

Polymorphism in the Beta Casein Gene and analysis of milk characteristics in Gir and Guzerá dairy cattle

A.H.N. Rangel¹, L.G. Zaros¹, T.C. Lima¹, L.H.F. Borba¹, L.P. Novaes¹,
L.F.M. Mota² and M.S. Silva³

¹Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil

²Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP, Brasil

³Programa de Doutorado Integrado em Zootecnia, Universidade Federal do Ceará, Fortaleza, CE, Brasil

Corresponding author: M.S. Silva

E-mail: mig_871@hotmail.com

Genet. Mol. Res. 16 (2): gmr16029592

Received December 20, 2016

Accepted April 24, 2017

Published May 25, 2017

DOI <http://dx.doi.org/10.4238/gmr16029592>

Copyright © 2017 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. Several human health problems have been related to the allergenic constitution of bovine milk due to the body's immune reaction to milk proteins. It is necessary find solutions to minimize the occurrence of such reactions, given the importance of milk as a source of animal protein. The aim of this study was to evaluate the allelic frequency of the CSN2 gene and to evaluate differences in the characteristics of Gir and Guzerá bovine milk. One hundred and fifty-six cows were used (68 Gir and 88 Guzerá) from the Felipe Camarão Experimental Station herd of the Agricultural Research Corporation of RN (EMPARN). DNA extractions were carried out from hair follicles of the animals; the gene was then amplified and sequenced in an ABI 3100 automatic sequencer. The obtained sequences were submitted to analysis using the Geneious 5.6.5[®] program. Data were submitted to analysis of variance and to the Tukey-

Kramer test at 5% probability and cluster analyses by main components were performed. Allele frequencies were 98 and 97% for the A2 allele and 0.96 and 0.93% for the genotype A2A2 for Gir and Guzerá, respectively. Gir and Guzerá animals showed differences in protein, lactose, and non-fat dry extract levels. Although correlations between milk yield and the production and some milk components are moderate, increases in milk yield are always greater than the increase in constituent yield. In addition, even though Guzerá animals have a higher percentage of protein, lactose, and non-fat dry extract, milk from Zebu breeds is an alternative for individuals sensitive to β -casein protein.

Key words: Allergy; CSN2; A2 milk; Gir; Guzerá; Milk yield

INTRODUCTION

Proteins present in milk have different classes that are related to the type of polypeptide chains, with casein, beta-lactoglobulin, and alpha-lactalbumin among them. However, caseins account for about 80% of all milk proteins, and are divided into four groups: alpha-s1 (39 to 46% of all caseins), alpha s2 (8 to 11%), beta (25 to 35%), and kappa (9 to 15%) (Kamiński et al., 2007).

β -casein has about 13 genetic variants, with the A1 and A2 variants being frequently found in dairy cattle (Farrell et al., 2004; Vercesi Filho et al., 2012). The existing difference between the A1 and A2 variants of β -casein is caused by mutation of the A2 allele to the A1 allele at position 67 (histidine A1 and proline A2), being important in the variation of protein content and in the composition of milk protein (Bonfatti et al., 2010). Casein contains proteins that transform into opiate compounds called β -casomorphins (BCM). BCM binds to the A1 allele of the β -casein and it is believed that an ingestion of milk containing presence of this allele causes allergy and other diseases in the human body. In contrast, the A2 allele of β -casein has no link to such health problems (Woodford, 2008).

Some authors have noted that the A2 allele of the β -casein (CSN2) has a positive association between milk yield and protein, and that some dairy breeds (Guernsey, Jersey, Holstein, Gir, Guzerá) present a greater frequency of this allele (Kamiński et al., 2007; Olenski et al., 2010). It is also possible to identify frauds in the production of mozzarella cheese from the use of an extension of the CSN2 gene as marker to identify addition of bovine milk to buffalo milk (Otaviano et al., 2008).

Zebu breeds produce A2 milk, wherein the mutation for A1 allele is found at a low frequency (De Noni et al., 2009; Vercesi Filho, 2012). Identifying mutations in the CSN2 gene in Zebu animals can be used as selection criteria in breeding programs in order to direct the breeding of animals so that the herd presents a higher frequency of genes that positively affect this trait. Thus, our aim was to evaluate the allelic frequency of the CSN2 gene and to evaluate the differences in the characteristics of Gir and Guzerá bovine milk.

MATERIAL AND METHODS

Sample collection, extraction, delineation, and sequencing

The study was conducted at the Felipe Camarão Experimental Station, Rockefeller

Farm, belonging to the Agricultural Research Corporation of Rio Grande do Norte (EMPARN), located in São Gonçalo do Amarante - RN (5°47'36" South, 35°19'44" West).

One hundred and fifty-six pure breed animals were used (88 Guzerá and 68 Gir) of both genders, including calves, young ones and adults. The animals were kept in pasture, receiving supplement concentrate. During the rainy season there was availability of *Brachiaria* sp or *Pennisetum purpureum* (cv. Mott) and in the dry season the animals had access to remaining pastures and sorghum or corn silage. Milking was carried out twice a day at 4:00 am and at 4:00 pm.

Hair samples (15 to 25 units) from the cow tail tassel were collected from each animal in order to preserve the bulbs. The collected material was packed into envelopes, identified and sent to the Animal Biotechnology Laboratory - ESALQ/USP for DNA extraction and genotyping. Gir and Guzerá milk production and composition data were obtained on the same day of sampling.

DNA extraction was performed from the hair follicles, following the precipitation method with salt as described by Regitano et al. (2001). The extracted DNA was submitted to PCR and visualized on agarose gel. Some DNA samples had to be re-extracted due to the low amount of DNA obtained. Salman and Laureano (2006) used the same DNA extraction method and claimed that the method proved to be very quick and simple compared to other methods, in addition to its lower cost.

The design of CSN2 primer pairs was previously conducted by Vercesi Filho et al. (2012) and comprises a partial region of intron 6 and exon 7. The GenBank accession numbers were JN051275 (CSN2-A1) and JN051276 (CSN2-A2). The primer sequences used were: F - 5'-CTGGCTTTCAGTAAAGGGCTCAACTG-3' and R - 5'-TGACCCCAATTTCTTAACCAAACCAA-3', generating a fragment of 362 bp.

Five microliters of DNA added to 20 µL reaction mix was used for PCR, being composed of the following reagents: 2.5 µL 10X Buffer (50 mM KCl, 10 mM Tris-HCl, pH 9.0); 0.75 MgCl₂ (50 mM); 0.5 µL dNTP (10 mM); 1.0 µL of each primer (2.0 pmol/µL); 0.1 µL Platinum[®] Taq DNA Polymerase High Fidelity (Invitrogen) enzyme; and Milli-Q water measuring a 25-µL final volume.

The sequencing reaction was performed according to the Big Dye Terminator Cycle Sequencing Ready Reaction Kit[®] protocol (Life Technologies). The following were used for each reaction: Save Money buffer, CSN2 primers (forward and reverse separately), Big Dye Kit, Milli-Q water and DNA. The program used in the sequencing reaction comprised the following steps: initial denaturation at 95°C for 1 min, followed by a new denaturation at 95°C for 15 s, annealing at 50°C for 15 s and extension at 60°C for 2 min, totaling 35 cycles. Two sequencing reactions were prepared for each individual, one using the forward primer and another using the reverse primer.

Samples were initially denatured for 2 min at 95°C; followed by denaturation for 30 s at 95°C; annealing at 60°C for 30 s; extension at 72°C for 30 s, and repeated for 32 reaction cycles. At the end of the cycles, a final extension step at 72°C was performed for 2 min. Specificity of the amplified products was verified on 1% agarose gel using standard low DNA mass ladder (Invitrogen). After amplification and visualization of the fragments, PCR purification products were performed using the PCR purification protocol for sequencing with AGENCOURT[®] AMPURE[®] XP through magnetic beads. The samples were applied on 1% agarose gel along with standard low DNA mass ladder (Invitrogen) in order to verify the concentration and quality of DNA existing in each sample.

The sequencing reaction purification protocol (LifeTech) was used for purification

and the samples were then applied to the ABI PRISM 3100 Genetic Analyser® sequencer (Applied Biosystems).

Nucleotide sequence readings of the amplified fragment for the beta-casein gene (A1 and A2) were aligned and edited. The Geneious 5.6.5® program was used (Drummond et al., 2011) in order to confirm the mutations and obtain the images.

Determination of allelic and genotypic frequencies of the breeds

Allele frequencies x_i for beta-casein alleles i and genotypic frequencies for the genotype ij were obtained using the equations:

$$x_i = \frac{2n_{ii} + \sum n_{ij}}{2n}; x_{ij} = \frac{n_{ij}}{n} \quad (\text{Equation 1})$$

where n_{ii} and n_{ij} correspond to the number of homozygotes and heterozygotes observed in the i allele, respectively; and n corresponds to the number of individuals analyzed.

Using the Hardy-Weinberg theorem, the expected genotypic frequencies at equilibrium were estimated from expanding the binomial:

$$(x_i + x_j)^2 = x_i^2 + 2x_i x_j + x_j^2 \quad (\text{Equation 2})$$

where x_i^2 is the expected frequency of homozygous for allele i ; $2x_i x_j$ is the expected frequency for heterozygotes ij ; and x_j^2 is the expected frequency of homozygous for allele j .

Statistical analysis

Only A2A2 genotypes for the CSN2 gene were used in the analysis to verify the differences in milk yield characteristics for Gir and Guzerá breeds. The effect of breed on milk yield characteristics was compared by the Tukey-Kramer test at 5% probability (PROC MIXED, Statistical Analysis System, version 9.3). The model used for variance analysis includes the effect of breed:

$$y_{ij} = \mu + a_i + e_{ij} \quad (\text{Equation 3})$$

where y_{ij} is the value observed in the characteristic breed i ; μ is the overall mean; a_i is the effect of breed, and e_{ij} is the residue.

The PRINCOMP procedure of the R program (R Development Core Team, 2015) was used to analyze the main components, applying a matrix of sample correlations; the variables are standardized for zero mean and the variance is equal to one. We opted to use a correlation matrix instead of a covariance matrix to mitigate possible discrepancies between the sharp variances and allow for comparisons between the eigenvectors in a component. The solution in implementing the correlation matrix is recommended when the variables are measured at very different scales from each other, since this is equivalent to a matrix of standardized

variables, as the analyzed characteristics consist of milk yield, linear metric evaluations and area, and subjective scores.

RESULTS AND DISCUSSION

The presence of more than one animal with A1 SNP (single nucleotide polymorphism) in the CSN2 gene was detected, where these animals were heterozygous (A1A2), whereas the presence of homozygous animals for the A1 allele was not observed (Figure 1). From the 156 subjects selected for CSN2 gene, only 5.77% of these animals were heterozygous (A1A2), while the majority were homozygous for the A2 allele.

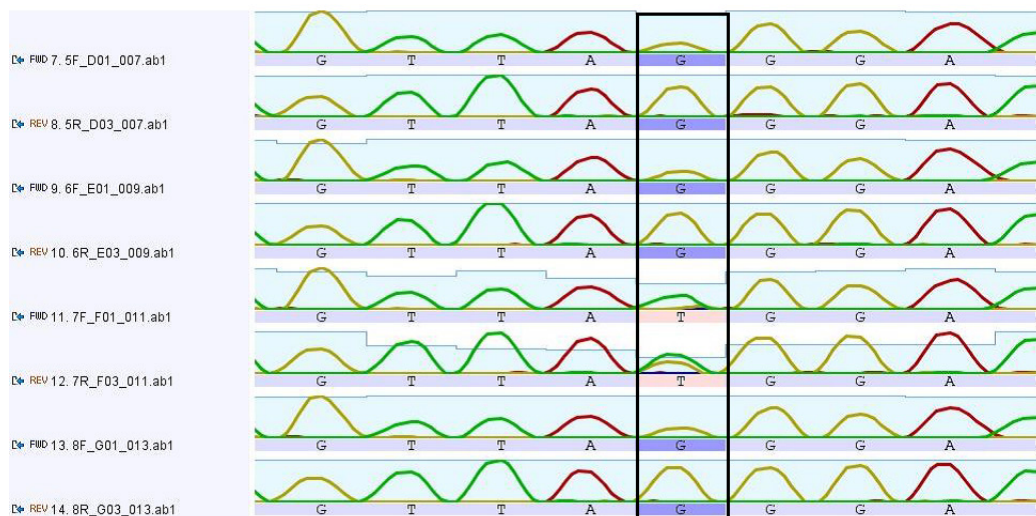


Figure 1. SNP CSN2 image obtained from the Geneious® software.

In analyzing the frequency of heterozygotes among breeds, we observed that Guzerá animals had a higher amount of heterozygous animals, being approximately 6.82% (6 animals) compared to Gir, which had 4.41% (3 animals) of the population (Table 1). None of the evaluated animals presented homozygosity for A1, due to Zebu breeds presenting higher frequency of A2 allele compared to taurine breeds (Zepeda-Batista et al., 2015; Silva, 2016). According to Silva (2016), as Brazil has animals originating from absorbent crossbreeding, pure breed animals generated by these crossings have the mitochondrial DNA of taurine origin, which may have caused the presence of A1 allele at a low frequency in Zebu breeds.

Table 1. Allelic and genotypic frequencies for the A1 and A2 alleles of the CSN2 gene in Gir and Guzerá breeds.

Breed	Allelic frequencies		Genotypic frequencies			HWE ¹	P value ²
	A1	A2	A1A1	A1A2	A2A2		
Gir	0.02	0.98	0	0.04	0.96	0.07	0.99
Guzerá	0.03	0.97	0	0.07	0.93	0.23	0.97

¹HWE = Hardy-Weinberg Equilibrium; ²P value = chi-square test.

Allelic frequencies for A2 allele have higher frequency than the A1 allele in the evaluated breeds. However, animals with the A1A1 genotype were not observed and the A2A2 genotype prevailed over the A1A2 genotype. Vercesi Filho et al. (2012), in a study with Gir animals found a frequency of 0.885 for A2 allele, and 0.084, 0.062, and 0.854 for genotypes A1A1, A1A2, and A2A2, respectively. The frequency of the A1 allele in different breeds varies from 0.06 (Guernsey), 0.01 to 0.12 (Jersey), and 0.31 to 0.66 (Dutch) (Kamiński et al., 2007). Some researchers have shown that the presence of the A2 allele in dairy cattle produces high-quality milk associated with a reduction in cholesterol and triglycerides (Tailford et al., 2003; Kamiński et al., 2007).

The beta casein A1 allele differs from the A2 allele due to a substitution of the amino acid at position 67 from a histidine (CAT) to a proline (CCT), caused by an SNP. The presence of the beta-casein A1 in milk causes the cleavage/breakdown of the peptide bond as a result of the digestion process, releasing bioactive peptides β -casomorphin-7 (BCM-7); the presence of the A2 allele prevents the hydrolysis of the peptide bond between residues 66a and 67a, inhibiting the release of BCM-7 (Kamiński et al., 2007; Sharma et al., 2013). The results observed for the A2 allele frequency compared to the low beta-casein A1 allele frequency indicate that these breeds present low percentage of deleterious variant of beta-casein causing low BCM-7 production.

The results obtained for genotypes classified as A2A2 of Gir and Guzerá breeds showed significant differences ($P < 0.05$) for the percentage of fat, lactose and non-fat dry extract (NFDE) (Table 2). Ribeiro et al. (2009) observed similar results for milk components in studying the milk yield and composition of Gir and Guzerá cows.

Table 2. Means and standard deviations obtained for the characteristics: fat (FAT), protein (PROT), lactose (LACT), total solids (TS), non-fat dry extract (NFDE), and milk yield (YIELD) of Gir and Guzerá cattle.

Characteristic	Gir	Guzerá	P value ²
Fat (%)	4.21 ± 0.64	4.28 ± 0.47	0.78
Protein (%)	3.28 ± 0.04	3.43 ± 0.03	0.02
Lactose (%)	4.61 ± 0.05	4.55 ± 0.04	0.04
Total solids (%)	13.02 ± 0.13	13.21 ± 0.12	0.31
Non-fat dry extract (%)	8.77 ± 0.07	8.97 ± 0.06	0.04
Milk yield (L)	12.63 ± 0.85	11.34 ± 0.75	0.26

²Means followed by different letters in the same column differ by Tukey test at 5%.

Average protein and NFDE levels were higher in Guzerá, while Gir presented higher levels of lactose, despite being the milk component that undergoes the least variation between breeds. The differences found are probably due the particularities among breeds, since both evaluated herds were kept under the same conditions of nutritional management.

Considering that lactose data are presented in relative and not absolute values, we report that despite the occurring differences, the low amplitude variation of lactose is due to the fact that it is related to the osmotic pressure regulation of the mammary gland in a way that a higher production of lactose determines a higher milk yield with the same lactose content (Peres, 2001).

The correlations for NFDE were significant only for protein (0.5329 and 0.5064), lactose (0.7072 and 0.4996), and total solids (0.5905 and 0.5605) for Guzerá and Gir, respectively (Table 3).

Table 3. Pearson correlation between milk constituents of fat (FAT), protein (PROT), lactose (LACT), total solids (TS), non-fat dry extract (NFDE), and milk yield (YIELD) of Guzerá (lower diagonal) and Gir (upper diagonal) of animals classified for the A2A2 allele.

Variables	FAT	PROT	LACT	TS	NFDE	Yield
FAT (%)	-	0.20*	-0.21*	0.81**	0.12	0.08
PROT (%)	0.44**	-	-0.34**	0.40*	0.50*	-0.15
LACT (%)	-0.22*	-0.17	-	0.26	0.49**	0.30*
TS (%)	0.73**	0.70**	0.07	-	0.56**	0.22
NFDE (%)	0.17	0.53**	0.70**	0.59**	-	0.07
YIELD (L)	-0.07	-0.25*	-0.14	-0.18	-0.18	-

*Significant at 5%, and **significant at 1% by t test.

Fat, protein and lactose are major constituents of total solids, and this positive correlation with fat is higher due to a higher variance, followed by protein and less variation in lactose, which is negatively correlated to fat in which it is the main osmotic agent of milk; therefore, the higher the amount of water drawn by lactose, leading to an increase in the milk volume, the lower or more diluted the fat percentage will be (Table 3).

Although there was a high correlation between milk yield and production and some milk constituents, the increase in milk yield is always higher than the increase in the production of constituents. This causes the correlation between milk yield (kg) and lactose to be negative. Therefore, a selection focusing only on milk yield can result in losses to the milk composition (Oltenucu and Broom, 2010). A positive correlation was observed due to the amount of fat being related to the amount of protein in milk, indicating that the increase in fat content will influence protein levels, although with a higher influence on Guzerá compared to Gir.

Guzerá and Gir breeds grouped by principal components (PCs) resulted in two groups, although most of the animals presented similarity in produced characteristics (Figure 2). The PCs used for the graphic dispersion indicates animals showing higher differences between and within the evaluated breeds, indicating that there are major differences between these animals in relation to the studied population.

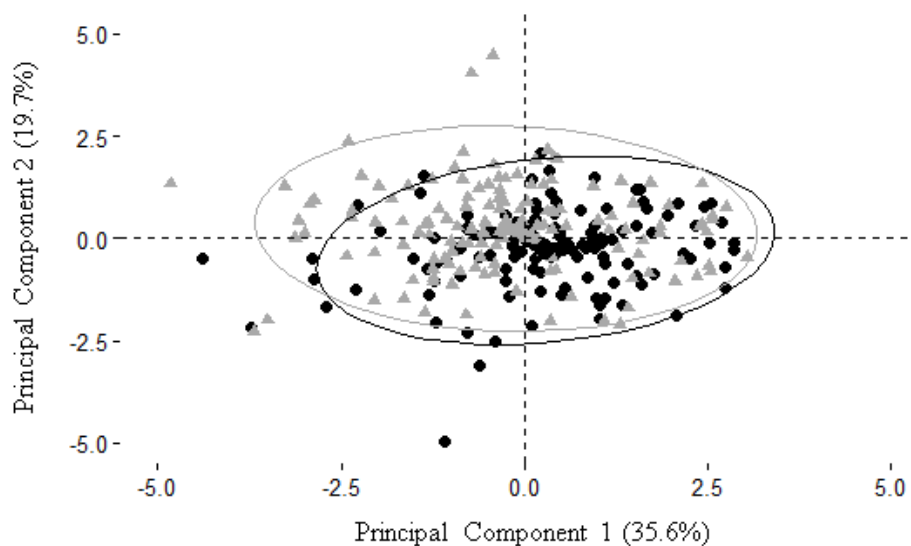


Figure 2. Grouping of Guzerá (triangles) and Gir (circles) animals in relation to milk constituents and yield.

Both PCs explained about 55% of the explained variances by components. The cut angle for Gir (black circle) and Guzerá breeds (gray circle) are quite different. A concentration of Gir animals with positive values for PC1 and negative for PC2 can be observed; the highest concentration of information in Guzerá animals is positive for PC2 and negative for PC1. Most of the analyzed information can be found in the overlapping area between the two circles, with those that are in the circle regarding Gir, and those that are practically in the circle regarding the Guzerá breed, thus corroborating the fact that these animals have similar genes and origin for the analyzed characteristic. However, the information generated in the analysis shows that there is a greater dispersion of data for the Guzerá breed.

CONCLUSION

The allelic frequency of the A2 allele and the genotypic frequency of A2A2 genotypes for the β -casein gene in the assessed Zebu breeds indicate that these breeds may produce less allergenic milk for individuals who are sensitive to the β -casein protein. Guzerá animals showed a higher percentage of protein, lactose, and non-fat dry extract, which provided a separation of some animals into two groups by the main component analysis.

REFERENCES

- Barroso A, Dunner S and Cañón J (1999). Technical note: use of PCR-single-strand conformation polymorphism analysis for detection of bovine beta-casein variants A1, A2, A3, and B. *J. Anim. Sci.* 77: 2629-2632. <https://doi.org/10.2527/1999.77102629x>
- Bonfatti V, Di Martino G, Cecchinato A, Vicario D, et al. (2010). Effects of β - κ -casein (CSN2-CSN3) haplotypes and β -lactoglobulin (BLG) genotypes on milk production traits and detailed protein composition of individual milk of Simmental cows. *J. Dairy Sci.* 93: 3797-3808. <https://doi.org/10.3168/jds.2009-2778>
- De Noni I, FitzGerald RJ, Korhonen HJ, Le Roux Y, et al. (2009). Review of the potential health impact of β -casomorphins and related peptides. *EFSA Sci. Rep.* 231: 1-107.
- Drummond AJ, Ashton B, Buxton S, Cheung M, et al. (2011). Research software for biologists, not computer scientists. Available at [geneious.com]. Accessed August 7, 2016.
- Farrell Jr HM, Jimenez-Flores R, Bleck GT, Brown EM, et al. (2004). Nomenclature of the proteins of cows' milk--sixth revision. *J. Dairy Sci.* 87: 1641-1674. [https://doi.org/10.3168/jds.S0022-0302\(04\)73319-6](https://doi.org/10.3168/jds.S0022-0302(04)73319-6)
- Kamiński S, Cieslińska A and Kostyra E (2007). Polymorphism of bovine beta-casein and its potential effect on human health. *J. Appl. Genet.* 48: 189-198. <https://doi.org/10.1007/BF03195213>
- Olenski K, Kamiński S, Szyda J and Cieslińska A (2010). Polymorphism of the beta-casein gene and its associations with breeding value for production traits of Holstein-Friesian bulls. *Livest. Sci.* 131: 137-140. <https://doi.org/10.1016/j.livsci.2010.02.023>
- Oltenu PA and Broom DM (2010). The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim. Welf.* 19: 39-49.
- Peres JR (2001). O leite como ferramenta do monitoramento nutricional. Uso do leite para monitorar a nutrição e o metabolismo de vacas leiteiras. UFRGS, Porto Alegre, 30-45.
- Otaviano AR, Lima ALF, Laureano MMM, Sena JAD, et al. (2008). β -casein gene polymorphism permits identification of bovine milk mixed with bubaline milk in mozzarella cheese. *Genet. Mol. Biol.* 31: 902-905. <https://doi.org/10.1590/S1415-47572008005000002>
- Regitano LDA, Regitano LCA and Coutinho LL (2001). Extração de DNA para aplicação em reação em cadeia da polimerase (PCR). Biologia molecular aplicada à produção animal. Embrapa Informação Tecnológica, Brasília, 179-186.
- Ribeiro AB, da Fonseca Tinoco AF, da Costa Lima GF, Guilhermino MM, et al. (2009). Produção e composição do leite de vacas Gir e Guzerá nas diferentes ordens de parto. *Rev. Caatinga* 22: 46-51.
- Salman AKD and Laureano MMM (2006). Protocolos para extração de DNA genômico de amostras de pelo de bovinos. *Circular Técnica* •••: 87.

- Sharma V, Sharma N, Jawed B and Nautiyal SC (2013). High resolution melt curve analysis for the detection of A1, A2 β -casein variants in Indian cows. *J. Microbiol. Biotechnol. Res.* 3: 144-148.
- Silva MR (2016). Influência do DNA mitocondrial na bovinocultura. *Rev. Eletrônica Interdiscip.* 1: 49-53.
- Tailford KA, Berry CL, Thomas AC and Campbell JH (2003). A casein variant in cow's milk is atherogenic. *Atherosclerosis* 170: 13-19. [https://doi.org/10.1016/S0021-9150\(03\)00131-X](https://doi.org/10.1016/S0021-9150(03)00131-X)
- Vercesi Filho AE, Camargo GMF, Cardoso DF, Zadra EF, et al. (2012). Identificação de alelos A1 e A2 para o gene da beta-caseínas na raça Gir Leiteiro. In: IX Simpósio Brasileiro de Melhoramento Animal. *Anais...* João Pessoa, Paraíba.
- Woodford KB (2008). A1 beta-casein, type 1 diabetes and links to other modern illnesses. An invited plenary paper to the International Federation Western Pacific Congress.
- Zepeda-Batista JL, Alarcón-Zúñiga B, Ruíz-Flores A, Núñez-Domínguez R, et al. (2015). Polymorphism of three milk protein genes in Mexican Jersey cattle. *Electron. J. Biotechnol.* 18: 1-4. <https://doi.org/10.1016/j.ejbt.2014.10.002>