

Technical note: Advantages and limitations of authenticating Palmera goat dairy products by pyrosequencing the melanocortin 1 receptor (MC1R) gene

B. Badaoui,*¹ A. Manunza,*¹ A. Castelló,*† M. D'Andrea,‡ F. Pilla,‡ J. Capote,§ J. Jordana,† A. Ferrando,† A. Martínez,# B. Cabrera,*† J. V. Delgado,# V. Landi,# M. Gómez,|| A. Pons,¶ M. El Ouni,** O. Vidal,†† and M. Amills*†²

*Department of Animal Genetics, Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB),

Campus Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

†Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

‡Department of Agricultural, Environmental and Food Sciences, University of Molise, Via de Sanctis,, 86100 Campobasso, Italy

§Instituto Canario de Investigaciones Agrarias, La Laguna 38108, Tenerife, Spain

#Departamento de Genética, Universidad de Córdoba, Córdoba 14071, Spain

ll Servicio de Ganadería, Diputación Foral de Bizkaia, 48014 Bilbao, Spain

¶Unitat de Races Autòctones, Servei de Millora Agrària, (SEMILLA-SAU), Son Ferriol 07198, Spain

*Livestock & Wildlife Laboratory, Arid Land Institute Medenine, 4119 Médenine, Tunisia

††Departament de Biologia, Universitat de Girona, Girona 17071, Spain

ABSTRACT

Inferring the breed of origin of dairy products can be achieved through molecular analysis of genetic markers with a population-specific pattern of segregation. The goal of the current work was to generate such markers in goats by resequencing several pigmentation genes [melanocortin 1 receptor (MC1R), v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT), tyrosinase (TYR), and tyrosinase-related protein 2 (TYRP2)]. This experiment revealed 10 single nucleotide polymorphisms (SNP), including 5 missense mutations and 1 nonsense mutation. These markers were genotyped in 560 goats from 18 breeds originally from Italy, the Iberian Peninsula, the Canary Islands, and North Africa. Although the majority of SNP segregated at moderate frequencies in all populations (including 2 additional markers that were used as a source of information), we identified a c.764G>A SNP in MC1R that displayed highly divergent allelic frequencies in the Palmera breed compared with the Majorera and Tinerfeña breeds from the Canary Islands. Thus, we optimized a pyrosequencing-based technique that allowed us to estimate, very accurately, the allele frequencies of this marker in complex DNA mixtures from different individuals. Once validated, we applied this method to generating breed-specific DNA profiles that made it possible to detect fraudulent cheeses in which Palmero cheese was manufactured with milk from Majorera goats. One limitation of this approach, however, is that it cannot be used to detect illegal manufacturing where Palmero dairy products are produced by mixing milk from Palmera and Majorera goats, because the c.764G>A SNP segregates in both breeds.

Key words: authentication of dairy products, cheese, goat, pyrosequencing, melanocortin 1 receptor gene (MC1R)

Technical Note

In general, the development of molecular tests to authenticate the breed of origin of dairy products with appellation of origin has been carried out using 2 main approaches (Fontanesi et al., 2011; Nicoloso et al., 2012). The first (probabilistic) approach relies on the inference of breed-specific allele frequency profiles through the genotyping of a moderate to high number of markers (e.g., microsatellites, amplified fragment length polymorphisms, or SNP). Then, maximum likelihood or Bayesian methods can be used to infer the probability that a given sample corresponds to a certain breed. This procedure can be a very powerful tool to assign individuals to specific breeds, but it would be difficult to apply it to the authentication of dairy products derived from multiple individuals (Fontanesi et al., 2011). The alternative (deterministic) approach consists of identifying alleles that segregate specifically in the breed of interest but not in other populations (where, ideally, they should be absent).

The latter method was applied successfully to authenticate dairy products from the ovine Massese breed (Fontanesi et al., 2011). Variability of the melanocortin 1 receptor (MC1R) was analyzed in sheep of different colors. In doing so, several 5' untranslated region (UTR) SNP were identified and one (c.-31G>A) was

Received May 2, 2014. Accepted July 23, 2014.

¹These authors contributed equally to this work.

²Corresponding author: marcel.amills@uab.cat

7294 BADAOUI ET AL.

subsequently genotyped in 14 ovine breeds. Although the A allele was fixed in the Massese breed, it segregated at variable frequency (0.06–0.50) in other sheep populations (Fontanesi et al., 2011). Thus, detection of the G allele in DNA isolated from Massese milk or cheese would indicate contamination with biological products from other breeds (Fontanesi et al., 2011).

In the current work, we wanted to investigate the feasibility of this approach to authenticate dairy products from a wide array of goat breeds. We focused our molecular analysis on the variability of 4 pigmentation genes; namely, MC1R, v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT), tyrosinase (TYR), and tyrosinase-related protein 2 (TYRP2), because their variation has been modulated not only by drift but also by artificial selection. In theory, this feature should maximize the probability of finding alleles with strong differences in allele frequencies among breeds. As a complementary source of information, previously reported data for the agouti signaling protein (ASIP) and tyrosinase-related protein 1 (TYRP1) genes (Badaoui et al., 2011, 2012) were also used in the current study.

Genomic DNA was extracted from hair follicles by using the DNeasy Blood and Tissue kit (Qiagen, Barcelona, Spain). To purify DNA from blood samples, we followed the protocols described by Zidi et al. (2010). Genomic DNA concentration was estimated by using a Nanodrop spectrophotometer ND-1000 (SG Servicios Hospitalarios, Barcelona, Spain). We also isolated total RNA, from either skin or liver samples, by using the RiboPure kit (Ambion, Life Technologies, Madrid, Spain) according to the instructions of the manufacturer. After assessing RNA quality with an Agilent 2100 Bioanalyzer equipment (Agilent Technologies, Barcelona, Spain), reverse transcription (RT) reactions were carried out using the Thermoscript RT-PCR System kit (Invitrogen, Barcelona, Spain) and following the protocols reported by Amills et al. (2003).

Although both genomic DNA and cDNA were used as templates in sequencing experiments concerning the *MC1R* gene, analysis of the variability of the *KIT*, *TYRP2*, and *TYR* genes was based exclusively on cDNA-amplified sequences. More specifically, we amplified and sequenced approximately 0.8, 3.5, 0.7, and 1.3 kb of the goat *MC1R* (GenBank accession number: JX444448), *KIT* (JX444449), *TYRP2* (JX444450), and *TYR* (JX444451) genes, respectively. Sequencing panels contained around 10 breeds (Malagueña, Murciano-Granadina, Saanen, Palmera, Majorera, Tinerfeña, Garganica, Cilentana Nera, Girgentana, and Payoya) when genomic DNA was used as template, and 6 breeds (Malagueña, Murciano-Granadina, Palmera, Majorera, Tinerfeña, and Payoya) when cDNA was used. In all

PCR, the composition of the amplification reactions was $1.5 \text{ m} M \text{ MgCl}_2$, $200 \mu M$ of each deoxynucleotide triphosphate (dNTP), $0.2 \mu M$ of each primer (Supplementary Table S1; http://dx.doi.org/10.3168/jds.2014-8316), 50 to 80 ng of genomic DNA (MC1R amplification) or 1.5 µL of cDNA (MC1R, KIT, TYRP2, and TYR amplification), and 0.5 U of BioTaq DNA polymerase (Bioline, Barcelona, Spain) in a final 20-µL volume. The thermal profile used in the amplification of the MC1R gene consisted of 35 cycles of 94°C for 45 s, annealing temperature (Supplementary Table S1) for 45 s, and 72°C for 1 min. Amplification of the KIT, TYRP2, and TYR genes was based on a thermal profile including 35 cycles of 94°C for 1 min, annealing temperature (Supplementary Table S1) for 1 min, and 72°C for 1.5 min. The BigDye Terminator v1.3 Cycle Sequencing Kit (Applied Biosystems, Sant Andreu de Llavaneres, Spain) was used to sequence the amplified fragments. The primers used for sequencing are listed in Supplementary Table S1. The coverage of the coding sequences was 100% (KIT), 83% (MC1R), 83% (TYR), and 50% (TYRP2).

Resequencing experiments revealed the existence of 10 SNP (Table 1), including 5 missense mutations (in MC1R, KIT, and TYR) and 1 nonsense mutation (in MC1R). Overall, the gene that showed the highest level of variability was MC1R. As previously seen in humans (Rana et al., 1999; Gerstenblith et al., 2007), most of the variation in MC1R was nonsynonymous. Two of the SNP found in the current study (MC1Rc.673C>T and c.748G>T) were previously identified by Fontanesi et al. (2009) in an independent sample of goats, thus confirming their widespread segregation in caprine breeds. To infer if any of these SNP displayed differential allele frequencies among breeds, we genotyped 560 individuals with a Sequenom MassARRAY iPLEX platform at the Spanish National Genotyping Centre (CeGen, Santiago de Compostela, Spain). The 560 analyzed goats were distributed in 18 populations covering 5 geographic areas: Italy (n = 122): Cilentana Nera (n = 26), Garganica (n = 41), Grigia Molisana (n = 41)= 13), Derivata di Siria (n = 9), Girgentana (n = 19), Jonica (n = 7), and Maltese (n = 7); Iberian Peninsula and Balearic Islands (n = 175): Malagueña (n = 43), Murciano-Granadina (n = 81), Azpi-Gorri (n = 13), Eivissenca (n = 12), and Rasquera (n = 26); Canary Islands (n = 92): Tinerfeña (n = 38), Palmera (n = 38) and Majorera (n = 16); North Africa (n = 117): Moroccan (n = 32) and Tunisian (n = 85); Switzerland (n = 54): Saanen (n = 54).

Genotypic frequencies of SNP mapping to pigmentation loci are shown in Supplementary Table S2 (http://dx.doi.org/10.3168/jds.2014-8316). For breeds with a small sample size (n < 20), estimates of genotypic

Table 1. Polymorphism of 4 goat pigmentation genes

Gene ¹	Polymorphism	AA substitution
KIT	c.1284T>G c.1887T>C	H428Q
MC1R	c.673C>T c.676A>G c.748G>T	Q225X K226E V250F
TYR	c.764G>A c.769C>A c.1455A>G	G255D P257T —
TYRP2	c.900T>C c.996C>T	

 1KIT = v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog; MC1R = melanocortin 1 receptor; TYR = tyrosinase; TYRP2 = tyrosinase-related protein 2.

frequencies should be considered as preliminary and hence interpreted with caution. Expected heterozygosities $(\mathbf{H_e})$ were calculated according to the formula $1 - \sum p_i^2$, where p_i is the frequency of the *i*th of k alleles. This method is implemented in the R package Genetics (http://cran.r-project.org/web/packages/ genetics/index.html) by using the function genotype. As shown in Supplementary Table S2, average H_e was 0.27 and ranged from 0.01 for the MC1R c.764G>A polymorphism to 0.44 for the TYR (c.769C>A) and TYRP2 (c.996C>T) loci. These values match those observed in previous studies. In this way, Pariset et al. (2009) reported H_e values ranging from 0.20 to 0.34 for 26 autosomal markers analyzed in Greek, Albanian, and Italian caprine breeds, whereas Nicoloso et al. (2012) studied the segregation of 29 SNP in 20 pigmentation genes and described H_e values that ranged between 0.25 and 0.29.

We used the population assignment method implemented in GenAlEx v6.41 (Peakall and Smouse, 2006) to infer the likelihood that a certain individual comes (or not) from a specific population. As shown in Supplementary Table S3 (http://dx.doi.org/10.3168/ jds.2014-8316), the probabilities of correctly assigning individuals to their corresponding breeds were generally low to moderate. The only exceptions were in the Girgentana breed, where this probability was around 90%, and in Palmera goats, where all individuals were correctly assigned. Indeed, close inspection of data reported in Supplementary Table S2 revealed a very distinctive allele at the MC1R SNP c.764G>A for the Palmera breed; that is, the frequency of the A allele was 1 in the Palmera breed and 0.12 and 0.05 in the Majorera and Tinerfeña breeds, respectively. In the remaining populations, the frequency of the A allele was null. We did not detect any other SNP with a pattern of breed-specific segregation, although it is worth noting that the T-allele of the MC1R c.673C>T polymorphism segregated exclusively in Italian breeds.

Similarly, Nicoloso et al. (2012) partially resequenced 36 pigmentation genes in 8 goats from 4 breeds and did not detect any breed-specific allele, although the frequencies of certain SNP were remarkably uneven among breeds.

Our finding that the A allele has a high frequency in the Palmera breed (whereas its frequency in Majorera and Tinerfeña breeds happens to be very low) is intriguing because these 3 breeds are assumed to have a common ancestry that traces back to the settlement of the Canary Islands by Imazighen tribes 2,500 YBP (years before present; Fregel et al., 2009). Our observation might be explained by (1) a founder effect; that is, the island of La Palma might have been populated by a few goats that, by chance, were AA; or (2) diversifying selection; that is, the A allele reached a high frequency because it was associated with a coat color that La Palma pre-Hispanic natives selected for.

Our discovery that the MC1R c.764 A allele segregates almost exclusively in the Palmera population allowed us to implement a molecular test aimed to authenticate dairy products from this breed. Palmero cheese has great economic and cultural importance in the agroindustry of the Canary Islands and it is protected with an appellation of origin (Freitas and Malcata, 2000). One of the commonest frauds is to produce Palmero cheese with milk from Majorera goats, which are more productive (in terms of milk yield) and better adapted to intensive management than their Palmera counterparts (J. Capote, personal observation). Thus, it was critical to devise a technique to allow discrimination of cheese samples from these 2 closely related breeds. With this aim, we developed a pyrosequencing protocol (Ronaghi et al., 1998) to estimate the allele frequencies of the MC1R SNP c.764 G>A. The main advantage of pyrosequencing is that it allows the processing of large batches of samples rapidly and reliably. A 94-bp fragment of the goat MC1R gene containing the SNP was PCR-amplified according to the following conditions: $2.5 \text{ m} M \text{ MgCl}_2$, $250 \mu M \text{ dNTP}$, $0.3 \mu M \text{ biotin-labeled}$ forward (5'-ATC AGG GCT TTG GCC TCA A-3') and reverse (5'-AGC GAG AGG TGC AGA AAG AAG G-3') primers, 1 U of AmpliTag Gold DNA polymerase (Applied Biosystems, Life Technologies), and 75 ng of genomic DNA in a final volume of 20 μ L. The thermal profile was 95°C for 10 min, 40 cycles at 95°C for 1 min, 62°C for 1 min, and 72°C for 1 min, with a final extension step of 7 min at 72°C. The biotinylated PCR product was prepared for pyrosequencing analysis according to the manufacturer's instructions (Qiagen). Pyrosequencing reactions included 18 µL of amplicon, the sequencing primer 5'-GGT GCA GAA AGA AGG G-3' at 0.3 μM, and PyroMark Gold Q96 SQA reagents (Qiagen). Reactions were run in aPSQ HS96 system 7296 BADAOUI ET AL.

pyrosequencing equipment (Qiagen) and data analysis was carried out by using the PSQ HS96A 1.2 software (Qiagen).

We typed some additional individuals from the Majorera (n = 6), Palmera (n = 11), and Tinerfeña (n = 10) breeds, and these genotyping data were combined with that generated using the Sequenom platform to obtain final estimates of the G and A allele frequencies, which were as follows: Palmera (7% G, 93% A), Majorera (91% G, 9% A), and Tinerfeña (98% G, 2% A). Then, we prepared pools of genomic DNA of GG and AA homozygous goats at proportions that recapitulated the allelic frequencies shown above for the MC1R c.764G>A polymorphism in the Palmera and Majorera breeds (frequencies in Tinerfeña and Majorera were very similar, so a third Tinerfeña pool was unnecessary).

Next, we tested our ability to differentiate pools by pyrosequencing the MC1R region containing the c.764G>A SNP. Given that interpretation of pyrograms is not trivial, we have provided a detailed explanation in Supplementary Figure S1 (http://dx.doi. org/10.3168/jds.2014-8316; in the graph, alleles G and A are named C and T because pyrosequencing was done on the reverse chain). As can be seen in Supplementary Figure S2, dramatic differences were found in the pyrograms obtained for the Majorera and Palmera breeds. Once we confirmed that the pyrosequencing technique allowed us to obtain reliable estimates of the c.764G>A allele frequencies, we applied it to the analysis of 3 cheese samples from Majorera, Tinerfeña, and Palmera goats. Genomic DNA was extracted from cheese by using the NucleoSpin Food kit (Macherey Nagel, Düren, Germany), and the MC1R region containing the polymorphic site was amplified as described above. As shown in Figure 1, pyrograms from Palmera compared with Majorera and Tinerfeña were clearly different, allowing us to detect fraudulent manufacture, in which Palmero cheese was produced with milk from other unrelated breeds (mainly Majorera). It can also be observed that the allele frequencies inferred from the Palmera cheese sample were different from those reported in Supplementary Table S2 (http://dx.doi. org/10.3168/jds.2014-8316), meaning that the A allele is not completely fixed (although its frequency is much higher than in any of the tested breeds). This means that, although we developed a tool to infer the origin (Palmera vs. Majorera and Tinerfeña) of cheese samples on the basis of pyrogram DNA profiles, it would be very difficult to determine if these cheeses are pure or if they were manufactured by mixing milk from 2 or more Canarian breeds. This limitation can be attributed to the fact that c.764G>A allele frequencies in Palmera goats may vary from farm to farm, making it impos-

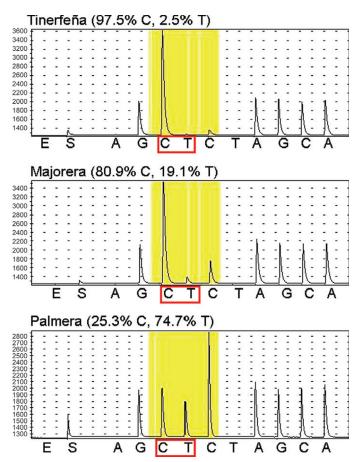


Figure 1. Pyrograms of melanocortin 1 receptor (MC1R) PCR products amplified from genomic DNA purified from Tinerfeña, Majorera, and Palmera cheese samples. The polymorphic position is boxed (red). Pyrogram interpretation is explained in Supplemental Figure S1 (http://dx.doi.org/10.3168/jds.2014-8316). Dramatic differences in the allelic frequencies of the c.764 G>A SNP between Palmera and the Majorera and Tinerfeña breeds were detected by the pyrosequencing technique and could be used to authenticate Palmera dairy products. The y-axis shows the relative light units as an indicator of nucleotide incorporation, and the x-axis is the programmed nucleotide dispensation order. Peaks are proportional to the amount of the respective nucleotide incorporated. E = enzyme-only control; S = substrate-only control. Color version available in the online PDF.

sible to discern contamination from the consequences of genetic drift.

In a previous paper (Amills et al., 2004), we developed a technique based on mitochondrial DNA that allowed us to determine whether dairy products had a Canarian origin. Herein, we optimized a pyrosequencing protocol to discern the origin of dairy products at a finer level; that is, the pyrograms obtained from Palmera products were very different from those generated in 2 closely related breeds (Majorera and Tinerfeña). Given that pyrosequencing-based methods allow processing of samples at a moderate to large scale, this approach may be used by the Palmero Cheese Appellation of Origin Regulatory Council to detect fraudulent production of

Palmero cheese with Majorera milk, which is forbidden by law. From a technical point of view, our findings reflect the high resolution and reliability of pyrosequencing in providing estimates of the frequencies of allele copies in a complex mixture of different individuals, a feature that illustrates the wide array of applications for this technique in the field of food authentication. Another general conclusion that can be derived from our work is that, as discussed by Fontanesi et al. (2011), it is very uncommon to find genetic variants that are fixed in one breed and completely absent from other breeds because of extensive gene flow and recurrent introgression events among goat breeds. As a whole, the picture that Fontanesi et al. (2011) obtained in sheep—that breed-specific alleles are scarce—agrees well with data presented in the current paper (Supplementary Table S2; http://dx.doi.org/10.3168/jds.2014-8316). In light of this, the implementation of deterministic DNA tests to detect fraud in which milks from distinct breeds are mixed should be based on generating large data sets of thousands of SNP (for each targeted breed) through next-generation sequencing and identifying a subset of tens or hundreds of markers with a breed-specific pattern of segregation.

ACKNOWLEDGMENTS

This research was partially funded by a grant (RZ2007-00005-C02-01) from the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (Spain).

REFERENCES

- Amills, M., J. Capote, A. Tomàs, L. Kelly, G. Obexer-Ruff, A. Angiolillo, and A. Sànchez. 2004. Strong phylogeographic relationships among three goat breeds from the Canary Islands. J. Dairy Res. 71:257–262.
- Amills, M., N. Jiménez, D. Villalba, M. Tor, E. Molina, D. Cubiló, C. Marcos, A. Francesch, A. Sànchez, and J. Estany. 2003. Identification of three single nucleotide polymorphisms in the chicken insulin-like growth factor 1 and 2 genes and their associations with growth and feeding traits. Poult. Sci. 82:1485–1493.

- Badaoui, B., M. D'Andrea, F. Pilla, J. Capote, A. Zidi, J. Jordana, A. Ferrando, J. V. Delgado, A. Martínez, O. Vidal, and M. Amills. 2011. Polymorphism of the goat agouti signaling protein gene and its relationship with coat color in Italian and Spanish breeds. Biochem. Genet. 49:523–532.
- Badaoui, B., A. Manunza, M. D'Andrea, F. Pilla, J. Capote, J. Jordana, A. Ferrando, J. V. Delgado, A. Martínez, O. Vidal, and M. Amills. 2012. Identification of c.483C>T polymorphism in the caprine tyrosinase-related protein 1 (TYRP1) gene. Ital. J. Anim. Sci. 11:63-67.
- Fontanesi, L., F. Beretti, S. Dall'Olio, B. Portolano, D. Matassino, and V. Russo. 2011. A melanocortin 1 receptor (MC1R) gene polymorphism is useful for authentication of Massese sheep dairy products. J. Dairy Res. 78:122–128.
- Fontanesi, L., F. Beretti, V. Riggio, S. Dall'Olio, E. G. González, R. Finocchiaro, R. Davoli, V. Russo, and B. Portolano. 2009. Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: Association with red and black coat colour phenotypes but with unexpected evidences. BMC Genet. 10:47.
- Fregel, R., V. Gomes, L. Gusmão, A. M. González, V. M. Cabrera, A. Amorim, and J. M. Larruga. 2009. Demographic history of Canary Islands male gene-pool: Replacement of native lineages by European. BMC Evol. Biol. 9:181.
- Freitas, C., and F. X. Malcata. 2000. Microbiology and biochemistry of cheeses with Appélation d'Origine Protegée and manufactured in the Iberian Peninsula from ovine and caprine milks. J. Dairy Sci. 83:584–602.
- Gerstenblith, M. R., A. M. Goldstein, M. C. Fargnoli, K. Peris, and M. T. Landi. 2007. Comprehensive evaluation of allele frequency differences of MC1R variants across populations. Hum. Mutat. 28:495–505.
- Nicoloso, L., R. Negrini, P. Ajmone-Marsan, and P. Crepaldi. 2012. On the way to functional agro biodiversity: Coat colour gene variability in goats. Animal 6:41–49.
- Pariset, L., A. Cuteri, C. Ligda, P. Ajmone-Marsan, and A. Valentini., and ECONOGENE Consortium. 2009. Geographical patterning of sixteen goat breeds from Italy, Albania and Greece assessed by single nucleotide polymorphisms. BMC Ecol. 9:20.
- Peakall, R., and P. E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6:288–295.
- Rana, B. K., D. Hewett-Emmett, L. Jin, B. H. Chang, N. Sambuughin, M. Lin, S. Watkins, M. Bamshad, L. B. Jorde, M. Ramsay, T. Jenkins, and W. H. Li. 1999. High polymorphism at the human melanocortin 1 receptor locus. Genetics 151:1547–1557.
- Ronaghi, M., M. Uhlen, and P. Nyren. 1998. A sequencing method based on real-time pyrophosphate. Science 281:363–365.
- Zidi, A., V. M. Fernández-Cabanás, B. Urrutia, J. Carrizosa, O. Polvillo, P. González-Redondo, J. Jordana, D. Gallardo, M. Amills, and J. M. Serradilla. 2010. Association between the polymorphism of the goat stearoyl-CoA desaturase 1 (SCD1) gene and milk fatty acid composition in Murciano-Granadina goats. J. Dairy Sci. 93:4332–4339.