Genetics and Molecular Research

Polymorphism in the *A2M* gene associated with high-quality milk in Murrah buffaloes (*Bubalus bubalis*)

A.C. Freitas[1], G.M.F. de Camargo[1,4], R.R. Aspilcueta-Borquis[2], N.B. Stafuzza[3], G.C. Venturini[1,5], F. Tanamati[1], N.A. Hurtado-Lugo[1], C.C. Barros[1] and H. Tonhati[1]

[1]Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, SP, Brasil
[2]Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados, Dourados, MS, Brasil
[3]Departamento de Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, SP, Brasil
[4]Universidade Tecnológica Federal do Paraná, Dois Vizinhos, PR, Brasil
[5]Departamento de Melhoramento e Nutrição Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista "Júlio de Mesquita Filho", Botucatu, SP, Brasil

Corresponding author: A.C. Freitas
E-mail: anafreitas.unesp07@gmail.com

Genet. Mol. Res. 15 (2): gmr.15027637
Received September 15, 2015
Accepted November 19, 2015
Published April 4, 2016
DOI http://dx.doi.org/10.4238/gmr.15027637

ABSTRACT. The study of genes associated with host defense mechanisms, such as the *A2M* gene, plays a critical role in preventing diseases that reduce milk yield and its constituents. The aim of this study was to identify polymorphisms in the *A2M* gene in Murrah buffaloes (*Bubalus bubalis*), and investigate their associations with milk yield, fat and protein production, fat and protein percentages, and somatic cell count. Hair follicle samples of 136 animals were collected for DNA extraction, and polymorphisms were identified by polymerase chain reactions and sequencing. Statistical analyses were performed to ascertain the allelic and

genotypic frequencies, the Hardy-Weinberg equilibrium, and association analysis was conducted between the polymorphisms and the traits studied. Comparative analysis between buffalo and bovine sequences revealed seven nucleotide substitutions. Alignments among the buffalo sequences identified three single nucleotide polymorphisms (SNPs), including one in exon 29, g.241A>G, which was used in subsequent statistical analyses. A Hardy-Weinberg test indicated that this SNP was in equilibrium in this population, and was significantly associated ($P < 0.05$) with fat production and fat and protein percentages. Therefore, this SNP can be used as a molecular marker for these traits.

Key words: Allele; *Bubalus bubalis*; Protein production; Somatic cell count

INTRODUCTION

Brazilian buffalo livestock is mainly produced for milk yield, which has a high financial return for farmers because milk has a high percentage of total solids that provide a high yield of derivatives (Tonhati et al., 2008). Dairy animals with large production capacities and superior constituent qualities have been selected for breeding programs that use molecular genetics to identify the genes responsible for phenotypic variation. Several candidate genes associated with buffalo milk components are described in the literature, such as casein alpha s1, s2 alpha, beta and kappa casein (Otaviano et al., 2005; Riaz et al., 2008), leptin and melatonin (Zetouni et al., 2013, 2014), ghrelin (Gil et al., 2013), and osteopontin (Rolim Filho et al., 2013).

The study of genes associated with host defense mechanisms plays a critical role in preventing diseases that reduce milk yield and its constituents, such as *DRB3* (Kumar et al., 2011), tumor necrosis factor alpha (*TNF-α*), and interleukin-1b (*IL-1b*) (Lahouassa et al., 2007). Another gene, *A2M* (alpha-2-macroglobulin), has been investigated in several animals, including the rat (Enghild et al., 1989), carp (Onara et al., 2008), and shrimp (Ma et al., 2010). The A2M protein is found in blood plasma globulin fraction, and acts as an inhibitor of non-specific proteases (Enghild et al., 1989). In addition, it plays important roles in host immunity, particularly in animals with an inflammatory process such as mastitis, as described by Wang et al. (2012). These authors identified one polymorphism in the *A2M* gene that is associated with low somatic cell counts (SCCs) and resistance to mastitis in dairy cattle (Wang et al., 2012).

Polymorphisms in *A2M* have not yet been identified in the buffalo. Understanding the associations between *A2M* and production traits would provide relevant information on allelic and genotypic variability, and supply important molecular markers for the genetic improvement of buffalo herds. Therefore, the aims of the current study were to identify polymorphisms related to *A2M* in Murrah buffaloes (*Bubalus bubalis*) and investigate their associations with milk yield, fat and protein production and percentages, and SCCs.

MATERIAL AND METHODS

Farm and animals

The Murrah buffaloes used in this study belonged to the Tapuio Farm, Taipu, Rio Grande

do Norte, Brazil. The grazing system used was the Voisin Rational Grazing System. The herd has approximately 1000 head of cattle, 468 of which are dams, with a production of around 3500 L/day. The farm implements the dairy buffalo milk-recording program maintained by Department of Animal Science, Universidade Estadual Paulista, Brazil. Milk recording and sampling were performed periodically to determine fat and protein content and SCC, as was the collection of biological material for DNA extraction.

Biological material collection and polymerase chain reactions (PCRs)

Hair follicle samples from 136 Murrah buffaloes were collected and stored at 4°C until DNA extraction. DNA was extracted by the phenol-chloroform-isoamyl alcohol method described by Sambrook et al. (1989).

The primers used to amplify exon 29 of *A2M* were designed based on the *A2M* sequence of the *Bos taurus UMD_3.1* reference assembly (GenBank accession No. AC_000162.1, 101,274,869-101,346,613). The primers were designed using the Primer3Plus program (Untergasser et al., 2007), and their quality was verified using the OligoAnalyzer 3.1 tool (http://www.idtdna.com/calc/analyzer).

The genomic DNA samples were amplified by PCR in a final volume reaction of 15 µL, which contained 70 ng genomic DNA, 15 pM of each primer, and 7.5 µL GoTaq® Colorless Master Mix (Promega, Madison, WI, USA). The appropriate annealing temperature was selected for a PCR performed in a T100™ ThermalCycler (Bio-Rad, Hercules, CA, USA) with thermal gradient software. The PCR conditions were as follows: initial denaturation at 95°C for 5 min, followed by 30 cycles at 95°C for 1 min (denaturation), 61.8°C for 1 min (annealing), extension at 72°C for 1 min, and a final extension at 72°C for 5 min.

After amplification, 2 µL PCR products were electrophoresed on 1.5% agarose gels stained with GelRed™ (Biotium, Hayward, CA, USA) in 1X TBE buffer at 90 V for 50 min. The PCR products were visualized in ultraviolet light, photographed using Gel-Doc™ apparatus (BioRad), and analyzed with the Image Analysis software (Kodak®) in order to confirm that the desired fragment was successfully amplified.

Sequencing

The PCR products were purified using Wizard® SV Gel and PCR Clean-Up System Kit (Promega) following the manufacturer recommendations. The samples were sequenced with both forward and reverse primers using an ABI PRISM® BigDye® Terminator v3.0 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) in an ABI 3730xl Automatic Sequencer (Applied Biosystems). For polymorphism identification, DNA sequences obtained from *A2M* were analyzed by the Codon CodeAligner program (http://www.codoncode.com/aligner/download.htm).

Allelic and genotypic frequencies

Allelic frequencies of the genotyped single nucleotide polymorphism (SNP) were calculated by simple allele counting (Falconer and Mackay, 1996) using PROC FREQ in SAS (SAS 9.2, SAS Institute, Cary, NC, USA). Possible deviations from the genotypic and allelic frequencies were tested using the chi-square test, in order to verify that the genotyped population was in Hardy-Weinberg equilibrium.

Association analysis

Relationships between the SNP and milk, fat, and protein production, fat and protein percentages, and SCC were investigated using the following model:

$$Y_{ijkl} = \mu + CG_i + S_j + M_k + b_l (I_{ijkl} - \bar{I}) + e_{ijkl} \quad \text{(Equation 1)}$$

where is the milk yield, fat and protein percentages of milk at 305 days, and SCC; is the constant inherent to all observations; is the fixed effect of contemporary group; is the random effect of sire, with mean 0 and variance ; is the fixed effect of genotype for SNP; is the regression coefficient of trait in relation to the age of the cow at measurement; is the quadratic regression coefficient of trait in relation to the age of the cow at measurement; is the age of the cow at measurement; is the average age of the cows at measurement; and is the residual random effect associated with trait with mean 0 and variance .

The contemporary group was a concatenation of year and season of calving. Analysis of variance was performed using the PROC MIXED procedure in SAS. The SCC was transformed to a logarithmic scale (SCCt) using the function described by Dabdoub and Shook (1984):

$$SCC_t = [\log_2 (SCC / 100{,}000)] + 3 \quad \text{(Equation 2)}$$

The means for milk yield, fat and protein production, fat and protein percentages, and SCC for the different genotypes were compared by Tukey's test ($P < 0.05$). After performing association analysis between the polymorphisms and production traits and milk quality, the significant SNPs had their additive effects and dominance deviations tested. For this analysis, the PROC GLM procedure by orthogonal contrasts was performed in SAS.

RESULTS

PCR and sequencing

The size of the sequenced PCR product was 303 bp, and contained part of intron 28, exon 29, and part of intron 29. Comparative analysis between the buffalo sequence and its bovine counterpart identified seven nucleotide substitutions; two synonymous in intron 28 and five in exon 29 (two synonymous and three nonsynonymous), as described in Table 1. These substitutions were identified in all of the buffalo DNA sequences.

Table 1. Positions of nucleotide substitutions identified in the *A2M* gene sequence from buffalo when compared to its homologous cattle sequence.

Position in the amplified fragment (303 bp)	Gene region	Nucleotide substitution	
		Cattle	Buffalo
53	Intron 28	A	T
72	Intron 28	T	C
89	Exon 29	C	T
122	Exon 29	C	T
149*	Exon 29	A (Glutamine)	C (Histidine)
232*	Exon 29	C (Threonine)	G (Arginine)
240*	Exon 29	T (Tyrosine)	C (Histidine/Arginine)

*Nucleotide substitution results in a change in the translated amino acid.

Three SNPs were identified in the buffalo *A2M* sequence (Table 2), one in intron 28 (g.38C>T), one in exon 29 (g.241A>G), and one in intron 29 (g.276G>A). SNP g.241A>G results in the change observed in the codon, where the first nucleotide is cytosine (C) and the second nucleotide could be adenine (A) or guanine (G), modifying the translated amino acid from histidine (CAT) to arginine (CGT). The obtained sequence, with substitutions and polymorphisms, was deposited on GenBank database under accession No. KF928932.

Table 2. Position of SNPs identified in the buffalo *A2M* gene sequence.

Position in the amplified fragment (303 bp)	Gene region	SNP
38	Intron 28	C/T
241*	Exon 29	A (Histidine)/G (Arginine)
276	Intron 29	G/A

*SNP results in a change in the translated amino acid.

Allelic and genotypic frequencies

We only conducted statistical analysis on the SNP identified in exon 29 (g.241A>G). The frequencies of the A and G alleles were 0.75 and 0.25, respectively, and the observed genotype frequencies were AA = 0.56, AG = 12.39, and GG = 0.05. The values for the expected genotype frequencies were AA = 0.57, AG = 12.37, and GG = 0.06. The frequencies observed were in equilibrium with the expected frequencies, according to the Hardy-Weinberg test, which was conducted using a chi-square test ($P < 0.05$), which meant that there was no indicative selection for this SNP in this population.

Association analysis

We identified associations between SNP g.241A>G and fat production ($P < 0.01$), fat percentage ($P < 0.01$), and protein percentage ($P < 0.05$) (Table 3). The GG genotype had the highest means for the fat production and percentage traits, statistically significantly differing from the other two genotypes in both cases. The three genotypes had similar phenotypic mean values for protein percentage, with homozygous genotypes differing from each other.

Table 3. Phenotypic means of the allelic substitution in SNP g.241A>C for milk yield (MY), fat production (FP), protein production (PP), fat percentage (%F), protein percentage (%P), and somatic cell count (SCC).

Genotype	MY (kg) (P = 0.115)	FP (kg)** (P = 0.003)	PP (kg) (P = 0.227)	% F** (P = 0.004)	%P* (P = 0.047)	SCC (P = 0.916)
AA	1797.17	114.21^{b}	77.09	6.62^{b}	4.27^{a}	4.55
GA	1814.26	112.83^{b}	76.62	6.42^{b}	4.17^{ab}	4.60
GG	2037.73	144.79^{a}	85.27	7.45^{a}	4.16^{b}	4.56

^{a,b}Values within columns with different superscripts differed significantly at **P < 0.01 or *P < 0.05.

DISCUSSION

The aim of the current study was to identify polymorphisms in *A2M* in dairy buffaloes, as well as to investigate associations with milk production, quality, and SCC. Investigating SCC in dairy cattle, Wang et al. (2012) reported that animals with low SCCs and high resistance to mastitis

had the “A” allele in SNP c.3535A>T. In our study, all of the buffaloes had this “A” allele in the same position as described by those authors, but in homozygosity (AA). We suggest that the “A” allele is fixed in this buffalo population, thus contributing to an increased resistance to mastitis.

Our results indicate that SNP g.241A>G is associated with fat production and percentage, as well as protein percentage, in buffalo milk. For fat production and percentage, the GG genotype had the highest phenotypic means, whereas for protein percentage, the AA genotype had the highest means, which significantly statistically differed from the GG genotype. Therefore, we identified a new molecular marker (g.241A>G) for the traits described above.

Another important finding was that SNP g.241A>G changes the codon from CAT to CGT. The first codon is translated in histidine amino acid, which plays an antioxidant role in biological functions (Rowley and Halliwell, 1983) and limits protein synthesis in milk (Huhtanen et al., 2002). The second is translated in arginine, a precursor in the synthesis of molecules that regulate essential metabolic pathways for health, growth, development, reproduction, and homeostasis (Wu et al., 2007). The changing of this amino acid can influence the biological role of the protein; however, functional studies are needed to investigate the effects of this amino acid substitution.

This is the first description of an *A2M* polymorphism in the buffalo, as well as associations between the SNP and different production traits. SNP g.241A>G was significantly associated ($P < 0.05$) with fat production and fat and protein percentage, and could be used as a molecular marker for these traits.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

A.C. Freitas was awarded a scholarship and G.C. Venturini a postdoctoral fellowship by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). D.F. Cardoso, G.M.F. Camargo, and F. Tanamati were granted scholarships and R.R. Aspilcueta-Borquis a postdoctoral fellowship by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). The CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) funded postdoctoral fellowships to N.B. Stafuzza and N.A. Hurtado-Lugo, and a productivity research fellowship to H. Tonhati.

REFERENCES

Dabdoub SAM and Shook GE (1984). Phenotypic relations among milk yield, somatic cell count and clinical mastitis. *J. Dairy Sci.* 67: 163-164.

Enghild JJ, Salvesen G, Thøgersen IB and Pizzo SV (1989). Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alpha 1-inhibitor-3. *J. Biol. Chem.* 264: 11428-11435.

Falconer DS and Mackay TFC (1996). Introduction to quantitative genetics. 4th ed. Longman Group Limited, Edinburgh.

Gil FMM, de Camargo GMF, Pablos de Souza FR, Cardoso DF, et al. (2013). Polymorphisms in the ghrelin gene and their associations with milk yield and quality in water buffaloes. *J. Dairy Sci.* 96: 3326-3331. http://dx.doi.org/10.3168/jds.2012-6362

Huhtanen P, Vanhatalo A and Varvikko T (2002). Effects of abomasal infusions of histidine, glucose, and leucine on milk production and plasma metabolites of dairy cows fed grass silage diets. *J. Dairy Sci.* 85: 204-216. http://dx.doi.org/10.3168/jds.S0022-0302(02)74069-1

Kumar S, Sangwan ML, Ahlawat S and Barwar A (2011). Polymorphism in *DRB3* exon by PCR-RFLP and its association with mastitis in Murrah buffaloes. *Int. J. Biotechnol.* 10: 232-234.

Lahouassa H, Moussay E, Rainard P and Riollet C (2007). Differential cytokine and chemokine responses of bovine mammary epithelial cells to *Staphylococcus aureus* and *Escherichia coli. Cytokine* 38: 12-21. http://dx.doi.org/10.1016/j.cyto.2007.04.006

Ma H, Wang B, Zhang J, Li F, et al. (2010). Multiple forms of alpha-2 macroglobulin in shrimp *Fenneropenaeus chinesis* and their transcriptional response to WSSV or *Vibrio* pathogen infection. *Dev. Comp. Immunol.* 34: 677-684. http://dx.doi.org/10.1016/j.dci.2010.01.014

Onara DF, Forlenza M, Gonzalez SF, Rakus KL, et al. (2008). Differential transcription of multiple forms of alpha-2-macroglobulin in carp (*Cyprinus carpio*) infected with parasites. *Dev. Comp. Immunol.* 32: 339-347. http://dx.doi.org/10.1016/j.dci.2007.06.007

Otaviano AR, Tonhati H, Sena JAD and Muñoz MFC (2005). Kappa-casein gene study with molecular markers in female buffaloes (*Bubalus bubalis*). *Genet. Mol. Biol.* 28: 237-241. http://dx.doi.org/10.1590/S1415-47572005000200010

Riaz MN, Malik NA, Nasreen F and Qureshi JA (2008). Molecular marker assisted study of kappa-casein gene in Nili-Ravi (Buffalo) breed of Pakistan. *Pak. Vet. J.* 28: 103-106.

Rolim Filho ST, Ribeiro HFL, de Camargo GMF, Cardoso DF, et al. (2013). Identification of polymorphisms in the osteopontin gene and their associations with certain semen production traits of water buffaloes in the Brazilian Amazon. *Reprod. Domest. Anim.* 48: 705-709. http://dx.doi.org/10.1111/rda.12144

Rowley DA and Halliwell B (1983). Formation of hydroxyl radicals from hydrogen peroxide and iron salts by superoxide- and ascorbate-dependent mechanisms: relevance to the pathology of rheumatoid disease. *Clin. Sci.* 64: 649-653. http://dx.doi.org/10.1042/cs0640649

Sambrook J, Fritsch EF and Maniatis T (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Tonhati H, Cerón-Muñoz MF, Oliveira JA, El Faro L, et al. (2008). Test-day milk yield as a selection criterion for dairy buffaloes (*Bubalus bubalis* Artiodactyla, Bovidae). *Genet. Mol. Biol.* 31: 674-679. http://dx.doi.org/10.1590/S1415-47572008000400012

Wang X, Huang J, Zhao L, Wang C, et al. (2012). The exon 29 c.3535A>T in the alpha-2-macroglobulin gene causing aberrant splice variants is associated with mastitis in dairy cattle. *Immunogenetics* 64: 807-816. http://dx.doi.org/10.1007/s00251-012-0639-8

Wu G, Bazer FW, Davis TA, Jaeger LA, et al. (2007). Important roles for the arginine family of amino acids in swine nutrition and production. *Livest. Prod. Sci.* 112: 8-22. http://dx.doi.org/10.1016/j.livsci.2007.07.003

Untergasser A, Nijveen H, Rao X, Bisseling T, et al. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35: W71-4. http://dx.doi.org/10.1093/nar/gkm306

Zetouni L, de Camargo GMF, da Silva Fonseca PD, Gil FMM, et al. (2013). Effects of a single nucleotide polymorphism in the leptin gene on the productive traits of dairy buffaloes (*Bubalus bubalis*). *Mol. Biol. Rep.* 40: 5159-5163. http://dx.doi.org/10.1007/s11033-013-2618-z

Zetouni L, de Camargo GMF, da Silva Fonseca PD, Cardoso DF, et al. (2014). Polymorphisms in the *MTRN1A* gene and their effects on the productive and reproductive traits in buffaloes. *Trop. Anim. Health Prod.* 46: 337-340. http://dx.doi.org/10.1007/s11250-013-0493-1