Transcript analysis at DGAT1 reveals different mRNA profiles in river buffaloes with extreme phenotypes for milk fat

Published version:

DOI:10.3168/jds.2017-12771

Journal of dairy science 100 (10) 2017 ; 8265-8276

DOI: 10.3168/jds.2017-12771

https://doi.org/10.3168/jds.2017-12771

Transcript analysis at *DGAT1* reveals different mRNA profiles in river buffaloes with extreme phenotypes for milk fat

M. Gu,[*],[†] G. Cosenza,[‡] I. Nicolae,[§] A. Bota,[#] Y. Guo,[†] L. Di Stasio,[*] A. Pauciullo[*]

[*]*Department of Agricultural, Forest and Food Science, University of Torino, Grugliasco (TO), Italy*
[†]*College of Animal Science and Technology, Beijing University of Agriculture, Beijing, China*
[‡]*Department of Agricultural Sciences, University of Naples 'Federico II', Portici (NA), Italy*
[§]*Research and Development Institute for Bovine Breeding, Balotesti, Bucharest, Romania*
[#]*Research and Development Station for Buffalo Breeding, Şercaia, Romania*

DGAT1 mRNA PROFILE IN LACTATING *RIVER BUFFALOES*

Corresponding author
Alfredo Pauciullo, Prof.
Department of Agricultural, Forest and Food Sciences, University of Turin
Largo P. Braccini, 2
10095 - Grugliasco (TO), Italy
Phone: +390116708598
Email: alfredo.pauciullo@unito.it

ABSTRACT

Buffalo *DGAT1* was mainly investigated for the characterization of the gene itself and for the identification of the K232A polymorphism similarly to what has been accomplished in cattle, whereas no information has been reported so far at mRNA level. The importance of *DGAT1* for lipid metabolism led us to investigate the transcript profiles of lactating buffaloes characterised by high (9.13 ± 0.23) and low (7.94 ± 0.29) for milk fat percentage, and to explore the genetic diversity at RNA and DNA level.

A total of 336 positive clones for the *DGAT1* cDNA were analysed by PCR and chosen for sequencing according to the differences in length. The clone assembling revealed a very complex mRNA pattern with a total of 21 transcripts differently represented in the two groups of animals ($P=0.0169$). Apart from the correct transcript (17 exons long), the skipping of the exon 12 is the most significant in terms of clone's distribution with 11.6% of difference between the two groups, whereas a totally different mRNA profile was found approximately in 12% of clones. The sequencing of genomic DNA allowed the identification of 10 polymorphic sites at intron level, which clarify, at least partially, the genetic events behind the production of complex mRNAs.

Genetic diversity was found also at exon level. The SNP c.1053C>T represents the first example of polymorphism in a coding region for the *DGAT1* in the Italian Mediterranean breed. In order to establish whether this polymorphism is present in other buffalo breeds, a quick method based on PCR-RFLP was set-up for allelic discrimination in the Italian Mediterranean and the Romanian Murrah (in total 200 animals). The alleles were equally represented in the overall population, whereas the analysis of the two breeds showed inverted frequencies which resulted different ($p<0.01$), likely indicating diverse genetic structure of the two breeds. The T allele might be considered as the ancestral condition of the *DGAT1* gene, being present in the great part of the sequenced species.

These data add knowledge at transcript and genetic level for the buffalo *DGAT1* and open the opportunity for further investigation on other genes involved in the milk fat metabolism for the river

buffalo, including the future possibility to select alleles with quantitative and/or qualitative favourable effects.

Key words: *DGAT1*, transcript analysis, alternative splicing, genetic diversity, river buffalo

INTRODUCTION

The Mediterranean river buffalo represents a fundamental economic resource for Italy, mainly for the milk used for different dairy productions. The growing interest at both national and international level for the most famous buffalo dairy product (mozzarella Campana PDO - Reg. EC 510/2006) led to a great development of the buffalo dairy industry which, in the last 10 years, doubled the number of buffalo stock, currently assessed in more than 400000 heads (www.faostat.org).

Despite such a great numerical increase, the productive level remains insufficient to satisfy the market demand and to meet the economic goals of farmers. Therefore, management, feeding and breeding improvements are still necessary to achieve these aims.

It is well known that among ruminants, the buffalo produces milk characterised by a higher level of fat. It varies between 7.5% at the beginning of the lactation (after the colostrum phase) and 12-14% at the end of the lactation (Arumughan and Narayanan, 1981; Catillo et al., 2002). As milk fat has a great influence on cheese-making properties and yield, one of the main goal of the Italian National Association of Buffalo Breeders (ANASB) is the increase of milk fat content, which contributes to the determination of the PKM (Production in Kg of Mozzarella), the genetic index used for the evaluation of EBVs. Therefore, the genetic improvement of buffalos for the fat content represents a fundamental step for the progress of this species.

Many candidate genes for lipid metabolism have been identified so far, including *FASN* (fatty acid synthase), *DGAT1* (diacylglycerol O-acyltransferase 1), *SCD* (stearoyl CoA desaturase), *ACACA* (acetyl-CoA carboxylase alpha). However, in the last 15 years, only the *DGAT1* has been recognised as strong functional candidate for the milk fat content (Winter et al., 2002; Grisart et al., 2002).

The *DGAT1* catalyses the last reaction step in the synthesis of triacylglycerol. In cattle, a non-conservative substitution at the exon 8 responsible for the amino acid change K232A has been associated to high and low milk fat percentage (Winter et al., 2002; Thaller et al., 2003; Grisart et al., 2004; Kühn et al., 2004) and, later, associated also to milk fat composition (Schennink et al., 2007; Schennink et al., 2008; Conte et al., 2010) and to milk fat globule structure (Argov-Argaman et al., 2013). This polymorphism has been deeply investigated worldwide and found in many cattle breeds.

Conversely, the *DGAT1* gene in river buffalo has received less attention so far, with information limited to the gene structure (Yuan et al., 2007; Mishra et al., 2007) and polymorphism detection (Yuan et al., 2007; Mishra et al., 2007; Raut et al., 2012; Silvia et al., 2016). In this respect, the K232A polymorphism has been investigated also in buffalo breeds (Tantia et al., 2006; Shi et al., 2012), which resulted monomorphic for the K allele.

Recently, new polymorphic sites have been identified and associated to fat trait. Cardoso et al. (2015) found that a variable nucleotide repeat (VNRT) in the promoter region of *DGAT1* explained the 32% of additive genetic variance of fat percentage, and de Freitas et al. (2016) reported a SNP in the exon 17 significantly associated with fat and protein percentage in Brazilian Murrah buffaloes.

Apart from these studies, no additional information is available and no investigation has been carried out at transcriptomic level for the buffalo *DGAT1*. Furthermore, no genetic diversity has been reported in *DGAT1* coding regions for the Italian Mediterranean breed.

To contribute to a more detailed knowledge of the river buffalo *DGAT1*, an investigation was undertaken to analyse the transcriptional profiles of buffalo cows characterised by extreme phenotypes (high and low) for milk fat percentage (FP), and to explore the genetic diversity at RNA and DNA level.

MATERIAL AND METHODS

Sample collection and nucleic acid isolation

Milk and blood samples were collected from eight unrelated lactating buffalos reared in Piedmont region (Northern Italy) and belonging to one farm. They were chosen among more than 500 lactating buffaloes ranked for milk FP, and separated in two groups at the extreme sides for this trait were selected: four buffalo cows (high group) with high FP (9.13 ± 0.23), and four buffalo cows (low group) with low FP (7.94 ± 0.29). The milk yield (kg/day) was comparable for the eight animals (8.74 ± 0.96). The selection was based on their monthly test-day milk FP records for the current and previous lactations, which were provided by the Italian National Association of Buffalo Breeders. The animals were comparable for age (approximately 6 years old), feeding system, number of lactation (third) and lactation stage (4th month).

Additional 200 blood samples were collected for DNA genotyping, 100 samples (Italian Mediterranean breed) from 8 buffalo farms in Campania region (Southern Italy) and 100 samples (Murrah breed) from Şercaia research station (Romania).

Total RNA was isolated from milk somatic cells using TRIzol® (Invitrogen, Carlsbad, CA) according to manufacturer's guidelines, whereas the remaining traces of DNA were removed with DNase I (Qiagen). The genomic DNA was isolated from blood samples according to the procedure described by Sambrook et al. (1989) and then resuspended in 100 µl TE buffer pH 7.6 (10 mM Tris, 1mM EDTA).

RNA and DNA concentrations and $OD_{260/280}$ ratios were measured with the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Average concentrations were 50 ng/µl for both RNA and DNA samples. A ratio higher than 1.8 was recorded for all the DNA samples, whereas a ratio higher than 2.0 was detected for RNA samples. These values are generally accepted as pure for DNA and RNA respectively, and therefore they indicated the absence of protein, phenol or other contaminants.

RT-PCR, cloning and sequencing

The reverse transcription of total RNA was performed by using an oligo dT_{18}. The mix was set up in a final volume of 20µl using ImProm-II™ Reverse Transcriptase (Promega) according to the standard protocol recommended by the firm. The PCR reaction was performed by using the following primers (*forward*) 5'-ATGGGCGACCGCGGCGG-3' and (*reverse*) 5'-TCAGGTGCCGGCTGCCGG-3', corresponding to the nucleotides 1-17 (exon 1) and complementary to the nucleotides 1453-1470 (exon 17), covering the whole river buffalo *DGAT1* cDNA (EMBL ID: DQ120929).

The PCR reaction mix (50 µl) comprised: 50 ng of total cDNA, 1X PCR Buffer (Promega, Madison, WI, USA), 2.5 mM $MgCl_2$, 5 pmol of each primer, dNTPs each at 200 µM, 1 U of *Taq* DNA Polymerase (Promega). PCR was performed under the following thermal conditions: 97°C for 4 min, 35 cycles at 97°C for 45 s, 60.5°C for 45 s, 72°C for 90 s, and the final extension at 72°C for 5 min.

The amplified products were first analysed by electrophoresis on 1.5% agarose gel in 0.5X TBE buffer, pool together according their classification (high or low FP group) and then cloned into pGEM®-T Easy Vector (Promega). The ligation products were transformed into JM109 High-Efficiency Competent Cells (Promega) following manufacturer's guidelines. White recombinant clones were randomly chosen and screened by PCR according to Pauciullo and Erhardt (2015) using the following combination of primers: exon 1 For 5'-ATGGGCGACCGCGGCGG-3' together with exon 11 Rev 5'-GCTTCATGGAGTTCTGGA -3' and exon 11 For 5'-TCCAGAACTCCATGAAGC-3' together with exon 17 Rev 5'-TCAGGTGCCGGCTGCCGG-3'. The strategy to divide the total *DGAT1* cDNA (1470 bp) into two sub-amplicons of 931 bp and 539 bp was necessary to allow the identification of cDNA populations with smaller size otherwise not detectable on the traditional gel agarose analysis.

All amplicons different in size and at least 4 amplicons with the same length were purified using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) and sequenced both directions in out-sourcing (Microsynth AG, Switzerland) using Sanger DNA sequencing technologies.

SNP discovery, validation and genotyping of the exon 13 c.1053C>T by DdeI

The genetic events responsible for the different transcripts were investigated by sequencing the genomic DNA of the same buffaloes for the amplicons reported in table 1. The sequencing results allowed the validation of the genetic diversity found at RNA level, as well as the discovery of variability at intron level.

Standard PCR mixture conditions were applied (as reported above), whereas the annealing temperature was adjusted according to the specific primer couples (table 1).

The entire panel of 200 animals was genotyped for the SNP c.1053C>T using a PCR-RFLP method. A DNA fragment 493 bp long spanning from the splicing acceptor site of the exon 12 to the 23rd nucleotide of the exon 15 of the buffalo *DGAT1* gene was amplified by using the following primers: forward 5'-AGGACATGGACTACTCCC-3' and reverse 5'-GGAGCATGGGCTTGTAGA -3', with the same PCR mixture conditions reported above. Thermal conditions were: 97°C for 4 min, 35 cycles at 97°C for 45 s, 63.2°C for 45 s, 72°C for 45 s, with the final extension at 72°C for 5 min. Product specificity was confirmed by Sybr Green (Sigma-Aldrich) stained 1.5% agarose gel electrophoresis.

Ten μl of each amplicon was digested with 5 U of *Dde* I endonuclease (5'-C↓TNAG-3') (New England Biolabs) over-night at 37°C. The digested products were analysed by electrophoresis in 2.5% agarose gel in 0.5X TBE buffer and stained with Sybr Green (Sigma-Aldrich).

Bioinformatics

Homology searches, comparison among sequences, and multiple alignments were performed by DNAsis-Pro (Hitachi Software Engineering Co., San Bruno, CA, USA). The same software was used to estimate the number of amino acids of the putative protein isoforms. Splice site prediction was performed by NNSPLICE ver. 0.9 (http://www.fruitfly.org/seq_tools/splice.html), whereas branch point prediction was carried out by SVM-BP finder software (http://regulatorygenomics.upf.edu/Software/SVM_BP/). Allelic frequencies and Hardy-Weinberg equilibrium were evaluated for the SNP c.1053C>T using PopGene software ver. 1.32 (University of Alberta, Canada). Contingency tables and χ2 were used to evaluate differences both in the distribution of clones between the low and high milk FP groups, both in the allele frequencies of the two breeds (table 3) using SAS system software ver. 9.4 (SAS Institute Inc). The significance level was set at $P<0.05$.

RESULTS

Comparative transcript analysis at DGAT1 gene

The transcripts of the *DGAT1* gene were isolated from eight lactating buffaloes divided into two groups and characterised by extreme phenotypic values for FP. A total of 336 positive recombinant clones (172 clones for the high FP and 164 clones for the low FP group) were analysed by PCR and agarose gel electrophoresis. Afterwards, all clones showing different length by gel analysis were chosen for the sequencing.

The comparative sequence analysis showed a total of 21 transcripts of different length, which were diversely distributed in the two groups (table 2). The χ2 calculation carried out on the common mRNA populations (clone numbers from 1 to 6) and representing about 90% of the total detected cDNAs evidenced significant differences between the milk FP groups ($\chi2>13.805$; $P=0.0169$).

The electrophoretic variability of the different clones for the amplicons covering the exons 1-11 and 11-17 is reported in the figure 1.

A slightly higher transcript variability was found in the high FP group (14 different mRNAs) compared to the low FP group (13 mRNAs). The most represented mRNA in both groups was the correctly assembled (in total 172 out of 336 clones, 51.2%), with a wider distribution in the high FP

(54.6%) compared to the low FP group (47.5%). This transcript normally encodes for a protein of 489 amino acids.

Apart from the correctly assembled transcripts, the two groups share five additional mRNA populations (table 2), among which the skipping of the exon 12 is the most prevalent, with a relevant difference between the high (12.8%) and low group (24.4%). The distribution of the population spliced of the last 66 bp of the exon 8 was also interesting, being more prevalent in the high FP group (8.1%) than in the low FP group (4.9%). The reverse situation was observed for the skipping of the exon 16 (2.3% in high *vs* 6.1% in low FP group).

Around 12% of clones in both group of animals (12.8% in the high *vs* 12.2% in low FP) showed a totally different mRNA profile. In the low FP group, half of this transcripts (6.1%) involved the splicing of the last 66 bp of the exon 8 in combination with the skipping of exon 12 and exon 16 singularly, or as loss of these exons together. Conversely, more variable was the situation for the high FP group of clones, where the skipping of the exon 8 was found in combination with the gain of intron 13 and 14 individually, and the loss of the exon 13 and 14 together for a total of 4.6% of clones.

Furthermore, more complex mRNA rearrangements were found in both groups of clones, but they represented only a minor part of the *DGAT1* gene transcripts (table 2).

Genetic diversity

In order to explore the genetic events responsible for the observed trascriptomic profiles and to validate putative SNPs identified by the comparative analysis of clone sequences, five DNA amplicons covering the whole *DGAT1* gene (table 1) were sequenced for the eight investigated animals.

Genetic variability was found both at exon and intron level. In particular, the SNP c.1053C>T (69th nucleotide of the exon 13) already detected by the comparative analysis of the clone sequences was confirmed at DNA level. A genotyping method based on PCR-RFLP was set up to establish the distribution of this SNP in the population. In particular, the transition c.1053C>T creates a restriction site for the endonuclease *Dde* I (5'-C↓TNAG-3').

The digestion of the PCR product (493 bp) with *Dde* I allowed the identification of both alleles (figure 2). The allele frequencies determined in the Italian Mediterranean and Murrah breeds are reported in table 3. No deviation from Hardy-Weinberg equilibrium was observed (table 3).

The sequencing of genomic DNA allowed the identification of further 10 polymorphic sites at intron level. In particular, 7 transitions and 2 transversions as single nucleotide polymorphisms (table 4), and one insertion/deletion of 11 nucleotides (GTAGTGGGGGC) in the intron 13 were observed. These mutations may partially explain the variability found at mRNA level.

The comparison with the Chinese buffalo *DGAT1* gene sequence (Yuan et al., 2007) and the Indian buffalo *DGAT1* gene (Mishra et al., 2007) and cDNA (Venkatachalapathy et al., 2008) showed 64 additional nucleotide differences (respectively 39, 13 and 12), with two sites in the exon 1 which seem to be typical of the Italian Mediterranean breed (table 4).

DISCUSSION

This study reports a comparative transcript analysis for the *DGAT1* gene between two groups of buffaloes, similar for milk yield, but characterized by extreme phenotypes (high and low) for milk fat percentage.

The analysis of the mRNA populations carried out on a total of 336 positive recombinant clones showed a very complex transcriptomic pattern for this locus, with a total of 21 mRNAs differently represented in both groups (P=0.0169).

The first element affecting the intricate mechanism of RNA maturation, which reflects the occurrence of many splicing events, is the extremely split architecture of the genes (Heyn et al., 2014). In this respect, the *DGAT1* gene is extremely fragmented (17 exons). In fact, except for the first two exons widely spaced out by two introns (about 3600 bp and 2000 bp, respectively), the rest of the gene is densely compacted in less than 2500 bp, which include 15 small exons (variable in size between 39 bp of the exon 10 and 156 bp of the exon 17) spaced out by 14 small introns (from 66 bp of the intron 10 to 215 bp of the intron 5) (Yuan et al., 2007; Mishra et al., 2007). A very similar split structure characterises other genes expressed in the mammary gland and well-studied from a trascriptomic point of view. The αs-casein genes (*CSN1S1* and *CSN1S2*) are good examples, being composed of 19 and 18 exons, respectively. In goats, the analysis of the αs1-casein gene (*CSN1S1*) transcripts showed different mRNA profiles for the A (normal protein yield: ~3.5g/l), F (defective: ~0.45g/l) and N (null: ~0.0g/l) alleles, with 5, 9 and 12 transcript populations respectively (Ramunno et al., 2005). At least three different mRNAs have been identified in the goat αs2-casein gene (*CSN1S2*) (Ramunno et al., 2001), and multiple transcripts have been found also in the homologous ovine gene (Boisnard et al., 1991). Analogous multiple mRNA profiles have been detected for both genes also in other species (for a review Rijnkels, 2002).

In both groups of buffaloes, the most represented *DGAT1* mRNA population was the correctly assembled (1470 bp) coding for a functional protein of 489 amino acids, whereas the most significant skipping event involved the exon 12.

The sequencing of genomic DNA including the exon 12 and its flanking regions evidenced a transversion (g.10874T>A) falling 7 bp upstream the acceptor splice site of this exon. It is known that the removal of intron sequences from pre-mRNA is carried out by the spliceosome machinery, which recognizes specific sites (donor site, branch point, polypyrimidine and acceptor site) in a complex molecular mechanism. Any deviation from consensus can result in the overall decrease of affinity for the spliceosome (Clark and Thanaraj, 2002; Cosenza et al., 2009). Therefore, in order to verify the influence of the transversion g.10874T>A on splicing sites, the sequence between the exon 11 and the exon 13 underwent computational splice site prediction and branch point/polyPy analysis. The results confirmed that the presence of the adenine alters both the poly-pyrimidine tract (negatively affected in terms of length) and the branch point (decrease of identification score), resulting in the skipping of the exon 12 (Figure 3).

Despite this skipping event, the mature mRNA did not undergo any frame-shift and the termination codon was kept as in the normal isoform. The putative protein 474 amino acids long (Table 2) was different from other predicted *DGAT1* isoforms derived from the buffalo genome project (https://www.ncbi.nlm.nih.gov/protein/595763152,594082162,594082160,594082158,594082156,594082154). In fact, the analysis of the *DGAT1* isoforms available in NCBI showed that the complete skipping of the exon 12 was reported only in combination with the splicing of the last 66 nucleotides of the exon 8 (found also in the present study), but it was not reported as single event (as instead observed herein).

Great part of the clones showed the splicing of the last 66 bp of the exon 8 alone or in combination with other skipping events. This alternative splicing is consequence of the incorrect identification of a splice donor site at the exon 8 (figure 3) and it is responsible for a protein isoform 22 amino acids shorter when compared to the full-length form. The same event was already observed in bovine *DGAT1* (Grisart et al., 2004) and more recently in yak (Liu et al., 2011).

In cattle, this spliced form was reported to be indirectly associated to the K232A mutation by a linkage disequilibrium condition with the SNP Nt1501 (C-T) at the exon 17 (Grisart et al., 2004). In particular, the amount of alternatively spliced mRNA increased 1.2 times in K allele for an unknown motivation (rather than the K232A amino acid change itself), and it was not connected to difference in mRNA expression levels (Grisart et al., 2004), as recently confirmed also by further transcriptomic studies (Finucane et al., 2008; Bionaz et al., 2012; Cui et al., 2014).

The result of our investigation indicates that in buffalo the amount of mRNA alternatively spliced of the last 66 bp of the exon 8 is independent from the K232A mutation. In fact, the ‘intronification’ of this exon portion is very frequent in river buffalo although the K232A change has not been observed in this species (monomorphic for the K allele). On the other hand, our result shows similarities with the findings of Grisart et al. (2004). In fact, the group of buffaloes characterised by higher milk FP also has a higher percentage (8.14%) of the alternatively spliced mRNA compared with the low FP group (4.88%). Furthermore, the ratio between the percentages of spliced mRNA in high *vs* low group is 1.66, which is not far from the 1.2 reported for the ratio K/A in the Holstein Friesian (Grisart et al., 2004). Considering these similarities, and independently from the K232A mutation, it is possible that this splicing event is related to the increase of fat production, although the reason for that still remains to be established.

The two groups of investigated buffaloes share also three mRNA populations with lower incidence on the total number of clones: the skipping of the exon 16 as well as the insertion of the intron 13 alone and in combination with the skipping of the exon 6 and 7. The investigation of the relative intronic regions showed several polymorphic sites (table 4), however we could not link any SNP to the exon splicing events. In fact, the bioinformatic analysis for the spliceosome complex did not evidence alterations from the normal condition. Conversely, the insertion of the intron 13 is most probably due to the unsuccessful identification of the corresponding donor splice site (GT) which allows the ‘exonification’ of this intron. Furthermore, it is interesting to notice that this intron has a triplet structure (87 bp coding for 29 amino acids) and its insertion does not alter the original reading frame, so that the same primary amino acid sequence upstream (1-366) and downstream (366-489) the insertion is maintained.

The incorrect identification of donor or acceptor sites characterises other alternative splicing events involving also other exons (6, 7 and 10), as reported in figure 3. In these and in other alternative spliced mRNAs (table 2), the original reading frame and the original termination codon are conserved despite the skipping events. Conversely, premature termination codons (PTCs) characterise 7 different mRNAs (table 2), which are found in low percentage probably for their rapid degradation via nonsense-mediated mRNA decay (NMD). This is a surveillance mechanism, which detects and rapidly degrades mRNAs containing PTCs, and it is a fundamental cellular tool to eliminate mRNAs encoding C-terminally truncated proteins, which may possess dominant-negative or deleterious gain-of-function activity (Shi et al., 2015).

Apart from these short isoforms, all the other putative proteins vary in size between 431 and 539 amino acids and their functionality remains to be investigated. *DGAT1* was reported to form a homotetramer, which requires the NH_2- terminus (McFie et al., 2010). According to the findings of Zhang et al., (2014), none of putative protein isoforms found in the present study showed alterations of the essential *DGAT1* homodimerization or heterodimerization domain with *MGAT2*. This is located in the NH_2- terminus of the protein (amino acids 35-80), which in our investigation was never affected by skipping events. The comparison with human *DGAT1* (EMBL ID: NP_036211) also confirmed both the high similarity for the interspace between two transmembrane domains (amino acids 149-169), never skipped out, and the conservation of the FY-DWWN motif (amino acids 361-367), which is invariant in all members of the *ACAT* gene family, with the tyrosine and tryptophans being critical for the enzyme activity (Oelkelrs et al., 1998). Conversely, the predicated catalytic domain of *DGAT1* (amino acids 407-426) is partially removed in the mRNA isoforms spliced out of the exon 16 (amino acids 418-438).

The formation of *MGAT2*/*DGAT1* heterodimers is expected to bring the intermediate substrate (i.e. 1,2-DAG) to the proximity of the next catalytic enzymatic step (i.e. *DGAT1*) and to largely increase the efficiency of TAG synthesis (Zhang et al., 2014). Therefore, the lack of the exon 16 is supposed to prevent the subsequent catalytic step, so reducing the TAG synthesis. This event, together with the others aforementioned, may explain, at least partially, the lower fat production in the group of buffaloes characterised by higher incidence of mRNA transcript skipped of the exon 16 (table 2).

The investigation at DNA level allowed the identification of further genetic diversity. The sequencing of the *DGAT1* amplicons (table 1) for the 8 investigated samples evidenced a total of 10 polymorphic sites at intron level and one conservative SNP at the exon 13 (table 4). The latter mutation (c.1053C>T) falls at the third position of a triplet coding for an alanine (GCC^{Ala}→GCT) and it is not responsible for amino acid replacement. However, it is the first example of polymorphism in a coding region of *DGAT1* in the Mediterranean Italian river buffalo breed. Therefore, we decided to genotype 100 Italian buffaloes for this SNP. In order to establish whether this polymorphism is present in other buffalo breeds, the Murrah breed was also analysed.

A PCR-RFLP method was set up to discriminate the genotypes. The restriction pattern of the homozygous CC was characterized by 2 fragments of 299 bp and 194 bp, whereas the band 299 bp long was further restricted into two fragments of 189 bp and 110 bp in the presence of the thymine. The restriction pattern of the heterozygous genotype showed only three fragments (299 bp, 194/189bp and 110 bp) because the bands of 194 bp and 189 bp could not be differentiated in the agarose gel. The same situation for the homozygous TT which shows only 2 bands (194/189bp and 110 bp).

The alleles were equally distributed in the overall population which resulted in equilibrium for Hardy-Weinberg (table 3), whereas the two breeds showed statistically different frequencies (p=0.002). In particular, the Mediterranean Italian population had a higher frequency of the T allele (0.540), which can be considered as the ancestral condition being present in the great part of the *DGAT1* sequences of domestic animals (bovine: AY065621; zebu: EF636701; sheep: EU301803; goat: LT221856; pig: AY116586; horse: XM_005613365; donkey: XM_014858168; bactrian camel: XM_010961176); human (human: NG_034192); wild primates (chimpanzee: XM_016960108; gorilla: XM_019032174; orangutan: XM_009244184; bonobo: XM_008972742); felinae (cat: XM_004000171; cheetah XM_015076920; leopard: XM_019435164); and other wild species (white rhinoceros: XM_004443011; giant panda: XM_002922090; lemur: XM_012648079 and star-nosed mole: XM_012734968). Conversely, the Romanian Murrah breed showed the predominance of the C allele (0.525). Such polymorphism, although tested only in two breeds, adds useful knowledge for genetic biodiversity as potential tool to characterise buffalo breeds.

The comparison of the DNA sequences from the eight investigated animals with the other available buffalo *DGAT1* sequences (table 4) evidenced 64 additional polymorphic sites, most of which were detected in introns. The distribution of the SNPs highlighted a particular allelic combination in the *DGAT1* of the Italian buffalo, very similar to that of the Indian buffalo (EMBL ID: DQ886485) for great part of the gene (intron 2 – exon 17), but identical to Chinese buffalo (EMBL

ID: AY999090) at the exon 2. Conversely, two sites (g.3593C>T and g.3614T>C) seem to be typical of the Italian Mediterranean breed and, if confirmed, they could be useful for identification/traceability purposes.

Great part of the genetic diversity found at *DGAT1* in the Italian Mediterranean buffalo is different from that reported in other breeds, thus evidencing that it likely originated after the breed divergence. However, the SNPs c.1053C>T at the exon 13 and g.11618G>A at the intron 16 are exceptions. In fact, the former is present also in Romanian Murrah (as proved in the present study), whereas the latter had been already evidenced by Mishra et al. (2007). In particular, these authors identified 19 SNPs in Indian Mehsana breed, but genetic diversity was evidenced also in Chinese Murrah and Nili-Ravi buffaloes (Yuan et al., 2007), in Indian Pandharpuri breed (Raut et al., 2012), in 4 Iranian breeds (Naserkheil et al., 2016) and Brazialian Murrah (Cardoso et al., 2015).

Despite the buffalo *DGAT1* gene shows considerable genetic variation, polymorphisms found in exons are still very limited and, so far, causative mutations of extreme fat phenotypes were not found. However, recently, interesting studies have been carried out to associate VNRT (Cardoso et al., 2015) and SNP (de Freitas et al., 2016) to milk fat traits. Therefore, the polymorphisms found in the present study contribute to increase the knowledge on the genetic diversity at the *DGAT1* and, in the future, they might be used for similar association studies in Italian river buffaloes, as already performed in other genes candidate for quali-quantitative variations of milk traits (Pauciullo et al., 2012a, Pauciullo et al., 2012b).

CONCLUSIONS

The genetic improvement of Italian river buffalo aims to improve the milk production traits, therefore the genes involved in fat metabolisms are important targets of study. The present investigation on *DGAT1* transcripts has revealed different mRNA profiles for buffaloes characterised by extreme phenotypes for milk fat, providing fundamental knowledge to a research field completely unexplored in this species. We have elucidated great part of the genetic events responsible for the transcriptomic differences, showing that mutations at intron level affect recognition sites of the spliceosome machinery. Furthermore, the detection of the first polymorphism at the *DGAT1* exon 13 adds useful information not only for genetic biodiversity itself as tool to characterise the breeds, but also for possible future linkage analysis with fat or other milk traits in river buffalo.

ACKNOWLEDGMENTS

This research was financially supported by Agricola Project (PA_RIC_LOC_15_01), Department of Agricultural, Forest and Food Sciences, University of Torino, Grugliasco (Italy). We are very grateful to buffalo breeders for providing the samples for this research.

REFERENCES

Argov-Argaman, N., K. Mida, B-C. Cohen, M. Visker, K. Hettinga. 2013. Milk fat content and DGAT1 genotype determine lipid composition of the milk fat globule membrane. PLoS One. 8: e68707. DOI: http://dx.doi.org/10.1371/journal.pone.0068707

Arumughan, C., K.M. Narayanan. 1981. Influence of stage of lactation on the triacylglycerol composition of buffalo milk fat. Lipids. 16: 155-164. DOI: 10.1007/BF02535433

Boisnard, M., D. Hue, C. Bouniol, J.C. Mercier, P. Gaye. 1991. Multiple mRNA species code for two non-allelic forms of ovine αs2-casein. Eur. J. Biochem. 201: 633-641. DOI: 10.1111/j.1432-1033.1991.tb16324.x

Bionaz, M., K. Periasamy, S.L. Rodriguez-Zas, R.E. Everts, H.A. Lewin, W.L. Hurley, J.J. Loor. 2012. Old and new stories: revelations from functional analysis of the bovine mammary transcriptome during the lactation cycle. PLoS One. 7: e33268. DOI: 10.1371/journal.pone.0033268

Cardoso, D.F., G.F.P. de Souza, R.R. Aspilcueta-Borquis, F.R. Araujo Neto, G.M.F. de Camargo, N.A. Hurtado-Lugo, D.C.B. Scalez, A.C. de Freitas, L.G. Albuquerque, H. Tonhati. 2015.

Variable number of tandem repeat polymorphisms in *DGAT1* gene of buffaloes (*Bubalus bubalis*) is associated with milk constituents. J. Dairy Sci. 98: 3492-3495. DOI: http://dx.doi.org/10.3168/jds.2014-8729

Catillo, G., N.P.P. Macciotta, A. Carretta, A. Cappio-Borlino. 2002. Effects of age and calving season on lactation curves of milk production traits in Italian water buffaloes. J. Dairy Sci. 85:1298-1306. DOI: http://dx.doi.org/10.3168/jds.S0022-0302(02)74194-5

Clark, F., and T.A. Thanaraj. 2002. Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human. Hum. Mol. Genet. 11: 451-464. DOI: https://doi.org/10.1093/hmg/11.4.451

Conte, G., M. Mele, S. Chessa, B. Castiglioni, A. Serra, G. Pagnacco, P. Secchiari. 2010. Diacylglycerol acyltransferase 1, stearoyl-CoA desaturase 1, and sterol regulatory element binding protein 1 gene polymorphisms and milk fatty acid composition in Italian Brown cattle. J. Dairy Sci. 93 :753-763. DOI: 10.3168/jds.2009-2581

Cosenza, G., A. Pauciullo, M. Feligini, A. Coletta, L. Colimoro, D. Di Berardino, L. Ramunno. 2009. A point mutation in the splice donor site of intron 7 in the αs2-casein encoding gene of the Mediterranean river buffalo results in an allele-specific exon skipping. Anim. Genet. 40: 791. DOI: 10.1111/j.1365-2052.2009.01897.x

Cui, X., Y. Hou, S. Yang, Y. Xie, S. Zhang, Y. Zhang, Q. Zhang, X. Lu, G.E. Liu, D. Sun. 2014. Transcriptional profiling of mammary gland in Holstein cows with extremely different milk protein and fat percentage using RNA sequencing. BMC Genomics. 15: 226. DOI: 10.1186/1471-2164-15-226

de Freitas, A.C., G.M. Ferreira de Camargo, N.B. Stafuzza, R.R. Aspilcueta-Borquis, G.C. Venturini, M.M. Dias, D.F. Cardoso, H. Tonhati. 2016. Genetic association between SNPs in the *DGAT1* gene and milk production traits in Murrah buffaloes. Trop. Anim. Health Prod. 48: 1421-1426. DOI: 10.1007/s11250-016-1110-x

Finucane, K.A., T.B. McFadden, J.P. Bond, J.J. Kennelly, F.Q. Zhao. 2008. Onset of lactation in the bovine mammary gland: gene expression profiling indicates a strong inhibition of gene expression in cell proliferation. Funct. Integr. Genomics. 8: 251-264. DOI: 10.1007/s10142-008-0074-y

Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford, P. Berzi, N. Cambisano, M. Mni, S. Reid, P. Simon, R. Spelman, M. Georges, R. Snell. 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 Res. 12: 222-231. DOI: 10.1101/gr.224202

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J.J. Kim, A. Kvasz, M. Mni, P. Simon, J.M. Frere, W. Coppieters, M. Georges. 2004. Genetic and functional confirmation of the causality of the *DGAT1* K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Natl. Acad. Sci. USA. 101: 2308-2403. DOI: 10.1073/pnas.0308518100

Heyn, P., A.T. Kalinka, P. Tomancak, K.M. Neugebauer. 2014. Introns and gene expression: Cellular constraints, transcriptional regulation, and evolutionary consequences. Bioessays. 37: 148-154. DOI: 10.1002/bies.201400138

Kühn, C., G. Thaller, A. Winter, O. Bininda-Emonds, B. Kaupe, G. Erhardt, J. Bennewitz, M. Schwerin, R. Fries. 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

Liu, W., Y. Yue, Y.Q. Lin, Z.X. Liu, S.Y. Jin, Y.O. Xu, Y.C. Zheng. 2011. Yak *DGAT1* gene: Cloning, tissue expression profile, splicing and polymorphism analysis. Livest. Sci. 142: 264-269. DOI: http://dx.doi.org/10.1016/j.livsci.2011.08.006

Yuan, J., J. Zhou, X. Deng, X. Hu, N. Li. 2007. Molecular cloning and single nucleotide polymorphism detection of buffalo *DGAT1* Gene. Biochem. Genet. 45: 611-621. DOI: 10.1007/s10528-007-9100-3

McFie, P.J., S.L. Stone, S.L. Banman, S.J. Stone. 2010. Topological Orientation of Acyl-CoA:Diacylglycerol Acyltransferase-1 (*DGAT1*) and identification of a putative active site histidine and the role of the N terminus in dimer/tetramer formation. J. Biol. Chem. 285: 37377-37387. DOI: 10.1074/jbc.M110.163691

Mishra, B., M.S. Tantia, S.T.B. Kumar, R. K. Vijh. 2007. Characterization of the *DGAT1* gene in the Indian buffalo (*Bubalus bubalis*). Genet. Mol. Biol. 30: 1097-1100. DOI: http://dx.doi.org/10.1590/S1415-47572007000600012

Naserkheil, M., S.R. Miraie-Ashtiani, M. Sadeghi, A. Nejati-Javaremi, D. Lee. 2016. Sequence analysis and identification of polymorphism in Diacylglicerol Aciltransferase1 (*DGAT1*) gene in Iranian buffaloes. Genetics in the 3rd Millennium. 14: 4324-4331

Oelkers, P., A. Behari, D. Cromley, J.T. Billheimer, S.L. Sturley. 1998. Characterization of two human genes encoding Acyl Coenzyme A: cholesterol acyltransferase-related enzymes. J. Biol. Chem. 273: 26765-26771. DOI: 10.1074/jbc.273.41.26765

Pauciullo, A., G. Cosenza, R. Steri, A. Coletta, L. Jemma, M. Feligini, D. Di Berardino, N.P.P. Macciotta, L. Ramunno. 2012a. An association analysis between OXT genotype and milk yield and flow in Italian Mediterranean river buffalo. J. Dairy Res. 79: 150-156. DOI: 10.1017/S0022029911000914

Pauciullo, A., G. Cosenza, R. Steri, A. Coletta, A. La Battaglia, D. Di Berardino, N.P.P. Macciotta, L. Ramunno. 2012b. A single nucleotide polymorphism in the promoter region of river buffalo stearoyl CoA desaturase gene (SCD) is associated with milk yield. J. Dairy Res. 79: 429-435. DOI: 10.1017/S0022029912000507

Pauciullo, A., and G. Erhardt. 2015. Molecular characterization of the llamas (*Lama glama*) casein cluster genes transcripts (*CSN1S1*, *CSN2*, *CSN1S2*, *CSN3*) and regulatory regions. PLoS ONE 10: e0124963. DOI:10.1371/journal.pone.0124963

Ramunno, L., E. Longobardi, M. Pappalardo, A. Rando, P. Di Gregorio, G. Cosenza, P. Mariani, N. Pastore, P. Masina. 2001. An allele associated with a non-detectable amount of αs2 casein in goat milk. Anim. Genet. 32: 19-26. DOI: 10.1046/j.1365-2052.2001.00710.x

Ramunno, L., G. Cosenza, A. Rando, A. Pauciullo, R. Illario, D. Gallo, D. Di Berardino, P. Masina. 2005. Comparative analysis of gene sequence of goat *CSN1S1* F and N alleles and characterization of *CSN1S1* transcript variants in mammary gland. Gene. 345: 289-299. DOI:10.1016/j.gene.2004.12.003

Raut A.A., A. Kumar, S.N. Kala, V. Chhokar, N. Rana, V. Beniwal, S. Jaglan, S.K. Samuchiwal, J.K. Singh, A. Mishra. 2012. Identification of novel single nucleotide polymorphisms in the *DGAT1* gene of buffaloes by PCR-SSCP. Genet. Mol. Biol. 35: 610-613. DOI: 10.1590/S1415-47572012005000043

Rijnkels, M. 2002. Multispecies comparison of the casein gene loci and evolution of casein gene family. J. Mammary Gland Biol. Neoplasia. 7: 327-345. DOI: 10.1023/A:1022808918013

Sambrook, J., E.F. Fritschi, T. Maniatis. 1989. Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, New York.

Schennink, A., W.M. Stoop, M.H.P.W. Visker, J.M.L. Heck, H. Bovenhuis, J.J. van der Poel, H.J.F. van Valenberg, J.A.M. van Arendonk. 2007. *DGAT1* underlies large genetic variation in milk-fat composition of dairy cows. Anim. Genet. 38: 467-473. DOI: 10.1111/j.1365-2052.2007.01635.x

Schennink, A., J.M.L. Heck, H. Bovenhuis, M.H.P.W. Visker, H.J.F. van Valenberg, J.A.M. van Arendonk. 2008. Milk fatty acid unsaturation: genetic parameters and effects of stearoyl-CoA desaturase (SCD1) and acyl CoA: diacylglycerol acyltransferase 1 (*DGAT1*). J. Dairy Sci. 91: 2135-2143. DOI: 10.3168/jds.2007-0825

Shi, D.S., J. Wang, Y. Yang, F.H. Lu, X.P. Li, Q.Y. Liu. 2012. *DGAT1, GH*, *GHR*, *PRL* and *PRLR* polymorphism in water buffalo (*Bubalus bubalis*). Reprod. Dom. Anim. 47: 328-334. DOI: 10.1111/j.1439-0531.2011.01876.x

Shi, M., H. Zhang, L. Wang, C. Zhu, K. Sheng, Y. Du, K. Wang, A. Dias, S. Chen, M. Whitman, E. Wang, R. Reed, H. Cheng. 2015. Premature termination codons are recognized in the nucleus in a reading-frame-dependent manner. Cell Discovery. 1: 15001. DOI: 10.1038/celldisc.2015.1

Silva, C.S., E.F. Silva, A.S. Matos, A.S. Schierholt, M.R. Costa, L.C. Marques, J.S. Costa, R.L. Sales, M.R. Gigueiro, J.R. Marques. 2016. Polymorphisms in the *DGAT1* gene in buffaloes (Bubalus bubalis) in the Amazon. Genet. Mol. Res. 15: 3. DOI: 10.4238/gmr.15038720

Tantia, M.S., R.K. Vijh, B.P. Mishra, B. Mishra, B.S.T. Kumar, M. Sodhi. 2006. *DGAT1* and *ABCG2* polymorphism in Indian cattle (*Bos indicus*) and buffalo (*Bubalus bubalis*) breeds. BMC Vet. Res. 2: 32. DOI: 10.1186/1746-6148-2-32

Thaller, G., C. Kühn, A. Winter, G. Ewlad, O. Bellmann, J. Wegner, H. Zuhlke, R. Fries. 2003. *DGAT1*, a new positional and functional candidate gene for intramuscular fat deposition in cattle. Anim. Genet. 34: 354-357. DOI: 10.1046/j.1365-2052.2003.01011.x

Venkatachalapathy, R.T., A. Sharma, S. Sukla, T.K. Bhattacharya. 2008. Cloning and characterization of *DGAT1* gene of riverine buffalo. DNA Seq. 19: 177-184. DOI: 10.1080/10425170701461748

Winter, A., W. Kramer, F.A.O. Werner, S. Kollers, S. Kata, G. Durstewitz, J. Buitkamp, J.E. Womack, G. Thaller, R. Fries. 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. Proc. Natl. Acad. Sci. USA. 99: 9300-9305. DOI: 10.1073/pnas.142293799

Zhang, J., D. Xu, J. Nie, J. Cao, Y. Zhai, D. Tong, Y. Shi. 2014. Monoacylglycerol Acyltransferase-2 is a tetrameric enzyme that selectively heterodimerizes with Diacylglycerol Acyltransferase-1. J. Biol. Chem. 289: 10909-10918. DOI: 10.1074/jbc.M113.530022

Amplicon	Region amplified	Primers	Sequence	T_a °C	Size
1	Exon 1	DGAT Ex 1 F*	5’-ATGGGCGACCGCGGCGG-3’	69.0	171 bp
		DGAT Ex 1 R	5’-GCCCACGTCTACGTCTCCGTA-3’		
2	Exon 2 - Exon 5	DGAT Ex 2 F	5’-ACCGCCTGCAGGATTC-3’	59.5	2355 bp
		DGAT Ex 5 R	5’-CAGCCACGGCAAAGATATTG-3’		
3	Exon 5 - Exon 9	DGAT Ex 5 F	5’-CAATATCTTTGCCGTGG-3’	59.0	855 bp
		DGAT Ex 9 R	5’-AGCTCGTAGCACAGGG-3’		
4	Exon 8 - Exon 14	DGAT Ex 8 F	5’-CCCCGACAACCTGACC-3’	61.5	842 bp
		DGAT Ex 14 R	5’-GTAGGTGATGGACTCGG-3’		
5	Exon 11 - Exon 17	DGAT Ex 11 F	5’-TTCACTTCTTTTCCTCCAC-3’	60.5	1021 bp
		DGAT Ex 17 R*	5’- TCAGGTGCCGGCTGCCGG-3’		

Table 1. Primer sequences, annealing temperature (T_a) and amplicon size used for the genetic diversity discovery/confirmation at the river buffalo *DGAT1*. Asterisks refer to primers designed on river buffalo *DGAT1* cDNA sequence (EMBL ID: DQ120929), whereas all the other primers were designed on genomic DNA sequence (EMBL ID: AY999090).

	N. of clones (%)		**Transcript size (bp)**	**Protein size (aa)**	**PTC**	**Rearrangement**
	Low FP	**High FP**				
1	78 (47.56)	94 (54.65)	1470	489		Correctly assembled
2	40 (24.39)	22 (12.80)	1425	474		Del ex 12
3	8 (4.88)	14 (8.14)	1404	467		Del of the last 66 bp of the ex 8
4	10 (6.09)	4 (2.33)	1407	468		Del ex 16
5	6 (3.66)	10 (5.81)	1557	518		Ins int 13
6	2 (1.22)	6 (3.49)	1337	168	*	Ins int 13 + Del ex 6 and ex 7
7	6 (3.66)	-	1359	452		Del of the last 66 bp of the ex 8 + Del ex 12
8	-	6 (3.49)	1362	453		Del ex 12 + ex 16
9	4 (2.44)	-	1570	263	*	Ins int 7
10	-	4 (2.33)	1491	496		Del of the last 66 bp of the ex 8 + Ins int 13
11	2 (1.22)	-	1341	446		Del of the last 66 bp of the ex 8 + Del ex 16
12	2 (1.22)	-	1296	431		Del of the last 66 bp of the ex 8 + Del ex 12 and ex16
13	2 (1.22)	-	1284	107	*	Ins int 3 + Del ex 6, ex 7 and ex 12
14	2 (1.22)	-	1452	483		Del of the last 18 bp of the ex 10
15	2 (1.22)	-	1389	462		Del of the first 81 bp of the ex 13
16	-	2 (1.16)	1620	539		Del ex 6 and the first 47 bp of the ex 7 + Ins int 8, int 12, int 13 and int 14
17	-	2 (1.16)	1476	491		Del of the last 66 bp of the ex 8 + Ins int 14
18	-	2 (1.16)	1225	400	*	Del of the last 66 bp of the ex 8 + Del ex 13 and ex14
19	-	2 (1.16)	1108	107	*	Ins int 3 + Del ex 6, ex 7, ex 12, ex 13 and ex 16
20	-	2 (1.16)	1710	107	*	Ins int 3, int 12 and int 13
21	-	2 (1.16)	1391	263	*	Ins int 7 + Del ex 13 and ex 14
Tot	**164** (100)	**172** (100)				

Table 2. Absolute and relative frequencies of clones carriers of *DGAT1* mRNA populations in two group of lactating buffaloes ranked for milk FP and divided in two groups (high and low) for fat production. Transcript and predicted protein size, Premature Termination Codon (PTC) and rearrangement events observed. For the clone number distribution (from 1 to 6): $\chi2$ =13.805, P=0.0169.

Breed	**Genotyped SNP c.1053C>T**			**N. of animals**	**Allele frequency**		**HW equilibrium**
	CC	CT	TT		C	T	$\chi2$ (P-value)
Italian Mediterranean	0.200	0.520	0.280	100	0.460	0.540	0.218 (0.640)
Romanian Murrah	0.260	0.530	0.210	100	0.525	0.475	0.392 (0.531)
Total buffalo population	0.230	0.525	0.245	200	0.493	0.507	0.504 (0.477)

Table 3. Genotyping data, allele frequencies, Hardy-Weinberg equilibrium ($\chi2$ test – $P<0.05$ – d.o.f = 1) for the SNP c.1053C>T at the *DGAT1* gene in two different river buffaloes breeds. The two breeds showed statistically different frequencies ($P=0.002$).

Position	**Nucleotide Genomic DNA**	**Nucleotide cDNA**	**A**	**B**	**C**	**D**
Exon 1	3593	89	*T*$^{\text{Val}}$	C$^{\text{Ala}}$	C$^{\text{Ala}}$	C$^{\text{Ala}}$
	3614	110	*C*$^{\text{Ala}}$	T$^{\text{Val}}$	T$^{\text{Val}}$	T$^{\text{Val}}$
Exon 2	7373	242	A$^{\text{Asn}}$	A$^{\text{Asn}}$	A$^{\text{Asn}}$	G$^{\text{Ser}}$
	7380	249	T$^{\text{Arg}}$	T$^{\text{Arg}}$	C$^{\text{Arg}}$	C$^{\text{Arg}}$
	7406	275	A$^{\text{Lys}}$	A$^{\text{Lys}}$	T$^{\text{Met}}$	T$^{\text{Met}}$
Intron 2	8921		C	T	Y	*
	8930/8931		C	-	C	*
	9011		-	A	-	*
	9019		-	C	-	*
	9056-9057		GC	CT	GC	*
	9079		G	C	G	*
	9083		-	T	-	*
	9095		G	A	G	*
	9100		G	A	G	*
	9111		G	A	G	*
	9113		-	C	-	*
	9175-9176		**-R**	AG	-G	*
	9180		-	T	-	*
	9188		**Y**	C	C	*
	9196		G	A	G	*
	9208-9209		**-R**	CG	-A	*
Intron 3	9460		**Y**	C	C	*
Exon 4	9529	369	G$^{\text{Lys}}$	G$^{\text{Lys}}$	G$^{\text{Lys}}$	A$^{\text{Lys}}$
Exon 5	9694	442	G$^{\text{Glu}}$	A$^{\text{Lys}}$	G$^{\text{Glu}}$	G$^{\text{Glu}}$
Intron 6	10109		**Y**	C	C	*
Exon 7	10158	608	T$^{\text{Leu}}$	T$^{\text{Leu}}$	T$^{\text{Leu}}$	C$^{\text{Pro}}$
Intron 11	10874		**W**	T	T	
Exon 13	11057	1040	T$^{\text{Leu}}$	T$^{\text{Leu}}$	T$^{\text{Leu}}$	C$^{\text{Pro}}$
	11070	1053	**Y**$^{\text{Ala/Ala}}$	T$^{\text{Ala}}$	C$^{\text{Ala}}$	C$^{\text{Ala}}$
	11100	1083	C$^{\text{Phe}}$	A$^{\text{Leu}}$	C$^{\text{Phe}}$	C$^{\text{Phe}}$
	11109	1092	C$^{\text{Asp}}$	A$^{\text{Glu}}$	C$^{\text{Asp}}$	C$^{\text{Asp}}$
Intron 13	11125		C	T	C	*
	11133		C	G	C	*
	11134		**M**	-	C	*
	11140-11141		CA	TG	CA	*
	11169		T	G	T	*
	11186-11187		CA	TG	CA	*
Exon 14	11214	1110	T$^{\text{Ser}}$	C$^{\text{Ser}}$	T$^{\text{Ser}}$	T$^{\text{Ser}}$
	11231	1127	A$^{\text{Gln}}$	A$^{\text{Gln}}$	A$^{\text{Gln}}$	T$^{\text{Leu}}$
Exon 15	11377	1201	A$^{\text{Lys}}$	G$^{\text{Glu}}$	A$^{\text{Lys}}$	A$^{\text{Lys}}$
	11426	1250	A$^{\text{Glu}}$	A$^{\text{Glu}}$	A$^{\text{Glu}}$	G$^{\text{Gly}}$
Intron 16	11618		**R**	G	R	*
	11634		**R**	G	G	*
Exon 17	11726	1392	C$^{\text{Ile}}$	C$^{\text{Ile}}$	C$^{\text{Ile}}$	T$^{\text{Ile}}$

A: Present study; B: Yuan et al. (2007) - EMBL ID: AY999090; C: Mishra et al., 2007 - EMBL ID: DQ886485; D: Venkatachalapathy et al., (2008) EMBL ID: DQ120929

Table 4. Polymorphisms detected by the comparison among Mediterranean river buffalo *DGAT1* gene sequence of the present study (**A**) with the Chinese water buffalo *DGAT1* sequence reported by Yuan et al. (2007) (**B**), the *DGAT1* gene of Indian water buffaloes representative of the breeds Murrah, Bhadawari, Tarai, Pandharpuri, Marathwada and Mehsana and reported by Mishra et al. (2007) (**C**), and the Indian *DGAT1* cDNA reported by Venkatachalapathy et al. (2008) (**D**). Mutations detected in the investigated samples (Y=C/T, R=A/G, W=A/T, M=A/C) are reported in bold. Grey cells identify nucleotides identical to the sequence of the present study. Dashes indicate deleted nucleotides, asterisks show unavailable sequences. Nucleotides in italics are typical of the Italian Mediterranean breed. Numbering of genomic DNA and cDNA is relative to the sequences AY999090 and DQ120929, respectively.

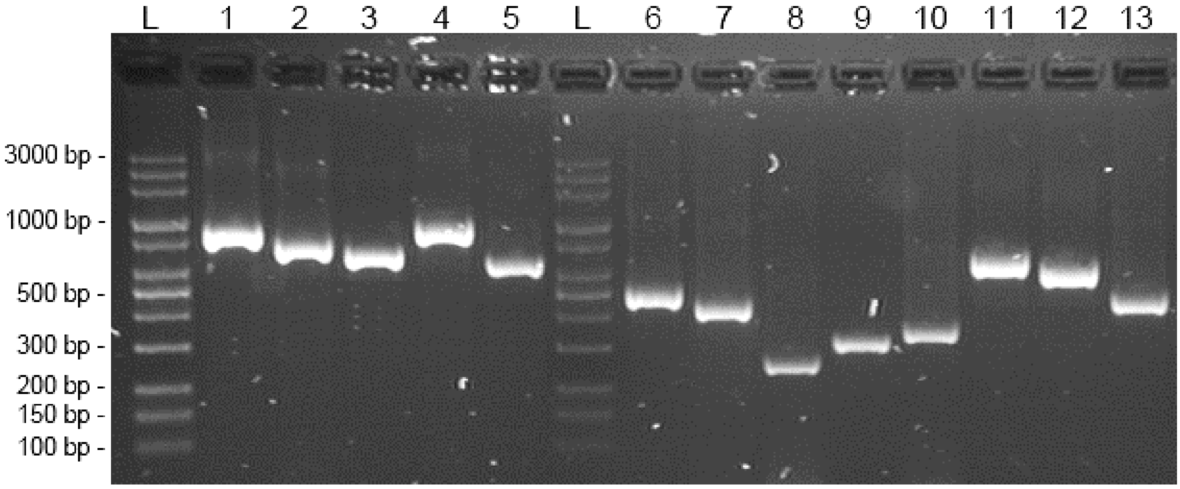

Figure 1. Electrophoretic pattern of the most prevalent transcripts for the river buffalo *DGAT1* in the mRNA regions from the exon 1 to the exon 11 (lines 1-5) and from the exon 11 to the exon 17. Lines 1 and 6, transcripts correctly assembled. Line 2, mRNA spliced out of the last 66bp of the exon 8. Line 3, skipping of the exon 6 plus the first 45 bp of exon 7 and insertion of the intron 8. Line 4, insertion of the intron 7 and splicing of the last 66bp of the exon 8. Line 5, insertion of the intron 3 and skipping of the exons 6 and 7. Line 7, transcript spliced out of the exon 12. Line 8, mRNA skipped out of the exons 12, 13 and 16. Line 9, skipping of the exons 13 and 14. Line 10, skipping of the exon 16. Line 11, insertion of the introns 12, 13 and 14. Line 12, insertion of intron 13. Line 13, mRNA skipped out of the exons 12 and 16. Line L, Mid Range DNA ladder 100bp-3kb (Jena Bioscience).

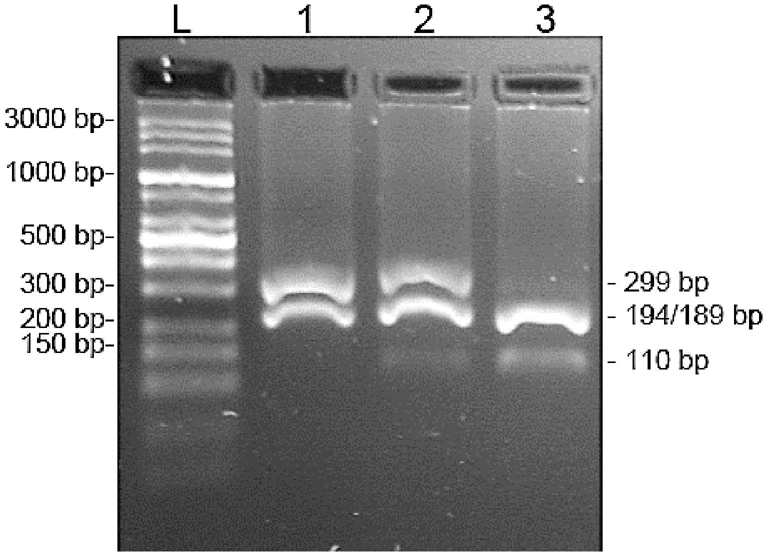

Figure 2. Genotyping of river buffalo *DGAT1* c.1053C>T SNP by *DdeI* I (5’-C↓TNAG-3’) PCR-RFLP. Line 1, CC homozygous samples; line 3, TT homozygous samples; line 2, heterozygous samples. Line L is Mid Range DNA ladder 100bp-3kb (Jena Bioscience).

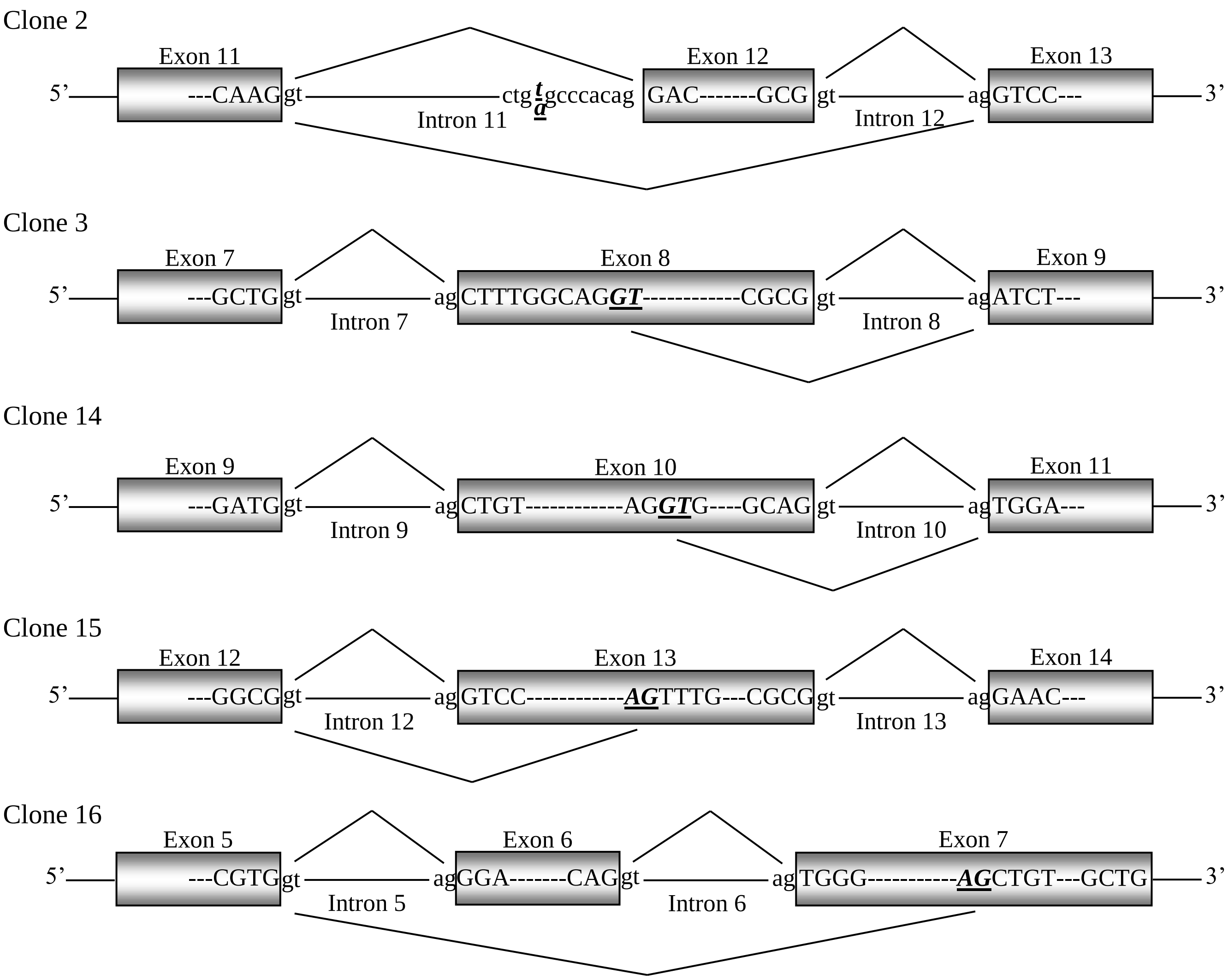

Figure 3. Schematic representations of the exon structures of the river buffalo *DGAT1* gene and the possible splicing combinations, normal (upper) and alternative (down).