

## SHORT COMMUNICATION

**The polymorphism in MUC1 gene in Nelore cattle**

F. R. P. de Souza<sup>1</sup>, D. B. Dentillo<sup>1</sup>, J. Meola<sup>1</sup>, F. H. Biase<sup>1</sup>, M. V. Andréa<sup>2</sup>, P. A. Vozzi<sup>1</sup>, R. B. Lôbo<sup>1</sup>, L. R. Martelli<sup>1</sup>

<sup>1</sup> Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Ribeirão Preto, SP, Brazil

<sup>2</sup> Departamento de Zootecnia, Centro de Ciências Agrárias e Ambientais, Universidade Federal da Bahia, Cruz das Almas, BA, Brazil

**Keywords**

Cattle; molecular marker; PCR; VNTR

**Correspondence**

Lucia Martelli, Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, 3900, Avenida Bandeirantes, 14049-900, Ribeirão Preto, SP, Brazil. Tel.: 55(16)3602-3081; Fax: 55 (16)3602-3076; E-mail: lrmartel@fmrp.usp.br

Received: 24 April 2006;

accepted: 7 December 2006

**Summary**

MUC1 is a transmembrane glycoprotein expressed on the apical surfaces of the uterine epithelial tissue with predicted functions in protection and cell–cell adhesion. These properties are closely related with the repetitive region [variable number of tandem repeats (VNTR)] of the extracellular domain and with the *O*-glycosylation in their serine and threonine residues. This study describes a polymerase chain reaction (PCR) protocol to analyse MUC1 bovine genetic polymorphism and demonstrates the existence of a VNTR within the sites for *O*-glycosylation. Oligonucleotide primers based on the *Bos taurus* mucin (MUC1) gene sequence GenBank AF399757 were used to amplify five VNTR MUC1 alleles from a study group of 56 pure Nelore bovines. The number of repeats varied between 10 and 24, being more prevalent than the alleles with less number of repeats. The DNA sequence analysis revealed two repeats and one of them presented 100% homology with the bovine *consensus* sequence already reported. The second repeat showed codons that translate to serine and proline amino acids, which are conserved in the MUC1 of several species. This study is the first description of allelic variation and the VNTR structure in the Nelore breed MUC1 gene, and we suggest that this genetic variability can be tested for association with variation in reproductive traits in Nelore cattle.

**Introduction**

MUC1 is a transmembrane glycoprotein mucin exclusively expressed in apical mammalian epithelial tissues (Duraismy *et al.* 2006). The broad functions of this glycoprotein are cell surface protection, lubrication and cell–cell adhesion but the glycoprotein can also play a role in growth, fetal development, renewal and in epithelial differentiation (Moniaux *et al.* 2001). In the female reproductive tract, MUC1 has been associated with embryo implantation control where it is thought to provide a barrier between the blastocyst and the epithelial cells of the endometrium (Brayman *et al.* 2004). For this role, studies of gene expression during implantation in ovines and other species show progesterone-dependent down-

regulation of MUC1 in the uterus (Johnson *et al.* 2001).

The structure of the MUC1 molecule is well defined in many species, including bovines. In Danish Holstein cattle, the mature protein comprises 558 amino acid residues that constitute a phosphorylated cytoplasmic tail, a hydrophobic transmembrane domain and an extracellular domain that includes an SEA domain, and a highly polymorphic variable number of tandem repeats (VNTR) region that encodes 20 amino acids. With the exception of this VNTR region, all the other domains of MUC1 are highly conserved (Pallensen *et al.* 2001). The VNTR region has also been called PTS because of the proline, threonine and serine residues that constitute the *consensus* amino acid sequence. The threonine

and serine residues serve as sites for the *O*-glycosylation responsible for both positive and negative adhesive properties of the molecule. The proline residues are associated with this role, giving the molecule an extended form and raising it well above the surface of the cell membrane (Patton *et al.* 1995; Gendler 2001; Dekker *et al.* 2002).

In the eight species studied, the PTS-conserved *consensus* amino acid sequence included one serine and two proline residues highly conserved. The remaining sequence presents high content of these residues with a significant degree of conservation (Pallensen *et al.* 2001; Rasero *et al.* 2002; Sacchi *et al.* 2004). The lack of perfect conservation between species is caused by the position of these residues varying within the repeat, as the VNTR varies in length. In bovines and caprines the repeats can show 80–100% of homology with the derived PTS *consensus* repeat. In Italian Friesian and Piedmontese breeds all repeats in the VNTR share on average 93% DNA sequence homology with the *consensus* sequence. In mouse MUC1, this homology between repeats is down to 75% on average (Rasero *et al.* 2002; Sacchi *et al.* 2004).

DNA polymorphism in the VNTR of dairy cattle and caprines has been used for association studies with improved selection schemes in mind. In taurus bovine subspecies, four alleles for the VNTR region have been described with size ranging from 778 base-pairs (bp) to approximately 1600 bp (Rasero *et al.* 2002; Sacchi *et al.* 2004).

The role of the MUC1 protein in the uterine environment suggests it may be an informative marker for variation in fertility traits. The aim of this study was thus to develop a polymerase chain reaction (PCR) protocol for analysing polymorphism of the MUC1 VNTR in the Brazilian Nelore breed for future association studies of traits such as early puberty, and the success rate of embryo implantation obtained with *in vitro* fertilization techniques (IVF).

## Materials and Methods

### Sample collection

The study material comprised DNA derived from 56 pure Nelore breed fertile cows selected in the Program for Genetic Improvement of the Nelore Breed (PMGRN) at the FMRP – University of Sao Paulo, Brazil. The sample was characterized by low values of relationship coefficient and inbreeding. We found a similar effective number of founders as reported by Vozzi *et al.* (2006) on the breed indicating that the sample was representative of the Brazilian Nelore cattle.

### DNA extraction

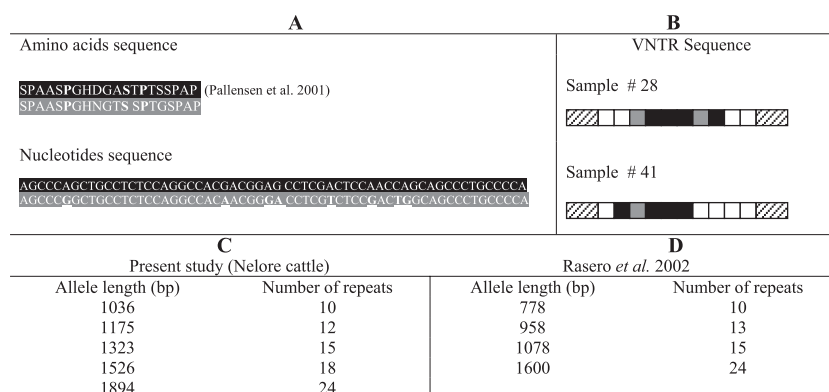
Molecular genetic analysis was performed on genomic DNA obtained by extraction from peripheral blood lymphocytes. DNA was extracted by precipitation in NaCl using standard techniques (Olerup & Zetterquist 1992).

### MUC1 polymorphism analysis

The forward and reverse primer sequences (5'-CGC AGA ACT ACG CCA GTT TCC-3' and 5'-AGA GCG GGT GGT CAT GGA TG-3') were based on the bovine sequence GenBank AF399757 and adapted from the primer sequences which flank the repetitive VNTR region of the bovine MUC1 published by Rasero *et al.* (2002). About 100 ng of genomic DNA was mixed with 15 pmol of each primer in a total volume of 25 µl containing 200 µM of each dNTP, 0.75 mM MgSO<sub>4</sub>, 0.5 U of Platinum Pfx *Taq* DNA polymerase, Platinum Pfx Amplification Buffer and PCR Enhancer Solution (Invitrogen, Carlsbad, CA, USA), both in the concentration of 1X. PCR was performed on a Whatman Biometra TGradiente Thermocycler at the following conditions: a denaturation step at 95°C for 5 min, 35 cycles of 95°C for 40 s, 58°C for 40 s, 68°C for 1 min and 30 s, and a final extension step of 68°C for 5 min. The amplified fragments were checked in a 1.5% agarose gel electrophoresis prepared with 0.5X Tris-borate-ethylenediaminetetraacetic acid (EDTA) buffer. The gel was photographed and allele lengths were estimated by the software Kodak Digital Science 1D Image Analysis Software with the Marker 1 kb Plus DNA Ladder molecular weight standard.

### DNA sequencing and VNTR fragment analysis

Polymerase chain reaction purification derived from two independent Nelore homozygotes (no. 28 and no. 41 in Figure 1b) for the same allele (1036 bp) was performed with GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences, Pittsburgh, PA, USA). For quantification, the samples were applied in a 2% agarose gel with Low DNA Mass Ladder (Invitrogen, Carlsbad, CA, USA). The sequencing was performed in the MegaBACE<sup>TM</sup> 1000 Analysis System (Amersham Biosciences) with DYEnamic ET Dye Terminator Cycle Sequencing Kit and Thermo Sequenase<sup>TM</sup> II DNA Polymerase and the sequence information was entered into GenBank as DQ644555 on 12 April 2006. The sequences were analysed by the Sequence Analyzer software using the Base Caller Cimarron 3.12 (Amersham Biosciences, Pittsburgh,



**Figure 1** MUC1 Gene repeat organization: (a) nucleotides and predictive amino acids sequences of each repeat. Conserved amino acids are shown in bold. The underlined nucleotides indicate the differences between the sequences. (b) Schematic representation of the VNTR sequence to show the organization of the MUC1 VNTR repeat in samples no. 28 and no. 41. The consensus sequence is black, and the second repeat is dark grey. The non-repeat sequence and low homology sequences are represented by striped and white boxes, respectively. (c) Allele length and the respective number of repeats in *Bos taurus indicus* (present study). (d) Allele length and the respective number of repeats in previous study of *Bos taurus taurus* (Rasero *et al.* 2002).

PA, USA). Both purification and sequencing reactions were performed by the DNA sequencing and PCR products purification services of Human Genome Research Center (CEGH).

Chromas software (<http://www.technelysium.com.au>) was used to compare the homology with other sequences in GenBank. The Tandem Repeat Finder Program (<http://tandem.bu.edu/trf/trf.html>) was used to identify and count the number of repetitions of the sequence (Benson 1999). The estimation of the number of repeats in other alleles was performed utilizing the following formula:

$$ETR = (NCR)(NR) - (L1036 \text{ bp})$$

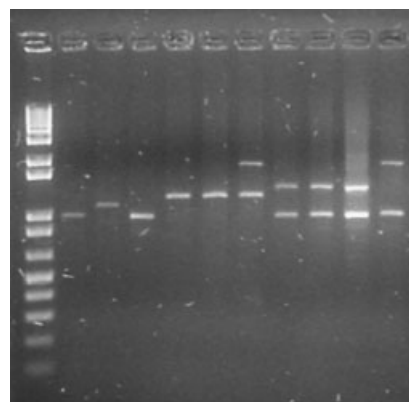
where ETR is the size (bp) of the non-repeat sequence, NCR the bp number in the consensus repeat, NR the number of repeats in the 1036 sequence and L1036 bp the bp length of the allele 1036.

$$NTR(X) = \frac{FNS(X) - ETR}{60}$$

where X is the allele, NTR the number of repeats and FNS the allele length.

## Results

Five alleles were identified in the VNTR MUC1 locus and they were initially named according to the size, based on estimates provided by the Image Analysis Software (Figure 2). The 1036-bp allele was the most prevalent, followed by the 1323- and 1175-bp alleles. The 1526- and 1894-bp alleles were the least frequent (Table 1).



**Figure 2** Gel electrophoresis showing the PCR products of 10 DNA samples. From left to right: 1 kb Plus DNA Ladder, alleles 1036, 1175, 1036, 1323, 1323, 1323/1894, 1036/1526, 1036/1526, 1036/1526 and 1036/1894 bp. Sizes are estimated by algorithm Kodak Digital.

The sequencing of two independent samples from animals (no. 28 and no. 41) homozygous for the 1036-bp allele gave rise to sequences of 967 bp (GenBank DQ644555) and 1011 bp long. In both sequences, Tandem Finder Repeat analysis demonstrated a VNTR formed by 10 repeats. The no. 28 sequence showed four repeats which share 100% of homology with the *consensus* repeat originally described by Pallensen *et al.* (2001) and Rasero *et al.* (2002). The other six repeats comprised sequences with partial homology to the *consensus* repeat, named 'imperfect' sequence by Rasero *et al.* (2002). Two of them presented the codons that translate to serine and proline in conserved positions between all species that have been studied to date. The VNTR

**Table 1** Allelic and genotypic frequencies in 56 Nelore cattle

Allele	Frequency	Genotype	Frequency
1036	0.71	1036/1036	0.5
1175	0.107	1036/1175	0.17
1323	0.133	1036/1323	0.17
1526	0.026	1036/1526	0.053
1894	0.0178	1036/1894	0.017
		1175/1175	0.017
		1323/1323	0.035
		1323/1894	0.017

sequence of the sample no. 41, comprised four *consensus* repeats, and six imperfect repeats. One imperfect repeat showed conserved codons for proline and serine residues. The number of repeats in the other alleles was estimated to be 12, 15, 18 and 24. With the exception of the alleles with 12 and 18 repeats, the other alleles were described in *Bos taurus taurus* by Rasero *et al.* (2002) (Figure 1c,d).

## Discussion

Mucin genes are characterized by the unusual presence of intragenic repeats within the transcript. Most genes with variable number of tandem repeats (VNTR) in the coding region are surface proteins involved in cell–cell interactions. The quantitative alterations in the cell-surface phenotypes are manifested as differential adhesiveness and different levels of flocculation in cells (Verstrepen *et al.* 2005). The wide variability in the functional efficiency of these molecules arises from length variation of the functional VNTR, caused by misalignment of tandem repeats and unequal crossing-over at meiosis.

In the reproductive tract of rodents, pigs and ruminants, the glycoprotein MUC1 is repressed by progesterone control at the time of embryo implantation. During the same period, in humans and rabbits, progesterone stimulates uterine MUC1 expression, and its repression is limited to the site of embryo implantation. These studies suggest that variables in the VNTR region of the molecule can influence the interaction between the embryo and uterine luminal cells by 'steric hindrance' or different patterns of *O*-glycosylation (Aplin *et al.* 2001).

For the function involving differential cellular adhesiveness, the MUC1 VNTR polymorphisms have been associated with genetic susceptibility to increased risk for metastasis in gastric carcinoma and for infertility (Carvalho *et al.* 1997; Horne *et al.* 2001; Goulart *et al.* 2004). In goats and bovines, MUC1 has been associated with yield, health and reproductive

traits (Hens *et al.* 1995; Patton *et al.* 1995; Sacchi *et al.* 2004). The present study shows five new allelic variations in the Nelore cattle in which the alleles with 12 and 18 repeats have not been previously described in *Bos taurus* Piedmontese and Italian Friesian cattle (Rasero *et al.* 2002). The higher frequency of the alleles with minor sizes presented in this study is in contrast to the report of Sacchi *et al.* (2004) who found the prevalence of alleles with moderate sizes in five cattle breeds.

The *consensus* sequence described by Spicer *et al.* (1995), differs from the sequence described by Pallesen *et al.* (2001) and from the findings we present. This variation may reflect genetic differences among breeds or exemplify the relative instability of the VNTR region in these populations. Despite the variation in the extent of homology for some repeats, the repeats typically maintain proline, threonine and serine residues, suggesting that the major selective pressure for MUC1 protein is to maintain the potential *O*-glycosylation sites.

We conclude that the high variability of MUC1 polymorphism in Brazilian Nelore cattle, the VNTR structure, and the presence of codons that translate for serine and threonine residues in the VNTR *consensus* sequence provide new information for finding association between the polymorphisms at the gene and the variation in fertility traits in cattle. In addition to these traits, the polymorphism may be associated with other economically important characteristics relevant to these breeds such as yield and overall health.

## Acknowledgements

This work was supported by CAPES, ANCP and PRONEX/CNPq. We thank Dr Jeremy A. Squire from University of Toronto for suggestions and text revision.

## References

- Aplin J.D., Meseguer M., Simón C., Ortiz M.E., Croxatto H., Jones C.J.P. (2001) MUC1 glycans and the cell-surface barrier to embryo implantation. *Biochem. Soc. Trans.*, **29**, 153–156.
- Benson G. (1999) Tandem repeats finder: a program to analyse DNA sequences. *Nucleic Acids Res.*, **27**, 573–580.
- Brayman M., Thathiah A., Carson D.D. (2004) MUC1: A multifunctional cell surface component of reproductive tissue epithelia. *Reprod. Biol. Endocrinol.*, **2**, 4.
- Carvalho F., Seruca R., David L., Amorim A., Seixas M., Bennett E., Clausen H., Sobrinho-Simões M. (1997)

- MUC1 gene polymorphism and gastric cancer – an epidemiological study. *Glycoconj. J.*, **14**, 107–111.
- Dekker J., Rossen J.W., Buller H.A., Einerhand A.W. (2002) The MUC1 Family: an obituary. *Trends Biochem. Sci.*, **3**, 126–131.
- Duraisamy S., Ramasamy S., Kharbanda S., Kufe D. (2006) Distinct evolution of the human carcinoma-associated transmembrane mucins, MUC1, MUC4 and MUC16. *Gene*, **373**, 28–34.
- Gendler S.J. (2001) MUC1, the renaissance molecule. *J. Mammary Gland Biol. Neoplasia*, **6**, 339–353.
- Goulart L.R., Vieira G.S., Martelli L., Inácio J., Goulart I.M., Franco J.G., Jr (2004) Is MUC1 polymorphism associated with female infertility? *Reprod. Biomed. Online*, **8**, 477–482.
- Hens J.R., Rogers G.W., Huot M.L., Patton S. (1995) Association of the epithelial mucin, PAS-1, with yield, health and reproductive traits in Holstein dairy cows. *J. Dairy Sci.*, **78**, 2473–2480.
- Horne A.W., White J.O., Margara R.A., Williams R., Winston R.M.L., Lalani E.-N. (2001) MUC1: a genetic susceptibility to infertility? *Lancet*, **357**, 1336–1337.
- Johnson G.A., Bazzar F.W., Jaeger L.A., Ka H., Garlow J.E., Pfarrer C., Spencer T.E., Burghardt R.C. (2001) Muc-1, integrin and osteopontin expression during the implantation cascade in sheep. *Biol. Reprod.*, **65**, 820–828.
- Moniaux N., Escande F., Porchet N., Aubert J.P., Batra S.K. (2001) Structural organization and classification of the human mucin genes. *Front. Biosci.*, **6**, 1192–1206.
- Olerup O., Zetterquist H. (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. *Tissue Antigens*, **39**, 225–235.
- Pallensen L.T., Andersen M.H., Nielsen R.L., Berglund L., Petersen T.E., Rasmussen L.K., Rasmussen J.T. (2001) Purification of MUC1 from bovine milk-fat globules and characterization of a corresponding full-length cDNA clone. *J. Dairy Sci.*, **12**, 2591–2598.
- Patton S., Gendler S.J., Spicer A.P. (1995) The epithelial mucin, MUC1, of milk, mammary gland and other tissues. *Biochim. Biophys. Acta*, **1241**, 407–424.
- Rasero R., Sacchi P., Rosati S., Cauvin E., Maione S. (2002) Molecular analysis of the length polymorphic MUC1 gene in cattle. *J. Anim. Breed. Genet.*, **119**, 342–349.
- Sacchi P., Caroli A., Cauvin E., Maione S., Sartore S., Soglia D., Rasero R. (2004) Analysis of the MUC1 gene and its polymorphism in *Capra hircus*. *J. Dairy Sci.*, **87**, 3017–3021.
- Spicer A.P., Duhig T., Chilton B.S., Gendler S.J. (1995) Analysis of mammalian MUC1 genes reveals potential functionally important domains. *Mamm. Genome*, **12**, 885–888.
- Verstrepen K.J., Jansen A., Lewitter F., Fink G.R. (2005) Intragenic tandem repeats generate functional variability. *Nat. Genet.*, **37**, 986–990.
- Vozzi P.A., Marcondes C.R., Magnabosco C.U., Bezerra L.A.F., Lobo R.B. (2006) Structure and genetic variability in Nelore (*Bos indicus*) cattle by pedigree analysis. *Genet. Mol. Biol.*, **29**, 482–485.