Genetic association of variations in the kappa-casein and β-lactoglobulin genes with milk traits in girolando cattle

Barbosa, Severino Benone Paes 1*, Araújo, Ítala Iara Medeiros de 1, Martins, Marta Fonseca 2, Silva, Elizabete Cristina da 1, Jacopini, Laís Aberrachid 1, Batista, Ângela Maria Vieira; Silva, Marcus Vinícius Barbosa da 2

Associação genética de variações nos genes da kappa-caseína e β-lactoglobulina com características do leite de bovinos da raça Girolando

1Departamento de Zootecnia da Universidade Federal Rural de Pernambuco, Recife-PE, Brasil
2Embrapa Gado de Leite, Rua Eugênio do Nascimento, Juiz de Fora - MG, Brasil

*Correspondence author email: severino.pbarbosa@ufrpe.br

SUMMARY

In dairy farm animals, one of the most important goal of the selection is the improvement of milk yield and composition. Several studies have demonstrated that the candidate genes of the kappa-casein (CSN3) and β-lactoglobulin (β-LG) are associated with milk yield, milk quality and health traits in dairy animals. Therefore the aim of this study was to detect polymorphisms in CSN3 and β-LG genes and its association with milk yield in up to 305 days (305MY) and predicted transmission capacity (PTA) for 305MY in Girolando cattle. Totally, 138 bulls and 729 cows (n=867) were sampled. The genotypes of both genes were obtained by the PCR-RFLP method using HinfI and HaeIII enzymes for CSN3 and β-LG genes, respectively. Statistical results revealed two alleles A and B for both genes. The genotypes and alleles more frequent for CSN3 and β-LG genes were respectively: AA (0.7324) and A (0.8558), and AB (0.4827) and A (0.5017). The x² test revealed that the two loci were at Hardy–Weinberg equilibrium (p<0.001). The allele substitution effects for the variants were not significant on 305MY and PTA for 305MY (p>0.05). The allele variants of β-LG and CSN3 might be more investigated before include them into future breeding schemes designed for Girolando dairy cattle with objective of improving milk traits as milk yield in up to 305 days (305MY) and predicted transmission capacity (PTA) for 305MY.

Key words: SNP genotyping, quantitative traits, dairy cattle, milk protein, dairy industry

RESUMO

Em rebanhos leiteiros, um dos objetivos mais importantes da seleção é a melhoria da produção e composição do leite. Vários estudos demonstraram que os genes candidatos da kappa-caseína (CSN3) e da β-lactoglobulina (β-LG) estão associados à produção de leite, qualidade do leite e características de saúde em animais leiteiros. Portanto, o objetivo deste estudo foi detectar polimorfismos nos genes CSN3 e β-LG e avaliar possíveis associações desses polimorfismos com a produção de leite em até 305 dias (305MY) e a capacidade de transmissão prevista (PTA) de leite em bovinos da raça Girolando. No total, 138 touros e 729 vacas (n = 867)
foram amostrados. A genotipagem foi realizada pelo método PCR-RFLP utilizando as enzimas HinfI e HaeIII para os genes CSN3 e β-LG, respectivamente. Os resultados estatísticos revelaram dois alelos A e B para ambos os genes. Os genótipos e alelos mais frequentes para os genes CSN3 e β-LG foram respectivamente: AA (0,7324) e A (0,8558) e AB (0,4827) e A (0,5017). O teste $\chi^2$ revelou que os dois loci estavam em equilíbrio de Hardy-Weinberg ($p < 0,001$). Os efeitos de substituição alélica para as variantes não foram significativos para as características 305 MY e PTA para 305MY ($p > 0,05$). Portanto, as variantes alélicas identificadas nos genes β-LG e CSN3 devem ser mais investigadas antes de serem incluídas nos programas de melhoramento desenhados para bovinos leiteiros Girolando objetivando melhorar as características do leite analisadas no presente estudo.

**Palavras-chave:** genotipagem de SNP, características quantitativas, bovinos leiteiros, proteína do leite, indústria leiteira
INTRODUCTION

The Girolando breed is one of the most important breeds of dairy cattle in Brazil, resulting from the crossing between Gyr (a Bos indicus breed) and Holstein (a Bos taurus breed) cattle breeds, since 1/2 Holstein + 1/2 Gyr until 13/16 Holstein + 3/16 Gyr, and has shown better adaptation to hot temperatures and tropical diseases (Facó et al., 2005; Bicalho et al., 2006). The purpose of the formation of the Girolando breed is to obtain animals of 5/8 composition Holstein + 3/8 Gyr with well-defined racial pattern and good dairy productivity (SILVA et al., 2016). This breed was selected for several years for increasing milk produced in the country.

Polymorphisms in candidate genes related to economically important traits in dairy cattle such as the candidate genes of kappa-casein (CSN3) and beta-lactoglobulin (β-LG) have been associated with milk composition, cheese production, milk quality and milk production (Botaro et al., 2009; Stipp et al., 2013; Singh et al., 2014; Selvaggi et al., 2014). Beta-lactoglobulina protein represents about 50% of total whey protein of ruminant milk (Fox & McSweenwey, 1998; Selvaggi et al. 2014). Milk β-LG gene was mapped to chromosome 11 in bovine and molecular studies already determined 15 alleles in this gene (Matejicek et al., 2008), out of which A and B variants are the most frequent and investigated (Zaglool et al., 2016). In bovines, the variant B has been more related to milk quality while variant A has been more associated with milk yield (Ng-Kwai-Hang et al. 2002; Tsiaras et al., 2005).

Caseins are milk proteins secreted by mammary gland cells, constitute about 78-82% of bovine milk proteins and are divided into four different types of caseins: alpha s1, alpha s2, beta, and kappa, where approximately 12% contribution comes from kappa-casein. Eleven variants have been reported for CSN3 (Prinzenberg et al., 2008; Caroli et al., 2009, Botaro et al., 2009) associated with composition, processing, milk quality and milk yield, being A and B the most frequent (Sitkowska et al., 2009; Dogru & Ozdemir, 2009; Ju et al., 2009). The A allele is favorable to milk yield, but with lower protein content, while B allele is related to higher fat and protein contents, but with lower milk yield (Botaro et al., 2009; Ju et al., 2009; Hamza et al., 2010).

Within this context, this present study aimed to identify A and B allele variants of CSN3 and β-LG genes and to investigate whether its association with milk yield in up to 305 days (305MY) and predicted transmission capacity (PTA) for 305MY to develop tools for the selection of animals participating in a test of progeny of the Girolando breed in Brazil.

MATERIAL AND METHODS

To detect polymorphisms in the CSN3 and β-LG bovine genes, samples of semen and or blood were collected from 867 unrelated animals (138 bulls and 729 cows) from different contemporary groups that participate of the Girolando Progeny Test coordinated by the Breeders Association of the Girolando Breed in partnership with Embrapa Dairy Cattle-National Dairy Cattle Research Center. The biological material used in this research represents all samples available in the DNA Bank at Embrapa until March 2012. These samples were collected during the milk control or linear classification of the
animals. All animals used in this study came from different private farms that participate of the Progeny Test and were a representative population of a population studied by SILVA et al. (2012).

The genomic DNA samples were extracted from blood and semen cells using a Dneasy Blood & Tissue Kit (Qiagen, Hilden, Germany), according to the manufacturer's recommendations. Subsequently, the quality of DNA genomic DNA including concentration and purity was verified using a Spectrophotometry (Nanodrop 1000, Thermo Fisher Scientific Inc., Wilmington, USA). After that, the extracted DNA samples were submitted to the PCR amplification technique (Polymerase Chain reaction), using the oligonucleotides initiators (primers) described by Barroso et al. (1998) and Medrano and Aguilar-Cordova (1990) for the CSN3 and β-LG genes, respectively (Table 1).

Table 1. Sequences of the forward (F) and reverse (R) primers for the amplification of the kappa-casein (CSN3) and β-Lactoglobulin (β-LG) genes.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Primers sequences (5’→3’)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSN3</td>
<td>F-TGTGCTGAGTATCCTAGTTATGG R-GCGTTGTCTTCTCTTTGATGTCCTCCTTAG</td>
<td>Barroso et. (1998)</td>
</tr>
<tr>
<td>β-LG</td>
<td>F-GTCCTTCTGACACCCGACTACA R-CAGGACACCCGGCTCCCGGTATATGA</td>
<td>Medrano &amp; Aguilar-Cordova (1990)</td>
</tr>
</tbody>
</table>

PCR amplification was prepared in a 25 µl volume containing of 400 nM and 125 nM of each primer, for the CSN3 and β-LG genes respectively; 50 ng of genomic DNA and 1x from GoTaq Green Master Mix (Promega, Madison, Wisconsin, USA). The gene amplification programs consisted: 94 ºc for 5 min, 35 cycles (94 ºc for 1min; 65 ºc for 1 min (60 ºc for β-LG); 72 ºc for 1 min) followed by a final extension for 5 min at 72 ºc (7 min for β-LG). For amplification, the cycler GeneAmp PCR System 9700 (Applied Biosytem, Forster City, CA, USA) was used. The restriction digestion of the PCR products was carried out with the HinfI and HaeIII enzymes (New England Biolabs, Inc., Ipswich, USA) according to manufacturers' recommendations. The digestion products were analyzed on 2% agarose gel.

Allelic and genotypic frequencies were calculated to determine the population structure by using Popgen version 1.32 package (Yeh et al., 1997), and the significance of differences between observed and expected genotype frequencies was tested based on Chi-square ($\chi^2$) test at one degree of freedom at one percent level to test whether the distribution of the genotypic frequencies was in the Hardy-Weinberg equilibrium.

For the association study, milk production data in up to 305 days (305MY) of 536 cows were used. For the Predicted Transmitting Ability (PTA) for 305MY data, information from 127 bulls and 536 cows were included. The group of evaluated sires presented genetic composition of 3/4 and 5/8 of Holstein breed and the group of cows presented the genetic composition equal to or greater than 5/8 of the Holstein.

Received on: 22/12/2018. Accepted on: 09/01/2019
Estimates of PTA for 305MY daughter j of sire i, μ is the overall population mean, Si is the fixed effect of sire i, β is the regression coefficient representing one-half the allele substitution effect (α/2), xij is the number of B alleles (0, 1 or 2) for the j daughter of sire i, and εij is the residual.

Estimates of PTA for 305MY were weighted by accuracy values to obtain weighted least squares estimates for the allele substitution effects. Data of P305 were analyzed following fixed effects model:

\[ Y_{ijklm} = \mu + S_j + G_{Ck} + C_{Gl} + O_m + \varepsilon_{ijklm}, \]

where, Yijklm represent the milk production of cow i, daughter of sire j, μ is a general constant, Sj is the fixed effect of jth sire, GCk is the fixed effect of kth contemporary group ( k=1, 2, ..., 52) (herds, calving birth and calving season); CGl is the fixed effect of lth genetic composition (l = 1, 2, 3, 4, 5, 6, these are representing the genetic composition of Girolando animals), Om is the effect of mth genotype ( m=AA, AB, BB) and εijklm is the residual.

The association study was carried out through regression analysis, using the GLM procedure of SAS 9.1 (SAS Institute, Inc., Cary).

RESULTS AND DISCUSSION

The use of the molecular PCR-RFLP technic chose in this study was effective to detect kappa-casein (CSN3) and beta-lactoglobulin (β-LG) polymorphisms. The CSN3 gene amplification product produced a DNA fragment of 453 bp that was digested with HinfI restriction enzyme, and for types of restriction patterns were obtained 426, 326, 100 and 27 bp as shown in Figure 1. The AA genotype was identified by the presence of three fragments (-326, 100 and 27 bp), the AB is characterized by the presence of the four fragments (-426, 326, 100 and 27 bp) and the BB has only two fragments (-426 and 27 bp).

For the β-LG gene, a 262 bp fragment was amplified with two polymorphic alleles represented by four 153, 109, 79 and 74 bp fragments from the enzymatic digestion of the PCR product with HaeIII (Figure 2). The AA genotype identified by the presence of two fragments (-153 and 109 bp), AB is characterized by the presence of the four fragments (-153, 109, 79 and 74) and the BB has three fragments (-109, 79 and 74 bp).
Knowledge about the proportions of the different genotypes and the different alleles of the candidate genes in the population through the study of the genotype and allelic frequencies and the association of these with the productive records of the animals may allow the development of strategies for marker assisted selection (MAS), making it possible to increase milk production and improve physical-chemical and technological characteristics, as well as increasing the speed and dynamism in the decision making regarding the selection or culling of animals for reproduction. According to Drogemuller et al. (2001) the use of MAS together with traditional selection tools can be more effective when the aim of breeding program is improving complex traits because with the MAS is possible getting faster genetic progress.

The genotypic and allelic frequencies for the CSN3 and β-LG genes are summarized in Table 2. The frequencies of alleles A and B at locus CSN3- Hinf I, were 0.86 and 0.14, with resulting in AA genotype being the most frequent (0.73). For β-LG gene, allelic frequency of A and B variants has been more close A = 0.52 and B = 0.48. The AB genotype from β-LG occurred at higher frequency demonstrating there was more variability in this gene in Girolando cattle analyzed in this study. The probability of deviations from the Hardy-Weinberg expectations for both genes were based on Chi-square test ($\chi^2$) and showed that all genotypic frequencies in the population were in Hardy-Weinberg equilibrium (P < 0.01) (Table 2).

Based on the results of this study, obtained for the CSN3 gene, was
observed a tendency for the fixation of higher AA genotype frequency in this population, corroborating with other studies showing this similar results for the zebu breeds, such as Gir (Silva & Del Lama, 1997; Kemenes et al., 1999) and taurine, such as the Holstein (Ng-Kwai-Hang et al., 1984; Lin et al., 1986; Hallen et al. 2011; Duifhuis-Rivera et al., 2014). The low frequency of the B allele, which was reported in previous studies to be associated with higher yield for cheese production, was expected because the frequency of this allele in both founding breeds (Holstein and Gir) of Girolando is also low (Famula & Medrano, 1994; Tsiaras et al., 2005; Azevedo et al., 2008).

The allele frequencies found for the β-LG gene differ from those obtained by Ng-Kwai-Hang et al. (1984), who studied the genotypic and allelic frequencies of 3,870 Holstein cows reared in 63 commercial herd in Quebec and they did not find the alleles equally distributed in the studied population with allelic frequencies of 0.39 and 0.61 for A and B, respectively, and frequencies genotypic of 0.1334 (AA), 0.5054 (AB) and 0.3602 (BB). The population studied by these authors was also in EHW and the proportion of heterozygous individuals was higher than that of homozygotes, as verified in the present study. On the other hand, Botaro (2007) obtained, using 74 Girolando animals in his study, greater frequency of genotype BB (0.45) than genotypes AB (0.34) and AA (0.21), finding no balance in the population.

Table 2. Allelic and genotypic frequencies for CSN3 and β-LG genes in Girolando cattle.

<table>
<thead>
<tr>
<th>Loci</th>
<th>Genotypes</th>
<th>Nº of genotype</th>
<th>Frequency</th>
<th>EHW (x²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Observed</td>
<td>Expected</td>
<td>Genotypic</td>
</tr>
<tr>
<td>CSN3- Hinf I</td>
<td></td>
<td>AA 635</td>
<td>635.02</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB 214</td>
<td>213.96</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BB 18</td>
<td>18.02</td>
<td>0.021</td>
</tr>
<tr>
<td>β-LG- HaeIII</td>
<td></td>
<td>AA 226</td>
<td>218.50</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB 419</td>
<td>433.99</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BB 223</td>
<td>215.50</td>
<td>0.26</td>
</tr>
</tbody>
</table>

**P<0.01

In the present study was no significant association (p>0.05) between the different genotypes and milk yield in up to 305 days (305MY) and predicted transmission capacity (PTA) for 305MY of Girolando cows and bulls analyzed. The average of the analyzed traits for the observed allelic substitution (α/2) for CSN3 and β-LG genes are shown in Table 3. The non-significance of the association of the variables with allele variants also resembles those obtained by previous studies with Holstein cows (Aleandri et al., 1990; Ng-Kwai-Hang et al., 1990; Tsiaras et al., 2005). The current results...
regarding the relationship between milk production and B allele are in agreement with those obtained by Bovenhuis et al. (1992), who genotyped 6,803 Holstein cows for the CSN3 gene and determined that the BB genotype produced lower amounts of milk compared to AA cows, although for these authors the results were significant. The results of Rachagani & Gupta (2008) indicated that genotype BB produced more milk than those of genotypes AA and AB, also a significant study. The conflicting results found in the literature for CSN3 showed that the A allele (Gonyon et al., 1987; Bovenhuis et al., 1992), or the B allele (Lin et al., 1986, 1989, Eenenanm & Medrano, 1991) are being related to higher milk production. According to Ng-Kwai-Hang et al. (1990), this fact can be attributed to the different number of samples analyzed, different genetic material and mainly to the rigor of the statistical analyzes. Considering that all variants of the CSN3 gene are located on region of chromosome 6, between 200 and 300 kb, potentially the proximity of the genes, for the effect of binding disequilibrium is great. However, this effect within a family can be altered due to differences between casein loci and other protein genes, which could explain the inconsistencies observed in previous studies (Cowan et al., 1992; Braunschweig et al., 2000). In addition, kappa-casein and beta-casein have been showed a strong link between the effects of gene binding (Bovenhuis et al., 1992).

Table 3. Estimates of allelic substitution (α/2) and p-value associated with CSN3 and β-LG in cows and bulls Girolando population.

<table>
<thead>
<tr>
<th>Variables</th>
<th>CSN3</th>
<th>p-value</th>
<th>β-LG</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>305MY (kg)</td>
<td>- 62.93</td>
<td>0.6924</td>
<td>- 34.98</td>
<td>0.7846</td>
</tr>
<tr>
<td>PTA for 305MY (kg)</td>
<td>- 14.16</td>
<td>0.6179</td>
<td>- 18.39</td>
<td>0.4130</td>
</tr>
</tbody>
</table>

The results found in this study resemble those obtained by Ng-Kwai-Hang et al. (1990), with animals of the Holstein breed that also did not find significant effect for the association of the β-LG gene with milk production. However, Aleandri et al. (1990) observed that AA genotype determined higher milk yield in first-lactating Holstein cows, as well as Bovenhuis et al. (1992) who observed that AA Hollander cows produced 93 kg more milk than cows of genotype BB. This result is not in agreement with Jairam & Nair (1983) that demonstrated higher milk yield for cows with BB genotype. In addition the A allele for CSN3 was the major allele and it may suggest that this enzymatic site was less informative to detect variability in Girolando cattle. Therefore, the allele variants of β-LG and CSN3 might be more investigated before include them into future breeding schemes designed for Girolando dairy cattle with objective of improving milk traits as milk yield in up to 305 days (MY305) and predicted transmission capacity (PTA) for 305MY.
REFERENCES


MEDRANO, J.F.; AGUILAR-CORDOVA, E. Polymerase chain reaction amplification of bovine β-lactoglobulin genomic sequences and


