

**MOLECULAR CHARACTERIZATION OF PARTIAL SEQUENCES OF
BRAZILIAN BLV GAG SEQUENCES**

Marcelo Fernandes Camargos¹, Jenner Karlisson Pimenta dos Reis², Rômulo Cerqueira Leite².

¹LANAGRO, Minas Gerais

²Escola de Veterinária, Universidade Federal de Minas Gerais

Corresponding author:

Marcelo Fernandes Camargos, LANAGRO/MG, Av. Rômulo Joviano s/n, Caixa Postal 50, 33.600-000 Pedro Leopoldo, MG, Brazil.

Tel.: 55-31-3660-9635; fax: 55-31-3661-2383

E-mail: mfcamargos@yahoo.com.br

Key words: bovine leukemia virus, GAG sequences, PCR

ABSTRACT

PCR amplification and sequencing of *gag* gene fragments of two Brazilian strains of BLV were performed. After comparison of the Brazilian sequences with four sequences originating from other countries, available in the GenBank, a high conservation of this gene was observed and the results confirmed that two of the available sequences in the GenBank had sequencing errors.

INTRODUCTION

The Bovine Leukemia Virus (BLV) *gag* gene codes for a polypeptide that originates the matrix protein, the capsid protein (CA, p24) and the nucleoprotein. The p24 protein has one of the most conserved regions of the *gag* gene, the major homology region (MHR) constituted by 20 aminoacids, which is present in the CA of most retroviruses and has an essential role in viral infectivity (Coffin et al. 1997, Willems et al. 1997). The p24 protein also contributes to *gag* polyprotein packing during viral assembly and it is important in the early stages of infection by interaction with cyclophilin A (Gamble et al. 1996). Two sequential epitopes were found in the p24 protein (Mager et al. 1994). The p24 aminoacid sequences of Human Immunodeficiency Virus (HIV-1) had a variation level of 6.2 % to 8.5 % (Iroegbu et al. 2000). There were only four BLV p24 *gag* sequences available in Genbank, and in two of these sequencing errors probably occurred (Rice et al. 1985, Couslton et al. 1990, Dube et al. 2000). The goal of this work was to characterize two Brazilian *gag* BLV partial sequences and compare them with the four sequences available in GenBank.

MATERIAL AND METHODS

DNA samples were obtained from blood of two BLV naturally infected bovines, identified as 30 and 384, using GFX Genomic Blood DNA Purification Kit[®] (Amersham Pharmacia Biotech, USA) according to manufacturer's recommendations. PCR primers forward (5'-ggaaatcgcaaccgcat-3' 760-777) and reverse (5'-tggggtgtggacgaggat-3' 1384-1401) had an annealing temperature of 50°C. The numbers correspond to those in BLV sequence M10987 (Rice et al., 1985). PCR and electrophoresis were done as described by Camargos et al. (2003). PCR products were purified with Wizard[®] PCR Preps DNA Purification System Kit[®] (Promega, USA) according to manufacturer's instructions. The purified PCR products had been directly sequenced using the DNA Sequencing Kit Big Dye Terminator Version III (Applied Biosystems, USA) and the automated sequencer ABI PRISM 377 DNA SEQUENCER[™] (Applied Biosystems, USA). The program Bioedit Sequence Alignment Editor version 5.0.9 (Hall 1999), was used for editing, alignment, translation of the nucleotide sequences and obtaining a consensus sequence from the sequences AF257515 (Argentina), D00647 (Australia), K02120 (Japan) and M10987 (Belgium). This consensus sequence was used to calculate the number of nucleotide and aminoacid substitutions. The software BlastP (<http://www.ncbi.nlm.nih.gov/Blast>) was used to search for conserved domains in the translated nucleotide sequences. The software SIFT (<http://blocks.fhrc.org/sift/SIFT>) (Ng and Henikoff, 2001) was used to verify if aminoacid substitutions would change BLV integrase properties. The software MEGA version 2.1 (Kumar et al. 2001) and the nucleotide substitution model Kimura 2-P was used to verify the genetic variation between partial nucleotide sequences of the BLV *pol* gene.

RESULTS AND DISCUSSION

PCR amplified 640 base pair fragments. After sequencing with forward and reverse primers the sequences were submitted to GenBank and received the following access numbers 30 (AY589727) and 384 (AY277948). The average nucleotide substitutions per sequence, compared to the consensus sequence, was 11.6 (2.1 %) and up to 6.1 % variation between the sequences was observed. Sixty-one of 70 substitutions were of the transition type (87.1 %), predominantly C→T and A→G. In all sequences deletions and insertions were observed. Insertions and deletions were not observed in BLV *pol* and *env* sequences (Coulston et al. 1990, Dube et al. 1997, Dube et al. 2000, Camargos et al. 2002, Monti et al. 2005). The sequences D00647 and K02120 probably had sequencing errors, as mentioned by others (Dube et al. 2000). When these sequences were taken out of the alignment deletions and insertions were not observed anymore. A consensus sequence was obtained with the remaining sequences to recalculate the substitution numbers, and 31 nucleotide substitutions were observed, with an average of 7.7 per sequence (1.4 %), most of them silent (74.2 %) and transitions (83.8 %). These results are similar to these obtained in previous reports (Coulston et al. 1990, Dube et al. 1997, Dube et al. 2000, Camargos et al. 2002). The sequences D00647 and K02120 were excluded from aminoacid alignment (Fig. 1).

	1	11	21	31	41	51	61	71	81	91
Consensus	GFGSPETIA-	-LFESLSEGW	TPYDWNQLLN	AVPSHGGDKQ	LLKSEINEEA	AEWDRAHRAP	QIPGTGSTQQ	QQAQPQFDGL	RQOYRQWILL	AWRKLPMQGH
30	<u>APGSOVWIOT</u>	<u>LRLAILLOADP</u>	<u>TPADLEQLCQ</u>	YIASPVDQTA	HMTSLTAAIA	AAEAANTLQG	FNPQNGTLTQ	QSAQPNAGDL	RSQYQNLWLQ	AWKNLPTRPS
384	-----T-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
AF257515	-----	-----	-----	-----	-----	-----	LT-KT-----	-----	-----	-----
M10987	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
	101	111	121	131	141	151	161	171	181	
Consensus	PTPLADIRQG	PKEPYQDFVD	RLQEALDAEQ	ADGEVKDYLT	QTLAYENANP	DCQSALRPLG	-FNATLEEKL	RACQDVG		
30	VQPWSTIVQG	<u>PAESYVEFVN</u>	<u>RLQISLADNL</u>	<u>PDGVPKEPII</u>	DSLSYANANK	ECQILQGRG	LVAAPVGQKL	QACAHWAPKI	KQP	
384	-----S-----	-----	-----	-----	-----	-----	-----	-----	-----	
AF257515	-----	-----	-----	-----	-----	-----	-----	-----	-----	
M10987	-----	-----	-----	-----	-----	-----	-----	-----	-----	v

Figure 1 - Alingment of partial aminoacid BLV *gag* sequences. The similarity between the evaluated sequences and the consensus sequence is indicated by a trace (-). The above consensus sequence presents the conserved domain of the p24 protein of the *gag* gene. The epitope sequences (2-26) and (112-136) are underlined and the MHR is in italics (107-126)

Seven aminoacid residues were changed, the residue 8 I→T (sequence 384); the residue 110 G→S (sequence 30); the residues 61 F→L, 62 N→T, 64 Q→K and 65 N→T (sequence AF257515) and the residue 180 I→V (M10987) which give an average of 1.75 (0.9 %) change per sequence, showing that this protein is highly conserved. The BlastP software showed that part of the p24 aminoacid sequence obtained belongs to the conserved domain p24 of the *gag* gene. Six of the seven changes were localized in this domain. All aminoacid changes evaluated by SIFT were predicted to be tolerable except the change I→T, but this change had low confidence which may be due to the low genetic variation and small number of sequences available. The residue 8 was the only one that varied in the BLV p24 epitopes (Mager et al. 1994). The four MHR residues conserved in all Retrovirus, Q-109, E-113, F-118 and R-121, were conserved in all BLV sequences (Wills & Craven 1991). To verify if aminoacid substitutions in epitopes, MHR and in the conserved domain of p24 would change the infectivity or reactivity of BLV it is necessary to do mutagenesis studies. It is necessary to do more sequencings of the BLV *gag* gene to obtain a better understanding of molecular and phylogenetic characteristics of this BLV gene.

REFERENCES

Camargos MF, Stancek D, Rocha MA, Lessa LM, Reis JKP, Leite RC 2002. Partial sequencing of env gene of bovine leukaemia virus from Brazilian samples and phylogenetic analysis. *J. Vet. Med.* B 49: 325-331.

Camargos MF, Stancek D, Rocha MA, Lessa LM, Reis JKP, Leite RC 2003. Development of a polymerase chain reaction and its comparison with agar gel

immunodiffusion test in the detection of bovine leukemia virus infection. *Braz. J. Vet. Res. Anim. Sci.* 40: 341-348.

Coffin JM, Hughes SH, Varmus HE 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, New York.

Coulston J, Naif H, Brandon R, Kumar S, Khan S, Daniel RCW, Lavin MF 1990. Molecular cloning and sequencing of an Australian isolate of proviral bovine leukaemia virus DNA: comparison with other isolates. *J. Gen. Virol.* 71: 1737-1746.

Dube S, Bachman S, Spicer T, Love J, Choi D, Esteban E, Ferrer JF, Poiesz BJ, 1997. Degenerate and specific PCR assays for the detection of bovine leukaemia/lymphoma virus *pol* DNA and RNA: phylogenetic comparisons of amplified sequences from cattle and primates from around the world. *J. Gen. Virol.* 78: 1389-1398.

Dube S, Dolcini G, Abbot L, Mehta S, Dube D, Gutierrez S, Ceriani C, Esteban E, Ferrer J, Poiesz B, 2000. The complete genomic sequence of a BLV strain from a Holstein cow from Argentina. *Virology* 277: 379-386.

Gamble TR, Vajdos FF, Yoo S, Worthylake DK, Houseweart M, Sundquist WI, Hill CP 1996. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* 87: 1285-1294.

Hall TA 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41: 95-98.

Iroegbu J, Birk M, Lazdina U, Sonnerborg A, Sallberg M 2000. Variability and immunogenicity of human immunodeficiency virus type 1 p24 gene quasispecies. *Clin. Diagn. Lab. Immunol.* 7: 377-383.

Kumar S, Tamura K, Jakobsen IB, Nei M 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* 17: 1244-1245.

Mager A, Masengo R, Mammerickx M, Letesson JJ 1994. T cell proliferative response to bovine leukaemia virus (BLV): identification of T cell epitopes on the major core protein (p24) in BLV-infected cattle with normal haematological values. *J. Gen. Virol.* 75: 2223-2231.

Monti G, Schrijver R, Beier D 2005. Genetic diversity and spread of Bovine leukaemia virus isolates in Argentine dairy cattle. *Arch Virol.* 150: 443-458.

Ng PC, Henikoff S 2001. Predicting deleterious amino acid substitutions. *Genome Res.* 11: 863-874.

Rice NR, Stephens RM, Burny A, Gilden RV 1985. The gag and pol genes of bovine leukemia virus: nucleotide sequence and analysis. *Virology* 142. 357-377.

Willems L, Kerkhofs P, Attenelle L, Burny A, Portetelle D, Kettmann R 1997. The major homology region of bovine leukaemia virus p24^{gag} is required for infectivity in vivo. *J. Gen. Virol.* 78: 637-640.

Wills J, Craven R 1991. Form, function, and use of retroviral gag proteins. *AIDS* 5: 639-654.