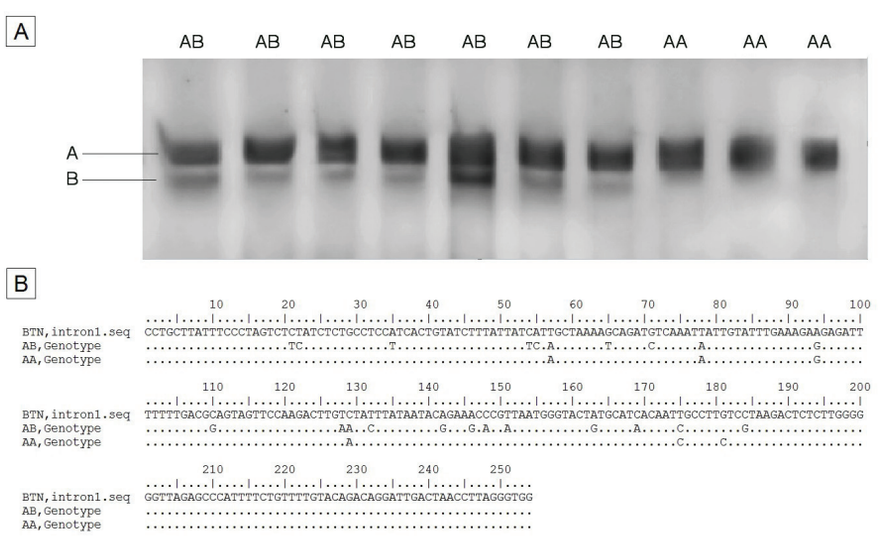

Fig. 3. SSCP-sequencing results for Holstein cattle partial 5′-UTR of the BTN genetic locus. A) PCR-SSCP of the two genotypes (AB, and AA) and the two alleles A and B. B) Multiple nucleic acid sequences alignment of two SSCP genotypes with their reference sequence.

Table 5. Single nucleotide polymorphisms in the 5′-UTR region of (A) *DGAT*2 and (B) *BTN* genes predicted significant by the UTRscan program

		Regulatory element	Position	Sequence
(A) *DGAT*2	reference sequence, *AB* and *AA* genotypes	IRES	108, 207	CCCGGCGGCCGGGGCATGGGCCAGGGGCGCGGG GTCTGGCGGCTTCCCGCGGGGCCCCGACCTGTAC TGGCTTCGTCATGAAGACCCTCATAGCCGCCTA
		uORF	123, 188	ATGGGCCAGGGGCGCGGGGTCTGGCGGCTTCCCG CGGGGCCCCGACCTGTACTGGCTTCGTCATGA
(B) *BTN*	reference sequence	IRES	174, 254	TTGCCTTGTCCTAAGACTCTCTTGGGGGGTTAGAG CCCAT TTTCTGTTTTGTACAGACAGGATTGACTAA CCTTAGGGTGG
		uORF	69, 206	TCAAATTATTGTATTTGAAAGAAGAGATTTTTTTGA CGCAGTAGTTCCAAGACTTGTCTATTTATAATACAG AAACCCGTTAATGGGTACTATGCATCACAATTGCCT TGTCCTAAGACTCTCTTGGGGGGT TAG
		MBE	112, 116	GTAGT
	AB genotype	IRES	178, 254	CTTGTCGTAAGACTCTCTTGGGGGGTTAGAGCCCA TTTTCTGTTTTGTACAGACAGGATTGACTAACCTTA GGGTGG
		uORF	154, 243	ATGGGTACTATGCGTAACAATCGCCTTGTCGTAAGA CTCTCTTGGGGGGTTAGAGCCCATTTTCTGTTTTGT ACAGACAGGATTGACTAA
		MBE	112, 116	GTAGT
	AA genotype	IRES	179, 254	TTCTCCTAAGACTCTCTTGGGGGGTTAGAGCCCATT TTCTGTTTTGTACAGACAGGATTGACTAACCTTAGG GTGG
		uORF	69, 206	TCAAATTATTGTATTTGAAAGAAGAGATTTTTTTGAC GCAGTAGTTCCAAGACTTGTCTATTTATAATACAGAA ACCCGTTAATGGGTACTATGCATCACAATTGCCTTGT CCTAAGACTCTCTTGGGGGGT TAG
		MBE	112, 116	GTAGT

was found more dramatic than in the *AA* genotype. This, in turn, may be attributed to the higher level of variability observed in this form of variation. The IRES-regulatory motif that extended from 179 to 256 bp in the reference sequence changed into 174 to 265 bp in the *AB* genotype maybe due to 184C>G SNP, and from 178 to 254 bp in the *AA* genotype maybe due to 181G>C SNP. The other modulated regulatory element is uORF. It extended from 69 to 206 bp in the reference sequence and changed into 154 to 243 bp only in the *AB* genotype due to the higher level of SNPs in this genotype. In converse to *DGAT*2 fragment, both PROMO and UTRscan were not confirmed by the HNRscan server. When NHR scan was used to analyze the possible modulation in the BTN protein characterization, any noticeable changes were not seen, and no nuclear receptors binding factors were found in the reference sequences as long as in the observed genotypes in this studied fragment (Table 6, B). This may indicate that the observed changes in *BTN* fragment were not related to any postulated interaction with nuclear receptors at least in this genotyped 5′-UTR *BTN* fragment.

DISCUSSION

The rapid progress in our understanding the bovine milk genetic variation shows that complex phenotypes are affected by many different types of genetic changes. It is well known that SNPs are the general form of genetic variations among individuals and are thought to be responsible for the majority of inherited traits, including the inherited quantity as well as quality of bovine milk. This study focused on the analysis of the SNPs consequences of both *DGAT*2 and *BTN* genes because of several reasons: (1) The polymorphism of both bovine *DGAT*2 and *BTN* genes was potentially related to some milk productive traits (Bhattacharya et al., 2006; An et al., 2011); (2) Understanding the cellular mechanisms encoded by *DGAT*2 and *BTN* genes may be important for elucidating the molecular processes of pregnancy, lactation, obesity, and several metabolic disorders (Mather, Jack, 1993; Cases et al., 2001); (3) Both *DGAT*2 and *BTN* genes polymorphism could be considered as one of the best biological markers since they are suggested to play a significant role in marker assisted

Table 6. Possible changes in the reference (Ref), *AB*, and *AA* genotypes of both (A) *DGAT*2 and (B) *BTN* fragments with their NHR binding sites as analyzed by the NHR scan program

	Genotypes	Site type	Sequence	Start	End	log (viterbi probability)	log (forward probability)	log (background state probability)	forward background
(A) *DGAT*2	Ref	IR1	GGGGCAT GGGCCA	118	130	–19.9281	–19.3780	–20.9351	1.5571
	AB	DR4	CGACCTCT ACTGGCCT	165	180	–24.6876	–24.1448	–25.1242	0.9794
	AA	Not found	Not found	Not found	Not found	Not found	Not found	Not found	Not found
(B) *BTN*	Ref	Not found	Not found	Not found	Not found	Not found	Not found	Not found	Not found
	AB								
	AA								

Table 7. Possible effects and consequences of the observed non-synonymous single nucleotide polymorphisms on the *DGAT*2 structure and function as determined by several bioinformatics tools

Nucleotide change	Genotype	Bioinformatics tools					
			SIFT	PANTHER	PROVEAN	I-Mutant 2.0	SNAP2
C204T (p.Ala7Val)	*AB*	score:	0.00	456^{PT*}	-0.513^{CF*}	7^{RI*}	64
		prediction:	deleterious	probably damaging	neutral	decreased	effect

RI* = reliability index score, PT* = preservation time, CF* = cutoff (–2.5)

selection (An et al., 2011); (4) The combination of the crucial effects of both *DGAT*2 and *BTN* genetic polymorphism and their respective intervening in triglyceride synthesis and in lipid droplet synthesis makes these genetic loci act as excellent candidate genes that may impart into animal genetic breeding; (5) The study of coding and non-coding SNPs on both *BTN* and *DGAT*2 genetic loci might have a considerable impact on their corresponding proteins expression, structure, and function.

First of all, the statistical analysis of the present study confirmed the high polymorphic nature of both PCR-SSCP genotyped fragments. This may necessarily imply the utilization of both genetic fragments in further explanations to take a snapshot for the consequences of genotyping variations in both studied loci, *DGAT*2 and *BTN*.

***DGAT*2 genotyping**

As a part of the enzymatic machinery involved in the formation of triglycerides (Cases et al., 2001), the DGAT2 enzyme is located in the endoplasmic reticulum as a transmembrane enzyme which transfers acyl chains into the nascent lipid polypeptides. Hence, the NHR scan program was utilized to explore the possible role of the observed SNPs in the changing of the DGAT2 enzyme characterizations with respect to its nuclear surface motifs (Sandelin, Wasserman, 2005). Interestingly, a noticeable change was observed after mutating the *DGAT*2 with the *AB* genotype SNPs, suggesting a possible role of some of these SNPs in changing such motifs, which may consequently change the nuclear receptor-mediated metabolic pathway in which the DGAT2 enzyme is involved. As well, the NHR scan in the present study showed that in the *AA* genotype, the DGAT2 enzyme has lost its repeated motif completely. This might be attributed to the presence of 168C>T, which is the only nsSNP that exists uniquely in this genotype. Though the precise mechanism by which triglycerides are deposited into lipid droplets is unknown (Yen et al., 2008), this finding may pinpoint some initial data which may help in precise understanding these motifs which may change the substrates with which they interact. Pairwise, understanding the functions of some of nsSNPs can greatly help understand the bovine genotype/phenotype variations that may have several effects on some milk synthesis characterization. As the genomic variations view among cattle is a common sense, the nsSNPs that exist through the coding region of this genome can, therefore, become very important since they modulate the amino acid composition. Such alterations can have an impact on protein structure, function, stability, and subcellular localization. Likewise, nsSNPs may affect gene expression by modifying DNA and transcription factor binding, inactivate active sites of enzymes or change

splice sites, thereby producing defective gene products (Masoodi et al., 2012). Therefore, the observed p.Ala7Val in the *AB* genotype is given a special attention in this study. To determine the functional effect of this nsSNP in DGAT2 protein, five widely used *in silico* tools were employed in this study, specifically SIFT, PANTHER, I-Mutant 2.0, SNAP2, and PROVEAN (Smigielski et al., 2000; Pauline, Steven, 2003; Capriotti et al., 2005; Choi et al., 2012; Tang, Thomas, 2016, respectively). However, not all the five utilized bioinformatics tools gave the same predictions on the deleterious effect of p.Ala7Val nsSNP on the DGAT2 protein. The difference in the results of these several prediction tools is due to the difference in features utilized by the methods, therefore, we would expect the outcomes to occur dissimilar at some point (de Alencar, Lopes, 2010). However, four out of the five tools provided sufficient indication for the deleterious effect of this p.Ala7Val nsSNP. As long as many nsSNPs may modify enzyme activity, destabilize protein structures or disrupt protein interactions (Patel et al., 2015), it is not unusual for this nsSNP to potentially alter some features of Holstein milk components. Therefore, we predict that the mutation from Ala to Val at residue position 7 in the native-type DGAT2 protein will be deleterious and would have drastic consequences on its structure and function. Unfortunately, we were not able to compare any data with other results since there are no previous SNPs based bioinformatics studies focusing on bovine exon 1 of the *DGAT2* genetic locus. In addition, the potential relationship between this p.Ala7Val could be clearly hypothesized when it was accompanied by full phenotypic records for Holstein milk production traits. Nonetheless, this study has provided a pilot data that filter out the likelihood of this nsSNP to affect the DGAT2 protein function, which likely has an impact on the quality of some bovine milk components as deleterious nsSNPs have a profound influence on the protein structure and interaction (Yadav et al., 2014). Therefore, these data provide the first evidence about the nature of such missense mutation in this gene. However, other missense mutations in DGAT2 proteins were observed in other species, too, e.g. in goats, which could have effects on both protein structure and function (An et al., 2011). However, we believed that the combination between SSCP-sequencing-based *in vitro* and SIFT/PANTHER/I-Mutant 2.0/SNAP2/PROVEAN based *in silico* analysis can provide a potent tool for the geneticists and breeders to identify the unknown SNPs that may be useful for establishing a possible impact on the biological function of DGAT2 proteins.

BTN genotyping

We found that all the observed polymorphisms of 5′-UTR of the bovine *BTN* gene were novel in comparison with previously published data (Winter et al., 2003). Though this study was conducted on a different bovine genetic locus, our results are in accordance with some data that revealed the highly polymorphic state of the exon 8 of *BTN* gene that extended to the protein level of the Holstein cross-bred cattle (Bhattacharya et al., 2006). Moreover, this sort of *BTN* variability was extended across other species to include both ovine and bubaline (Bhattacharya et al., 2007). However, the size of fat globules, which are synthesized by *BTN* gene-encoded products, varies from species to species and this variation might be due to the differences in the *BTN* genotypes. Thus, it was demonstrated through the *in vitro* study portion of this fragment that *BTN* genotyping through PCR-SSCP sequencing can explore this kind of variability. The *in silico* study portion of this fragment, in which the PROMO program was utilized first, showed it was very difficult to obtain an accurate view of the effect of 5′-UTR SNPs on the genetic expression of the *BTN* fragment. In contrary, UTRscan program has overcome these limitations by showing the involvement of 12 SNPs (positioned between 110 to 184 bp in the *AB* genotype) and 3 SNPs (positioned in 129 to 181 bp of the *AA* genotype) in the changing of the positions of both IRES and uORF regulatory elements. This, in turn, may subsequently alter the level of *BTN* gene expression through these postulated re-positioning effects on both the *AB* and *AA* genotypes. Noteworthy, the studied fragment of *BTN* locus that belongs to the *AB* genotype is highly polymorphic and may be regarded as a mutational hotspot, or the highest variable site, leading to a considerable diversity. The high genetic dissimilarity that we observed in the *BTN* locus of this studied population might be due to several factors, such as the type, location, or even the climate of the breed used since the nature of polymorphisms was potentially highly dependable on such factors in the determination of its final genetic polymorphism (Berman, 2011). However, the differences in the breeding styles among countries have led to the suggestion that genetic differentiation between countries may exist as a potential consequence of the local selection, environmental effect, genetic drift, and genetic isolation that could not be excluded from this explanation (Hammami et al., 2009).

CONCLUSION

Our data analysis showed that Holstein cattle breed can be differentiated by SSCP using the previously described *DGAT*2 and *BTN* genetic fragments. Likewise, the results of our study suggest that the application of computational tools might provide an indispensable approach to select functional SNPs in both *DGAT*2 and *BTN* loci. The current *in vitro* genotyping technologies suggest the likelihood for *DGAT2* and *BTN*

genes to be as markers of choice at the moment in an efficient and cost-effective manner as a highly variable genetic polymorphism was revealed in both loci. The stepwise prediction of the effect of the DGAT2 p.Ala7Val nsSNP (SIFT > PANTHER > SNAP2 > I-Mutant 2.0), the prediction matching among the tools and the trajectory analysis revealed that this p.Ala7Val nsSNP is damaging and highly deleterious nsSNP affecting the stability of the DGAT2 protein.

ACKNOWLEDGEMENT

Sincere thanks are due to the Al-Qadisya Holstein cattle station owners for providing us with some logistic technical assistance.

REFERENCES

Al-Shuhaib MBSA (2017): A universal, rapid, and inexpensive method for genomic DNA isolation from the whole blood of mammals and birds. Journal of Genetics, 96, 171–176. doi: 10.1007/s12041-017-0750-6.

An XP, Song SG, Hou JX, Zhu CM, Peng JX, Liu XQ, Liu HY, Xiao WP, Zhao HP, Bai L, Wang JG, Song BY, Cao BY (2011): Polymorphism identification in goat DGAT2 gene and association analysis with milk yield and fat percentage. Small Ruminant Research, 100, 107–112. http://dx.doi.org/10.1016/j.smallrumres.2011.05.017

Bahrami A, Behzadi Sh, Miraei-Ashtiani SR, Roh SG, Katoh K (2013): Genetic polymorphisms and protein structures in growth hormone, growth hormone receptor, ghrelin, insulin-like growth factor 1 and leptin in Mehraban sheep. Gene, 527, 397–404. doi: http://dx.doi.org/10.1016/j.gene.2013.05.066

Berman A (2011): Invited review: Are adaptations present to support dairy cattle productivity in warm climates? Journal of Dairy Science, 94, 2147–2158. doi: 10.3168/jds.2010–3962.

Bhattacharya TK, Misra SS, Feroz DS, Shukla S, Kumar P, Sharma A (2006): Effect of butyrophilin gene polymorphism on milk quality traits in crossbred cattle. Asian-Australasian Journal of Animal Science, 19, 922–926. doi: 10.5713/ajas.2006.922.

Bhattacharya TK, Sheikh FD, Shukla S, Kumar P, Sharma A (2007): Differences of ovine butyrophilin gene (exon 8) from its bovine and bubaline counterpart. Small Ruminant Research, 69, 198–202. doi: 10.1016/j.smallrumres.2005.12.006.

Capriotti E, Fariselli P, Casadio R (2005): I-Mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research, 33, 306–310. doi: 10.1093/nar/gki375.

Cases S, Stone SJ, Zhou P, Yen E, Tow B, Lardizabal KD, Voelker T, Farese RV (2001): Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. Journal of Biological Chemistry, 276, 38870–38876. doi: 10.1074/jbc.M106219200.

Cheunsuk S, Sarataphan N, Kengwikum K (2010): Genotypic effect of butyrophilin gene on milk fat in Thai-Holstein dairy cow. Journal of Biotechnology in Living Production, 5, 22–31. doi: 10.3168/jds.2009-2550.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012): Predicting the functional effect of amino acid substitutions and indels. PLOS ONE, 7, e46688. doi: 10.1371/journal.pone.0046688.

Coleman RA, Lewin TM, Muoio DM (2000): Physiological and nutritional regulation of enzymes of triacylglycerol synthesis. Annual Review of Nutrition, 20, 77–103. doi: 10.1146/annurev.nutr.20.1.77.

de Alencar SA, Lopes JCD (2010): A comprehensive in silico analysis of the functional and structural impact of SNPs in the IGF1R gene. Journal of Biomedicine and Biotechnology, 2010, Article ID 715139. doi: 10.1155/2010/715139.

Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, Messeguer X (2003): Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic Acids Research, 31, 3651–3653. doi: 10.1093/nar/gkg605

Felius M, Theunissen B, Lenstra A (2015): Conservation of cattle genetic resources: the role of breeds. Journal of Agricultural Science, 153, 152–162. doi: 10.1017/S0021859614000124.

Gasser RB, Hu M, Chilton NB, Campbell BE, Jex AJ, Otranto D, Cafarchia C, Beveridge I, Zhu X (2006): Single-strand conformation polymorphism (SSCP) for the analysis of genetic variation. Nature Protocols, 1, 3121–3128. doi: 10.1038/nprot.2006.485.

Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, Gennarino VA, Horner DS, Pavesi G, Picardi E, Pesole G (2010): UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated region of eukaryotic mRNAs. Nucleic Acids Research (Database issue), 38, D75–D80. doi: 10.1093/nar/gkp902.

Grisart B, Farnir F, Karim L, Cambisano N, Kim JJ, Kvasz A, Mni M, Simon P, Frere JM, Coppieters W, Georges M (2004): Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proceedings of the National Academy of Sciences of the United States of America, 101, 2398–2403. doi: 10.1073/pnas.0308518100.

Gupta N, Pandey A, Malik G, Gupta SC (2009): Single nucleotide polymorphism (SNP) in growth hormone gene of Jakhrana, a prominent milk goat breed in India. Small Ruminant Research, 81, 35–41. doi: 10.1016/j.smallrumres.2008.11.001.

Hammami H, Rekik B, Gengler N (2009): Genotype by environment interaction in dairy cattle. Biotechnology, Agronomy, Society and Environment, 13, 155–164.

Heid HW, Keenan TW (2005): Intracellular origin and secretion of MFG. European Journal of Cell Biology, 84, 245–258. doi: 10.1186/1746-4358-4-7.

Hradecká E, Čítek J, Panicke L, Řehout V, Hanusová L (2008): The relation of GH1, GHR and DGAT1 polymorphisms with estimated breeding values for milk production traits of German Holstein sires. Czech Journal of Animal Science, 53, 238–245.

Kale DS, Yadav BR, Mukherjee A, Prasad J (2013): Single strand conformation polymorphism within butyrophilin gene and its relationship with milk yield in Indian River buffaloes. Buffalo Bulletin, 32, 253–259.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015): The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10, 845–858. doi: 10.1038/nprot.2015.053.

Komisarek J, Waskowicz K, Dorynek Z (2006): Analysis of the relationship between two single nucleotide polymorphisms of the butyrophilin (BTN1A1) gene and milk production traits in Jersey cattle. Annals of Animal Science, 6, 45–52.

Kuhn C, Thaller G, Winter A, Bininda-Emonds ORP (2004): Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics, 167, 1873–1881. doi: 10.1534/genetics.103.022749.

Liao PY, Lee KH (2010): From SNPs to functional polymorphism: the insight into biotechnology applications. Biochemical Engineering Journal, 49, 149–158. doi: 10.1016/j.bej.2009.12.021.

Masoodi TA, Talluri VR, Shaik NA, Al-Aama JY, Hasan Q (2012): Functional genomics based prioritization of potential nsSNPs in EPHX1, GSTT1, GSTM1 and GSTP1 genes for breast cancer susceptibility studies. Genomics, 99, 330–339. doi: 10.1016/j.ygeno.2012.04.006.

Mather IH, Jack LJW (1993): A review of the molecular and cellular biology of butyrophilin, the major protein of bovine milk-fat globule membrane. Journal of Dairy Science, 76, 3832–3850. doi: 10.3168/jds.S0022-0302(93)77726-7.

Ng PC, Henikoff S (2006): Predicting the effects of amino acid substitutions on protein function. Annual Review of Genomics and Human Genetics, 7, 61–80. doi: 10.1146/annurev.genom.7.080505.115630.

Ogg SL, Weldon AK, Dobbie L, Smith AJ, Mather IH (2004): Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk-lipid droplets. Proceedings of the National Academy of Sciences of the United States of America, 101, 10084–10089. doi: 10.1073/pnas.0402930101.

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989): Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proceedings of the National Academy of Sciences of the United States of America, 86, 2766–2770.

Patel SM, Koringa PG, Reddy BB, Nathani NM, Joshi CG (2015): In silico analysis of consequences of non-synonymous SNPs of Slc11a2 gene in Indian bovines. Genomics Data, 5, 72–79. doi: 10.1016/j.gdata.2015.05.015.

Pauline CN, Steven H (2003): SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Research, 31, 3812–3814. doi: 10.1093/nar/gkg509.

Robenek H, Hofnagel O, Buers I, Lorkowski S, Schnoor M, Robenek MJ, Heid H, Troyer D, Severs NJ (2006): Butyrophilin controls milk fat globule secretion. Proceedings of the National Academy of Sciences of the United States of America, 103, 10385–10390. doi: 10.1073/pnas.0600795103.

Sandelin A, Wasserman WW (2005): Prediction of nuclear hormone receptor response elements. Molecular Endocrinology, 19, 595–606. doi: 10.1210/me.2004-0101.

Smigielski EM, Sirotkin K, Ward M, Sherry ST (2000): dbSNP: a database of single nucleotide polymorphisms. Nucleic Acids Research, 28, 52355.

Spelman R, Georges M, Snell R (2002): Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research, 12, 222–231. doi: 10.1101/gr.224202.

Tang H, Thomas PD (2016): PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. Bioinformatics, 32, 2230–2232. doi: 10.1093/bioinformatics/btw222.

Thaller G, Kramer W, Winter A, Kaupe B, Erhardt G, Fries R (2003): Effects of DGAT1 variants on milk production traits in German cattle breeds. Journal of Animal Science, 81, 1911–1918. doi: 10.2527/2003.8181911x.

Wakimoto K, Chiba H, Michibata H, Seishima M, Kawasaki S, Okubo K, Mitsui H, Torii H, Imai Y (2003): A novel diacylglycerol acyltransferase (DGAT2) is decreased in human psoriatic skin and increased in diabetic mice. Biochemical and Biophysical Research Communications, 310, 296–302. doi: 10.1016/j.bbrc.2003.09.015.

Winter A, Kramer W, Werner FA, Kollers S, Kata S, Durstewitz G, Buitkamp J, Womack JE, Thaller G, Fries R (2002): Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA: diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proceedings of the National Academy of Sciences of the United States of America, 99, 9300–9305. doi: 10.1073/pnas.142293799.

Winter A, Eckeveld Van M, Bininda Emonds ORP, Habermann FA, Fries R (2003): Genomic organization of the DGAT2/MOGAT gene family in cattle (Bos taurus) and other mammals. Cytogenetic and Genome Research, 102, 42–47. doi: 10.1159/000075723.

Yadav P, Chatterjee A, Bhattacharjee A (2014): Identification of deleterious nsSNPs in α, μ, π and θ class of GST family and their influence on protein structure. Genomics Data, 2, 66–72.

Yeh FC, Yang R, Boyle T (1999): POPGENE Version 1.31. Microsoft Window-based freeware for population genetic analysis. University of Alberta, Edmonton.

Yen CE, Stone SJ, Koliwad S, Harris C, Robert V, Farese Jr RV (2008): DGAT enzymes and triacylglycerol biosynthesis. Journal of Lipid Research, 49, 2283–301. doi: 10.1194/jlr.R800018-JLR200.

Zenger KR, Khatkar MS, Cavanagh JA, Hawken RJ, Raadsma HW (2006): Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. Animal Genetics, 38, 7–14. doi: 10.1111/j.1365-2052.2006.01543.x.

Corresponding Author:

Mohammed Baqur S. Al - Shuhaib, Al-Qasim Green University, College of Agriculture, Department of Animal Production, Babil, Iraq, phone: +96 47 707 115 693, e-mail: mohammed79@agre.uoqasim.edu.iq, http://orcid.org/0000-0002-6458-2068