## RESEARCH



**Open Access** 

# A detailed molecular analysis of complete Bovine Leukemia Virus genomes isolated from B-cell lymphosarcomas

Gonzalo Moratorio<sup>1,2</sup>, Sabrina Fischer<sup>1,2</sup>, Sergio Bianchi<sup>1,3</sup>, Lorena Tomé<sup>1,4</sup>, Gonzalo Rama<sup>1,5</sup>, Gonzalo Obal<sup>1,4</sup>, Federico Carrión<sup>1</sup>, Otto Pritsch<sup>1,4</sup> and Juan Cristina<sup>2\*</sup>

## Abstract

It is widely accepted that the majority of cancers result from multiple cellular events leading to malignancy after a prolonged period of clinical latency, and that the immune system plays a critical role in the control of cancer progression. Bovine leukemia virus (BLV) is an oncogenic member of the Retroviridae family. Complete genomic sequences of BLV strains isolated from peripheral blood mononuclear cells (PBMC) from cattle have been previously reported. However, a detailed characterization of the complete genome of BLV strains directly isolated from bovine tumors is much needed in order to contribute to the understanding of the mechanisms of leukemogenesis induced by BLV in cattle. In this study, we performed a molecular characterization of BLV complete genomes from bovine B-cell lymphosarcoma isolates. A nucleotide substitution was found in the glucocorticoid response element (GRE) site of the 5' long terminal repeat (5'LTR) of the BLV isolates. All amino acid substitutions in Tax previously found to be related to stimulate high transcriptional activity of 5'LTR were not found in these studies. Amino acid substitutions were found in the nucleocapsid, gp51 and G4 proteins. Premature stop-codons in R3 were observed. Few mutations or amino acid substitutions may be needed to allow BLV provirus to achieve silencing. Substitutions that favor suppression of viral expression in malignant B cells might be a strategy to circumvent effective immune attack.

### Introduction

Bovine leukemia virus (BLV) is a B-lymphotropic oncogenic member of the Retroviridae family that infects cattle worldwide and is the causative agent of enzootic bovine leukosis (EBL), a neoplastic proliferation of B cells [1,2]. BLV infection is characterized by a long period of viral latency and by the absence of viremia. This is thought to be related to the transcriptional repression of viral expression in vivo [3]. Latency is likely a viral strategy to evade the host immune response, thereby allowing tumor development [4,5]. In fact, B lymphocytes harboring an integrated provirus do not produce detectable levels of viral RNA or proteins [6]. Nevertheless, when these cells are isolated and cultured in vitro, a marked increase in viral transcription occurs, suggesting that the provirus is maintained at a repressed stage in vivo [7].

Regarding genome organization, as in all retroviruses, BLV has the gag, pro, pol, env structural genes (from 5' to 3' of the genome) required for the production of infectious virions [8]. In addition to these genes, the BLV genome contains an X region located between the env gene and the 3' long terminal repeat (3'-LTR) [9], as also observed in other Deltaretroviruses [10]. This region contains the open reading frames of four regulatory proteins: the transactivator protein, Tax [11]; the Rex protein, which stabilizes and allows exportation through the cytoplasm of viral RNA [12] and two accessory proteins R3 and G4 whose small open reading frames (ORF) are located in the region between the env gene and the tax/rex genes [13]. Deletion of R3 and G4 genes of BLV in an infectious and tumorigenic BLV molecular clone induced loss of the leukomogenic phenotype and



© 2013 Moratorio et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

<sup>\*</sup> Correspondence: cristina@cin.edu.uy

<sup>&</sup>lt;sup>2</sup>Laboratorio de Virología Molecular, Centro de Investigaciones Nucleares, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay

Full list of author information is available at the end of the article

G4 exhibited oncogenic potential both in vivo and in vitro [14,15].

The BLV transcriptional promoter is located in the 5' long terminal repeat (5'-LTR) and is composed of the U3, R and U5 regions. Gene expression is induced at the transcriptional level by the virus-encoded transactivator Tax [16].

Few complete genomic sequences of BLV strains are available in the databases. These sequences are from different sources: peripheral blood mononuclear cells (PBMC) [17], tumor cells, experimentally infected sheep, and cell lines (FLK). The degree of genetic variation among these strains and those directly isolated from bovine B-cell lymphosarcomas remains unknown. For this reason, and in order to contribute to the understanding of the mechanisms of leukemogenesis induced by BLV, we performed a detailed characterization of the complete genome of three BLV isolates from B-cell lymphosarcomas of three cows from different farms, and we compared them with all available and corresponding full length sequences from BLV isolates from other sources.

### Materials and methods

Lymphosarcoma samples were obtained by certified veterinary doctors following appropriate ethical guidelines from national and international veterinary associations. The project was also read and approved by Institut Pasteur-Montevideo, Uruguay.

### Animals

Lymphosarcoma samples were obtained from three dairy cows proven to be infected with BLV by PCR and ELISA (VMRD Inc., Pullman, WA, USA).

### DNA extraction and PCR amplification

DNA samples were extracted from lymphosarcoma tissue and FLK cells (as a control), using the QiAmp DNA Blood Mini kit from QIAGEN, according to the instructions supplied by the manufacturer. PCR amplification of overlapping genome fragments covering the complete genome of BLV was achieved using Phusion DNA Polymerase (New England BioLabs) and specific primers designed for this study (synthesized by Integrated DNA Technologies, Leuven, Belgium and shown in Additional file 1). The location of each amplicon is shown in Additional file 2. Reagents for PCR were from New England BioLabs. The final reaction mixture (50  $\mu$ L) contained 1x HF buffer, 200 µM dNTP, 200 nM of each primer, and 1 U Taq polymerase. The cycle for the PCR amplification were as follows: 98°C for 30 s, then 30 cycles of denaturation at 98°C for 10 s, annealing at 55-65°C for 30 s, and extension at 72°C for 1-3 min, followed by a final extension at 72°C for 10 min. The PCR reactions were carried out using an Eppendorf Mastercycler Gradient PCR Thermal Cycler.

### Amplicon purification and cloning

Amplicons were resolved by 1% agarose gel electrophoresis, stained with ethidium bromide and purified using QIAquick PCR Purification Kit from QIAGEN, according to instructions from the manufacturers, and cloned into pGEM T- Easy vector (Promega). Electrocompetent XL1-Blue bacteria were transformed by colonies and were expanded and small-scale plasmid purification was performed using the GFX DNA purification kit (GE Healthcare, Piscataway, NJ, USA).

### Sequencing

Both strands of purified plasmids were sequenced in order to avoid discrepancies by using specific and universal T7 or SP6 primers and the Big Dye DNA sequencing kit (Perkin-Elmer) on a 373 DNA sequencer apparatus (Perkin-Elmer). Complete genome sequences were obtained from B-cell lymphosarcomas and deposited in the EMBL database under accession numbers EMBL:HE967301 to EMBL: HE967303 (LS1to LS3). Complete genome sequences were obtained for all available and comparable BLV strains by using All-round Retrieval of Sequence and Annotation (ARSA) at the DNA Data Bank of Japan (DDBJ) [18].

### Sequence alignment

Sequences were aligned using the CLUSTAL W program [19].

### **Protein sequences**

Protein sequences were obtained by means of in silico translation of nucleotide to amino acid sequences. This was done by using software from the MEGA program [20].

### **Results and discussion**

## Comparison of the 5'-LTR genome region of BLV strains isolated from lymphosarcomas and other origins

BLV initiates transcription at the U3-R junction of the 5'-LTR induced by Tax protein [16]. Transactivation requires the presence of three 21-bp enhancer elements (called Tax-responsive elements, TxRE) located in the U3 region of the 5'-LTR [21]. Each TxRE contains an octanucleotide core sequence corresponding to an imperfectly conserved cyclic AMP-responsive element (CRE), which binds cellular transcription factors like CRE-binding protein (CREB), CRE-modulator  $\tau$  isoform (CREM $\tau$ ), and activating transcription factors 1 and 2 (ATF-1 and ATF-2) [22]. TxRE also contains an E-box sequence, which overlaps each of the three CRE motifs, and binds proteins that belong to the basic helix-loop-helix (bHLH) family of transcription factors, including c-Myc, Max, USF or TFE3 [23]. The U3 region also contains a PU.1/Spi-B binding site [24] and a glucocorticoid responsive-element (GRE) [25]. In addition, BLV expression is regulated by 5'-LTR sequences downstream of the transcription initiation site: a 64-bp downstream activator sequence (DAS) at the 3' end of the R region [26] and an interferon regulatory factor binding site in the U5 region [27]. A scheme showing the positions of all these elements in BLV 5'-LTR is shown in Figure 1.

Comparison of the 5'-LTR genomic sequences of the three BLV lymphosarcoma isolates (LSI) with all available complete BLV genome sequences, revealed that this genome region is highly conserved (Figure 1). The only significant difference between LSI and those isolated from other cell types, e.g. PBMC or FLK cells, is a base substitution found at position 150 (G to A) in the third enhancer element of this region, at the GRE binding site (Figure 1). It has been previously found that GRE confers responsiveness to glucocorticoids such as dexamethasone in the presence of the Tax transactivator [28]. However, in the absence of Tax, mutation of the GRE significantly decreases basal LTR activity as shown in reporter-based assays [25]. This raises the possibility that this substitution may have allowed a better silencing of viral transcription in the lymphosarcoma strains, as a strategy to avoid recognition by the host immune response [25].

## Comparison of deduced amino acid sequences from structural proteins of BLV LSI with those of other origins

In order to detect differences among BLV LSI and isolates from PBMC and other origins, the amino acid sequences of structural proteins encoded by gag, pro, pol, and env genes were aligned. Gag is a polyprotein precursor that is cleaved in the mature virions giving rise to the following: matrix (p15- MA), capsid (p24-CA) and nucleocapsid (p12-NC) proteins (see Figure 2A). NC proteins among all retroviruses share as a major characteristic the presence of a high percentage of basic residues as well as zinc binding domains involved in RNA packaging, both of which are well conserved in all BLV isolates. Indeed, previous studies have shown that substitutions in either basic amino acid residues or zinc finger domains led to a significant reduction in viral RNA packaging [29]. In that sense, a proline to serine (P340S) substitution was observed in the NC protein of all BLV LSI (Figure 2A). This substitution could potentially increase side chain hydrophylicity and be involved in the elimination of the structural restriction related to proline presence. Interactions of NC with RNA sequences, besides those RNA secondary structures of the RNA packaging signal, has been demonstrated for other retroviruses, e.g., murine leukemia virus (MLV) and spleen necrosis virus (SNV). These interactions play an important role in the RNA packaging of these viruses [30,31].

BLV protease (PR) is an aspartic protease with a functional activity involved in gag processing and thus in virion maturation. Previous work proposed a molecular model for BLV PR as well as its substrate specificity, cleavage type sites and inhibitor sensitivity [32]. The comparison of amino acid sequences of PR of BLV LSI with all other sequenced BLV isolates examined in this study is shown in Figure 2B. Only one amino acid substitution (V1651) was found among the BLV lymphosarcoma isolates and is not related to sites previously reported to be involved in BLV PR function via mainchain atoms of peptide substrates or residues predicted to form cleavage subsites [32,33]. Two substitutions can be observed at positions 37–38 in lymphosarcoma BLV isolate LS1, as compared to other genomic sequences including LS2 and 3 isolates (see Figure 2B).

Two amino acid substitutions can be found in the polymerase precursor of all three BLV isolates, one located in the RT (T378A), the other in the endonuclease region (S573P) (see Figure 3).

This substitution could involve important structural changes, but unfortunately, the structure of BLV polymerase as well as other related Deltaretroviruses, like HTLV-1, is currently unknown.

Further studies will be needed to establish if these substitutions can affect polymerase fidelity or processivity.

The Env protein complex is composed of two component subunits: gp51 surface (SU, N- terminal portion) and gp30 transmembrane (TM, C-terminal portion), which remain associated as a functional trimer with three SU subunits linked by disulphide bonds to a spike of three TM subunits [34]. The gp51protein recognizes and binds to cellular receptors, thereby initiating conformational changes that lead to fusion of viral and cellular membranes by gp30 oligomers [35].

Previous studies have shown that the N-terminal portion of mature gp51 plays an important role in virus infectivity [36]. This region is composed of conformational epitopes F, G and H [37] followed by the structural strong turn GYDP, which is conserved in all oncogenic retroviruses [38]. This motif separates the conformational epitope region from the C-terminal domain of gp51 that contains the linear epitopes A, B, D and E [39] (see Figure 4). Comparison of Env protein of BLV LSI with other previously described isolates, reveals an amino acid substitution in SU conformational epitope region (D134N) in a location previously shown to be related to neutralization [39] (Figure 4). However, this substitution has been previously described as a signature of BLV strains circulating in Uruguay, and it is not specific for LS samples [40].

### Comparison of deduced amino acid sequences from nonstructural proteins of BLV LSI and other origins

Previous studies on the functional domains of the BLV Tax protein have identified a putative zinc finger motif

LS1 LS2 LS3 EF600696 FJ914764 K02120 AF257515	1       50       CRE       E       Box       CRE       E       Box       100         TGTATGAAAGATCATGCCGACCTAGGCGCCACCGCCACCGCCGTAAA       CCAGACAGAGACGTCAGCTCGCCAGAGAAAGCTGGTGACGGCAGCTGGTGGGCTAG       100
LS1 LS2 LS3 EF600696 FJ914764 K02120 AF257515	CAT Box PU.1/Spi-B       GRE       150       CRE       Box PROMT       PAS       200 <u>AATCCCCGTA</u> CCTCCCCCAACT       CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
LS1 LS2 EF600696 FJ914764 K02120 AF257515	CAP SITE       U3  →R       250       300         GTTAGCGGCACCAGAAG       GTTCTTCTCCTGAGACCTCGGGCCAGCCTCGGGCCTCTGGGCCGGCC
LS1 LS2 LS3 EF600696 FJ914764 K02120 AF257515	350     DAS     E Box     400       GCGGTCAGGTAAGCCAACCGGGTGGTTCTCGGCGGAGACCACCGCGAGC     TCTATCTCCGGTCCTCGACCGTCGACACCTCC
LS1 LS2 EF600696 FJ914764 K02120 AF257515	$\begin{array}{c c} 450 \\ R \mid \rightarrow \textbf{U5} \\ \hline \textbf{IRF} \\ \hline \textbf{500} \\ \hline \textbf{CCTTTGCCTCTGACCCGGGCT} CCAAGGGGGGTCTGGCTTGGCCCGGGGTTTGTTTCCTGTTTCTGTTTCTGCGGGCCGGGGCTCTCCCTT \\ \hline \textbf{CC} \hline \hline \textbf{CC} \\ \hline \textbf{CC} \hline \hline \textbf{CC} \\ \hline \textbf{CC} \hline \hline CC$
LS1 LS2 LS3 EF600696 FJ914764 K02120 AF257515	U5  → tRNA <sup>Pro</sup> Bind site 600 CGGCGCCCTCTAGCGGCAGGAGAGACCGGCAAACAAT GGGGGGCTCGTCCGGGAT TGATCACCCCGGAACCCTAACAACTCTCTGGACCCACCCCCTC 
LS1 LS2 LS3 EF600696 FJ914764 K02120 AF257515 Figure 1 Alig	GGCGGCATTTTGGGTCTCCCTTCAAATTATATC GT

**Figure 1 Alignment of S'LTR nucleotide sequences of BLV strains.** BLV strains isolated from PBMC cells and previously described are shown by accession number on the left side of the figure. BLV isolates from B-lymphosarcoma tumors (LS1 through LS3) are shown by name. Identity with BLV strain LS1 is indicated by a dot. The U3, R, and U5 regions are indicated on top of the alignment. The three TxRE enhancer regions are shown in green, cyclic AMP- responsive element (CRE) sequences are underlined and E-Box sequences are shown in italics. Binding sites for PU.1/Spi-B are shown in bold. The glucocorticoid responsive element (GRE) binding site is shown in bold, italics and underlined. Nuclear factor kB (NF-kB) binding sites are shown double underlined. The CAT box and GATAA box promoters (PROMT) sequences are indicated in yellow and in bold italics, respectively. The polyadenylation site (PAS) is shown in magenta and the CAP site is shown in light blue. The tRNA proline primer binding sites are shown in red. The downstream activator sequence (DAS) and the interferon regulatory factor (IRF) binding sites are shown in dark and light grey, respectively.



LS2 LS3	GASIPFKLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPV	RKPNGAWRFVHDLRATNALTKPI	PALSPGPPDLTAIPTHLPHIICLDLKDAFFQIPVED
1.53		· · · · · · · · · · · · · · · · · · ·	
EF600696			
ET914764	т.	⊥	
K02120			P
AF257515	.т.	T	R.PP
	101	150	20
LS1	RFRSYFAFTLPTPGGLQPHRRFAWRVLPQGFINSPALFERA	LQEPLRQVSAAFSQSLLVSYMDD	ILIASPTEEQRSQCYQALAARLRDLGFQVASEKTRQ
LS2	•••••••••••••••••••••••••••••••••••••••		
153 FF600606	α		
AF033818			
FJ914764			
K02120	F.LSS		YS.
AF257515			
	201	250	200
1.91	ZUI TOSDUDELCOMULEOTUTVOSLOTLOISSDISLUOLOAVLCI		KGTDDDRATIOLSDFOLOGTAFLROALSHNARSPVN
LS2			KGIDDEKAI IQHSEEQHQGIAEHKQAHSHMAKSKIN
LS3			
EF600696			
AF033818	N		
FJ914764			L
NUZIZU AF257515			т.
CTC1C2713			·····
	301	350	400
LS1	EQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQA:	SPWGLLLLLGCQYLQTQALSSYA	KPILKYYHNLPKASLDNWIQSSEDPRVQELLQLWPC
LS2			•••••••••••••••••••••••••••••••••••••••
LS3			
2EF600696			т
FJ914764			
K02120			T
AF257515			V. <b>T</b> FC
	401	450	50
LSI	ISSQGIQPPGPWKTLITRAEVFLTPQFSPDPIPAALCLFSD	GATGRGAYCLWKDHLLDFQAVPA	PESAQKGELAGLLAGLAAAPPEPVNIWVDSKYLYSL
LS3			К
EF600696			
AF033818	E		L
FJ914764	E		L
K02120			· · · · · · · · · · · · · · · · · · ·
AF25/515	V		·····
			endonuclease 600
	501	550	
LS1	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCQ
LS1 LS2	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCQ
LS1 LS2 LS2	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC
LS1 LS2 LS2 EF600696 AF033818	501 LRTLVLGAWLOPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCO
LS1 LS2 EF600696 AF033818 FJ914764	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S.V.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S
LS1 LS2 EF600696 AF033818 FJ914764 K02120	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVD <mark>QLLPLETPEQ</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S.V.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTL/LGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVD <mark>OLLPLETPEO</mark>	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC SVS. S. S. S. S. S. V.
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVDOLLPLETPEQ	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC SV SSSSS
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTL/VLGAWLOPDPVPSYALLYKSLLRHPAIFVGHVRSHSS	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTVS(GATUA 6 & KDCH PPPONUMT	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCO
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2 LS3	501 LRTL/LGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS2 LS3 EF600696	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 PVDTYSGATHASAKRGLTTQMTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S
LS1 LS2 LS2 EF600696 AF033818 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC SVS. SS. SS. CS. CS. C
LS1 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120	501 LRTL/LGAWLOPDPVPSYALLYKSLLRHPAIFVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTL/ULGAWLOPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLBTPEQ 650 FVDTYSGATHASAKRGLTTQWTI	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S.V.S.S.S.S.S.S.S.S.S.S.S.V. GELLEAIVHLGRPKKLNTDQGANYTSKTFVRFCQQF R.A.
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 F7033818 F7914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 F7914764 K02120 AF257515	501 LRTL/LGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 PVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 FF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 FF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS1 LS2	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S.V.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S
LS1 LS2 LS2 EF600696 FF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF257515 LS1 LS2 LS1 LS2 LS3 EF600696	501 LRTL/LGAWLOPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTT 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 LS2 LS2 LS1 LS1 LS1 LS2 LS3 LS3 LS3 LS3 LS3 LS3 LS3 LS3 LS3 LS3	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 PF003818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS2 LS2 LS2 LS3 EF600696 AF033818 FJ914764	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC
LS1 LS2 LS2 PF600696 FF038818 7J914764 (02120 AF257515 LS1 LS2 LS3 FF600696 AF033818 7J914764 (02120 LS1 LS2 LS3 EF600696 AF033818 R7914764 (02120	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 EF600696 FF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S. V. S.
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTL/LGAWLOPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 AF033818 AF033818 LS1 LS2 LS3 LS3 EF600696 AF033818 LS1 LS1	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 750 TQALSRALWTHNQINLLPILKTR 850 DGPEDAHNRSSDG	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S. V. S.
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 EF600696	501 LRTLVLGAWLOPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC 
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS2 LS3 EF600696 AF033818 FJ914764 FJ914764	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S
LS1 LS2 LS2 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 7J914764 K02120 AF257515 LS1 LS3 EF600696 AF033818 3P600696 AF033818 CJ914764 K02120 AF257515	501 LRTLVLGAWLQPDPVPSYALLYKSLLRHPAIPVGHVRSHSS 	550 ASHPIASLNNYVDQLLPLETPEQ 650 FVDTYSGATHASAKRGLTTQMTI 	WHKLTHCNPRALSRWPNPRISAWDPRSPATLCETCC S

I         Leader peptide        gp51         50         ND         100           LS1         MPKERRSRRROPI I RWVSLTLTLLALCKPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF           LS2	1     Leader peptide    gp51     50     ND 100       LS1     MKKEKKSRKRPOH I KWYSLTUTLALOR PUTWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRRF       LS2	$ $ $\rightarrow$ conformational epitopes F,G,H	conformational enitoned F G H	~		$ _{ ightarrow}$ conformat:	ional epitopes F,G,H		
LS2       LS3         LS3      K.         AF033818	LS2	▶	$ $ $\rightarrow$ contormational epicopes r, G, n	LS1	1 Leader peptide	►  →gp51 ALCEPIOTWECSLSLGNOO	50 WMTAYNOEAKESISIDOILEAHNOSPE	ND CAKSPRYTI.DSVNGYPKIYWPPP <mark>OGRI</mark>	100 RRF
LS3      K.         EF600696      K.         AF033818	LS3      K.         AF033818	1 Leader peptide  →gp51 50 ND 100 LS1 MPKERRSRRPOPTIRWSLTLTLALCEPTOTWCSLSLGNOOWMTAYNORAKESISIDOILEAHNOSPECAKSPRYTLDSVNGYPKIYWPPP <mark>OGRRPF</mark>	I Leader peptide  →gp51 50 ND 100 LS1 MPKRERSREPOPTIEWSLULTILALCEDITOTWCSLSLGNOOWMTAYNORAKESISIDOILEAHNOSPECAKSPRYTLDSVNGYPKIYWPPPOGREF	LS2		· · · · · · · · · · · · · · · · · · ·			
EF600696      K.         AF033818	EF600696      K.         AF033818	1     Leader peptide       →gp51     50     ND     100       LS1     MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	I Leader peptide  →gp51 50 ND 100 LS1 MPKERRSRRPQPIIRWVSLTLTLLALCRPIQTWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF LS2	LS3		<b>.</b>			
AF033818	AF033818	1     Leader peptide     →gp51     50     ND     100       LS1     MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	LS2 LS3	EF600696	K				
K02120	F0314764	1     Leader peptide      -gp51     50     ND     100       LS1     MPKERRSRRPQPIIRWVSLTLLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	I       Leader peptide       →       ND       100         LS1       MPKERRSRRPQPTIRWVSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	AF033818		Q	T	.PRF	• • •
AF257515      Q	AF257515	1         Leader peptide         ⊣gp51         50         ND         100           LS1         MPKERRSRRPOPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRRF           LS2	I       Leader peptide	FJ914764 K02120		Q.I		· · · · · · · · · · · · · · · · · · ·	• • •
Conformational epitopes F,G,H-          101       ND       150CD8*-T Cell epitope       epitope E       200         LS1       GARAMVTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHSIFTIAWSWGYDDITFSLKKTPDPPOPIFPQLNSDWVPSVRSWAL       LS2	Conformational epitopes F,G,H          101       ND       150CD8*-T Cell epitope       epitope E       200         LS1       GARAMV TYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFILHLKQCHGIFTIALWEIWGYDDI TFSLHK FDDFPDF FPQLNSDWVPSVRSWAL       S.       S.         LS2           S.         LS3           S.         LS3             LS4             LS2             LS3             EF600696             AF033818             FJ914764             K02120             AF257515             201       epitope B       epitope B'       250 epitopes D D'_       TMHR       epitope A       300	1         Leader peptide        gp51         50         ND 100           LS1         MPKERRSRRPOPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF           LS2	I       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRPQPI I RWVSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFS I SI DQ I LEAHNQSPFCAKSPRYTLDSVNGYPK I YWPPPQGRRRF         LS2	AF257515		QF		F	
Conformational epitopes F,G,H-          Interpretation       ND       150CD8'-T Cell epitope       epitope E       200         LS1       GARAMV TYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCH SI FULWEIWGYDG IN FRILKK FDDFPOPI FPQLNSDWVPSVRSWAL       LS2	Conformational epitopes F,G,H-          Interpretation       Interpretation         101       ND       150CD8'-T Cell epitope       epitope E       200         LS1       GARAMV TYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHSIFELAKKIPDFPOP       FPQLNSDWVPSVRSWAL         LS2       H.       S       S         LS3       D.       S       S         LS4       D.       S       S         LS5       D.       S       S         LS3       D.       D       S         LS4       D.       S       S         LS5       D.       S       S         LS4       D.       S       S         LS5       D.       S       S         LS6       S       S       S         LS1       D.       S       S         LS2       D.       S       S         LS3       S       S       S         S       D.       S       S         K02120       D.       D       S         AF257515       D.       D       S         201       epitope B       epitope B'       250 epitopes D D'       TMHR       <	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	I       Leader peptide       -gp51       50       ND 100         LS1       MPKERRSRRPQPIIRWSLTLTLLALCRPIQ       TRRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2						
ND       150CD8'-T Cell epitope       epitope E       200         LS1       GARAMV TYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCH SI FELTWEIWGYDD ITFSLHKI PDPPOPT FPQLNSDWVPSVRSWAL       LS2	ND         150CD8'-T Cell epitope         epitope E         200           LS1         GARAMV TYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCH SI FELTWEIWGYDD, ITFSLHKI PDPPOPT FPQLNSDWVPSVRSWAL	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	I       Leader peptide       Implicit epitopes r, s, n         1       Leader peptide       Implicit epitopes r, s, n         LS3       Implicit epitopes r, s, n       ND 100         LS3       Implicit epitopes r, s, n       Implicit epitopes r, s, n         FF600696            FJ914764            K02120            AF257515            LS3 </th <th></th> <th></th> <th>Conformatio</th> <th>hal epitopes F,G,H<math>_{\leftarrow} </math></th> <th></th> <th></th>			Conformatio	hal epitopes F,G,H $_{\leftarrow} $		
LS1       GARAMVTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHSTFELHWEIWGYDEUTFSLHKTPDPPOPDFPQLNSDWVPSVRSWAL         LS2	LS1       GARAMV       GARAMV       TYDCEPRCPYVGADRFDCPHWDNA       SQANQGSFYVNHQILFLHLKQCH       IFPLTWEIWCYDE/ITFSLHKTPDPPOPD       FPQLNSDWVPSVRSWAL         LS2       H       S       S       S       S         LS3       H       S       S       S         EF600696       D       S       S         AF033818       H       D       S         FJ914764       D       D       G         K02120       D       D       S         AF257515       D       D       D         201       epitope B       epitope B'       250 epitopes D D'       TMHR       epitope A       300	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIERWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRRF         LS2	I       Leader peptide       I		101	ND	150CD8 <sup>+</sup> -T Cell epitop	e epitope E	200
LS2	LS2	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIERWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF         LS2	I       Leader peptide        gp51       50       ND 100         LS1       MPKERRSRRPQPI IRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFS IS IDQI LEAHNQSPFCAKSPRYTLDSVNGYPK IYWPPPQGRRRF         LS2	LS1	GARAMV TYDCEPRCPYVGADRFDCPH	HWDNA <mark>SQANQGSFYVNHQI</mark>	LFLHLKQCHGIFTLTWEIW <u>GYDP</u> LITE	SLHKIPDPPQPDFPQLNSDWVPSVRSV	WAL
LS3	LS3	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIRWSLTLTLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF         LS2	I       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRPQPTIRWSULTLTLIALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRRF         LS2	LS2		· · · · · · · · · · · · · · · · · · ·	.s	•••••	
EF600696	EF600696	1         Leader peptide        gp51         50         ND 100           LS1         MPKKERSRRPOPTIRWSLTLTLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRFF           LS2	I       Leader peptide        gp51       50       ND 100         LS1       MPKERRSRRPQPI I RWVSLTLTLLALCRPIQ TWRCSLSLGNQQWMTAYNQEAKFS I SI DQ I LEAHNQSPFCAKSPRYTLDSVNGYPK I YWPPPQGRRRF         LS2	LS3	H	· · · · · · · · · · · · · · · · · · ·	.S		
AF033818	AF033818	1         Leader peptide        gp51         50         ND 100           LS1         MPKKERSRRPOPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRFF           LS2	I       Leader peptide       I -gp51       50       ND 100         LS1       MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ       TRRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	EF600696		D			• • •
K02120       D         AF257515       D         201       epitope B       epitope B'       250 epitopes D D'       TMHR       epitope A       300         LS1       LLNQTARAF       LLVYNKTISSSGP       GLALPDAQIFWV       MTSSFNTTQGWHHPSQR       LLPPISLVNLSTASSAPPT       RVR	K02120	1       Leader peptide        -gp51       50       ND 100         LS1       MPKKERSRRPOPTIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRF         LS2	I       Leader peptide        gp51       50       ND 100         LS1       MPKERRSRRPQPIIRWVSUTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	AF033818	н	ש	с		
AF257515	AF257515	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPQP11RWSLTLTLLALCRP10TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTDSVNGYPKIYWPPPQGRRFF         LS2	I         Leader peptide         -gp51         50         ND 100           LS1         MPKERRSRRPOPI IRWVSLTUTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF           LS3	K02120		שש	· · · · · · · · · · · · · · · · · · ·		•••
201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300 LS1 LLNQTARAF <mark>PDC<u>AICWEPSPPWAPE</u>ILVYNKTISSSGP</mark> GLALPDAQIFWV <u>NTSSFNTTQGWHHPSQR<mark>LLFNVSQGNALLLPPISLVNLSTASSAPPT</mark>RVR</u>	201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIEWVSLTLTLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	I       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRRPQP I IRWVSLTLTLLALCRP IQ       TWRCSLSLGNQQWMTAYNQEAKFS IS IDQILEAHNQSPFCAKSPRYTLDSVNGYPKI YWPPPQGRRRF         LS2	1000100		<b>D</b> D			
201         epitope B         epitope B'         250 epitopes D D'         TMHR         epitope A         300           LS1         LLNQTARAF         PDCAICWEPSPPWAPE         ILVYNKTISSSGP         GLADAQIFWV         MTSSFNTTQGWHHPSQR         LLPPTSLVNLSTASSAPPT         RVR	201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300	1       Leader peptide      gp51       50       ND 100         LS1       MPKERKSRRRPQPIIRWVSLTLTLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	I       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRPQDITRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRRP         LS3	AF257515					
LS1 LLNQTARAFPDC <u>AICWEPSPPWAPE</u> ILVYNKTISSSGPGLALPDAQIFWV <u>NTSSFNTTQGWHHPSQRLLFNVSQGNALLLPDISLVNLSTASSAPPT</u> RVR		1     Leader peptide      -gp51     50     ND 100       LS1     MPKERRSRRPOPI IRWSSITILLALCRIG TWRCSLSLGNQQWMTAYNQEAKFSISIDQI LEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRFF       LS3	I       Leader peptide       I -gp51       50       ND 100         LS1       MPKERRSPRRPOPTIRWSLITITLALCRPTOTWRCSLSLEGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	AF257515		epitope B'	250 epitopes D D'	TMHR epitope A	300
	LS1 LLNQTARAFPDC <u>AICWEPSPPWAPE</u> ILVYNKTI <b>SSSG</b> PGLALPDAQIFWU <u>NTSSFNTTQGWHHPSQRLLF</u> NVSQGNALLLP <u>PISLVNLSTASSAPPT</u> RVR	1     Leader peptide      -gp51     50     ND 100       LS1     MPKERRSRRPOPITRWSLTTTLAACCP1QTWRCSLSGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPYTLDSVNGYPKIYWPPQCRRRF       LS2	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERKSKRPQ91 IKWSLTLTLALCCP107 TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRRF         LS2           LS3           EF600696        Q.          AF033818        Q.          K02120        Q.          AF257515        Q.          Conformational epitopes F,G,H         F.         101       ND       150CD8'-T Cell epitope epitope E       200         LS1       GARAWYTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCH       S.          LS2          S.         LS3             LS1       GARAWYTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCH       TITINE INFORMATION FOR POLY PSVRSWALL          LS2              LS3              LS3              LS3	AF257515	201 epitope B		AQIFWV <u>NTSSFNTTQGWHHPSQR<mark>LLF</mark>N</u>	IVSQGNALLLP <u>PISLVNLSTASS</u> APPTI	RVR
LISZ	LS2	1     Leader peptide      -gp51     50     ND 100       LS1     MPKERRSRRPQP1IRWSLTLTLLACRPIQ     NRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQCRRFF       LS2	1       Leader peptide        -qp51       50       ND 100         LS1       MPKERRSRRRPQP1 IRWSLTLTLLALCRPIQ       TRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	AF257515 LS1	201 epitope B LLNQTARAF <mark>PDC<u>AICWEPSPPWAPE</u></mark>	<i>ILVYNKTI<b>SSSG</b>P</i> GLALPD			
EF600696	1.5.5	1     Leader peptide      -qp51     50     ND 100       LS1     MPKERRSRRPQPI IRWSLTLTLLALCRPIG     mRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQCRRF       LS2	I         Leader         peptide         peptide         peptide         peptide         peptide         peptide         ND 100           LS1         MPKERRSRRRPOT IRWUSLTULLALCEPIG         TWRCSLSLENQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF           LS2	AF257515 LS1 LS2 LS3	201 epitope B LLNQTARAF <mark>PDC<u>AICWEPSPPWAPB</u> </mark>	<u>ILVYNKTI<b>SSSG</b>P</u> GLALPD			• • •
AF033818V	دی EF600696	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRRPQFI IRWVSLTLTLLALCPIG TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRF         LS3	I         Leader peptide          gp51         50         ND 100           LS1         MPKERRSRRRPOPI IRWSSITTLALCEPIO         TWRCSLSLEGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPOGRRFF           LS2	AF257515 LS1 LS2 LS3 EF600696	201 epitope B LLN <u>O</u> TARAF <mark>PDC<u>AICWEPSPPWAPE</u></mark>	<u>ILVYNKTI<b>SSSG</b>P</u> GLALPD			 
FJ914764I	L53 EF600696 AF033818	1         Leader peptide          -qp51         50         ND         100           LS1         MPKERRSRRPQFI RWSSITTLALCEPIC         TWRCSLSLEGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPOGERRF           LS2	1       Leader peptide        gp51       50       ND 100         LS1       MKERRSRRRPDI I RWSLITTLALCEPIC TWRCSLSGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKI WPPPOGRRRF         LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818	201 epitope B LLN <u>Q</u> TARAF <mark>PDC<u>AICWEPSPPWAPE</u></mark>	<i>ILVYNKTI<b>SSSG</b>P</i> GLALPD			· · · · · · · · · · ·
K02120S	LS3 EF600696 AF033818 FJ914764	1         Leader peptide          -qp51         50         ND 100           LS1         MPKERRSRREPOI I RWSULTILALCEPIC         TWRCSLSLGNQQWMTAYNQEAKES IS IDQILEAHNQSPECAKSPRYTLDSVNGYPKI YWPPPQGRREF           LS3	1       Leader peptide        -gp51       50       ND 100         LS1       MpKERRSRPDD1 IRWSLATTILALCEPTOTWRCSLSGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS3	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764	201 epitope B LLNQTARAF <mark>PDC<u>AICWEPSPPWAPE</u></mark>	<i>ILVYNKTI<b>SSSG</b>P</i> GLALPD. G.	L	V	  K
301  →gp30 Fusion peptide 350 GD21 400	LS3         EF600696         AF033818	1         Leader peptide         -gp51         50         ND 100           L51         MEKERBRRE PQD I RAVSLALLAR PIQ TWRCSLSLGNQQMMTAYNQEAKFS IS TDQI LEAHNQSPFCAKSPRYTLDSVNGYPK IYWPPPQGRREF           L52	1       Leader peptide      gp51       50       ND 100         Ls1       MPKERSRPPOPI IRVSITULALURPITETWRCSLSLGNQQMMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYMPPPOGRRFF         Ls2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAF <mark>PDCAICWEPSPPWAPE</mark>	<i>ILVYNKTI<b>SSSG</b>P</i> GLALPD. G.	LI	VVV	   
LSI KSPVAALTLGLALSVGLTGINVAVSALSHORLTSLIHVLEODOORLITAINOTHYNLLNVASVVAONRRGLDWLYIRLGFOSLCPTINEPCCFLRIONDS	LS3	1         Leader peptide          -gp51         50         ND 100           L51         MPKERSERRECOTINGUESTICALLEPTETWCSLSLGNQQWMTAYNQEAKESISIDQILEAHNQSPCAKSPRUTLDSUNGYPKIYWPPPCERREF           L52	I         Leader peptide        gp51         50         ND 100           LS1         MMXERRARKPOPTIKWWSIATTLELACEPTOTWCSLSLONQOWNTATNYQEAKPSISTDQILEAHNQSPFCAKSPRTTLDSVNGYPKIYWPPPOGRRRF         ND 100           LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDCAICWEPSPPWAPE: 	ILVYNKTI <b>SSSGP</b> GLALPD. .G	L	GD21 DULYIRLGFOSLCPTINEPCCFLRIO	      400 <b>NDS</b>
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAI<mark>NQTHYNLLNVASVVAQNRRGLD</mark></u> WLYIRLGFQSLCPTINEP <i>CCFLRIQNDS</i> LS2	LS3	1         Leader         peptide          -qp51         50         ND 100           MPKERERPOPT         NN 100         MPKERERPOPT         NN 100         NN 100           LS3	I         Leader peptide        gp1         50         ND 100           LS1         MEXERBARROPTI RWSUITTLALCEPT TWRCSLSLONQOWNTATNOBAKESTSTB01LEAHNOSPFCAKSPRTTLDSVNGYPRITYPPPOGRAFIE         1         0         ND 100           LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2	201 epitope B LLNQTARAFPDCAICWEPSPPWAPE 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G	L	GD21 	•••• ••• ••• ••• ••• ••• ••• ••• ••• •
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGFQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3	LS3	1         Leader         peptide          -qp51         50         ND 100           L51         MMCENERERPOPTIENVEDTITLELACEPTOTIVESLELENQOMMETATIVERATESTSTEDULEAHNOSPECARSPERTITLESUNGPERTURDIVEDUPOLOGRAFIE           L53	1         Leader peptide        gp51         50         ND 100           LS1         MMXERSHERFUG I HWVSLTUTLALCHUP THRCSLSLANQOWNTATWQEARFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYFKIYWPPPQCHREF         LS3           LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2 LS3	201 epitope B LLNQTARAF <mark>PDC<i>AICWEPSPPWAPE</i> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2</mark>	ILVYNKTI <b>SSSGP</b> GLALPD. .G	L	GD21 WLYIRLGFQSLCPTINEP <u>CCFLRIQ</u>	400 NDS
LS1 RSPVAALTLGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGFQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 EF600696	LS3	1         Leader peptide        gp13         50         ND 100           LS1         MEXEMBERGEDU LEWINGTPRICESLELONQUMMTATINGEREFSISTDILEAHNOSPPCAKSPRYTLDSWORPKHYMPPPCGEREF           LS2	I         Leader peptide        gp51         50         ND 100           LS1         MXXERSBRERVp1 IMVVSITULALCHUT TWRCSLSLANQQMMTATVQSARSISIDDILEAHNQSPFCAKSPRYTDSVNGYPKI YMPPPQCRRRF           LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696	201 epitope B LLNQTARAF <mark>PDC<i>AICWEPSPPWAPE</i></mark> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G	L	GD21 	400 NDS
LS1 RSPVAALTLGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINQTHYNLLNVASVVAQNRRGLD</u> WLYIRLGFQSLCPTINEPCCFLRIQNDS LS2 LS3 EF600696 AF033818 FI914764	LS3	1         Leader peptide        gp31         50         ND 100           LS1         MMEMBERSREPUT INVSITUTLALER TWRESCISION OWNTATING BARFSISTIDULEANNOSPECAKSPRYTLDSVNOTPKIYMPPD GERRE           LS2	I Leader peptide        gp51         50         ND 100           LS1         Impose intervent Traces Lennogement Stress Lennogement St	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818	201 epitope B LLNQTARAF <mark>PDC<i>AICWEPSPPWAPE</i></mark> 301  -g <b>p30 Fusion peptide</b> RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G.	L	GD21 	400 NDS
LS1 RSPVAALTEGLALSVGETGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <u>CCPLRIQN</u> DS LS2 EF600696 AF033818 FJ914764 K02120	LS3	1         Leader peptide          -gp31         50         ND 100           LS1         MEXEMPSREBUG INVOLVENTIALIANCE TARGESISTED (LEANNOSPECAKSPRYTLDSVNOVPENTIVPEP OCERNE           LS2	I Leader peptide        gp51         50         ND 100           LS1         IMMARKSERHOPTIKWSLINILALOUT UNCSLSLENQOMMTAYNQEAKFSISIDQILEAHNOSPECAKSPRYTLDSVNGYPKIYWPPPCARAHE           LS2	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120	201 epitope B LLNQTARAF <mark>PDC<i>AICWEPSPPWAPE</i></mark> 301  -g <b>p30 Fusion peptide</b> RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G	L	GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ	400 NDS
LS1 RSPVAALTEGLALSVGETGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <u>CCPLRIQN</u> DS LS2 EF600696 AF033818 FJ914764 K02120 AF257515	LS3	1         Leader         peptide         j=gp51         50         ND 100           LS1         MARRESHENPOL HAVSLICHLANCENETRIKESISIONQUMPTAYNQEAKPSISIDULEAHNQSPFCAKSPRYTLDSVNYKYTVPPPQCARRE           LS3	I         Leader peptide        gp51         50         ND 100           LS1         Image: Im	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G.	L	GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ	400 NDS 
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <u>CCPLRIQN</u> DS LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	LS3	1         Leader         peptide         j=gp51         50         ND 100           LS3	I Leader peptide         i-gp11         50         ND 100           LS1         Internament-opinite         i-gp11         50         ND 100           LS2         Internament-opinite         i-gp11         50         ND 100           LS3         Internament-opinite         i-gp11         50         ND 100           LS3         Internament-opinite         i-gp11         ND 100           LS3         Internament-opinite         i-gp11         ND 100           LS3         Internament-opinite         i-gp11         Internament-opinite           LS3         Internament-opinite         internament-opinite         PR           LS3         Internament-opinite         Internament-opinite         Provide           LS3         Internament-opinite         Internament-opinite         Provide           LS3         Internament-opinite         Internament-opinite         Provide           LS3         Internament-opinite         S         Provide         Provide           LS3         Internament-opinite         S         Provide         Provide         Provide           LS3         Internament-opinite         S         Provide         Provide         Provide         Provide         Provide	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2	ILVYNKTI <b>SSSGP</b> GLALPD. .G.	L	GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ	
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCPLRIQNDS</i> LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500	LS3	1         Leader peptide          -gp51         50         ND 100           LS1         MFERERSHARD'E LANGENTILLALGER TORCESLENNOQWITA YNDERAFFSISIDQILEARNOSPECARSPRITTLOSINGYPERINGPECORRER           LS2	Ising         Ising <thising< th="">         Ising         <thi< td=""><td>AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515</td><td>201 epitope B LLNQTARAFPDC<u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM</td><td><u>ILVYNKTISSSGP</u>GLALPD. </td><td>L. L. S. L. A. I. 350 RLITAI<mark>NOTHYNLLNVASVVAQNRRGI</mark></td><td>GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ</td><td>400 NDS   </td></thi<></thising<>	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM	<u>ILVYNKTISSSGP</u> GLALPD. 	L. L. S. L. A. I. 350 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark>	GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ	400 NDS   
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 LS3 FJ914764 K02120 AF033818 HJ914764 K02120 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 IIRLGDLQPLSQRVSTDWQWPWNWDLGLTAWVRETHSVLSLFLLALFLLFLAPCLIKCLTSRLLKLLRQAPHFPEISLTPKPDSDYQALLPSAPEIYSH	Lb3	1         Leader peptide         ]-gp51         50         ND 100           LS1         MFERRIKANDO'E HEWSENTTLALADERET TWECSLALENDOWNTAVDRAKEFS IS IDQUILEANDOSPECKSPRTTLOSVNSTPETUSPPOCERNE           LS2	I         Leader         periode         (-gp51         50         ND 100           LS3	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2	201 epitope B LLNQTARAF PDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	<u>ILVYNKTISSSGP</u> GLALPD. 	L. L	GD21 DWLYIRLGFQSLCPTINEPCCFLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPEI	
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 IIRLGDLQPLSQRVSTDWQWPWNWLGLTAWVRETIHSVLSLFLLALFLLFLAPCLIKCLTSRLLKLLRQAPHFPETSLTPKPDSDYQALLPSAPEIYSH LS2 LS3	LS3	1         Leder peptide         ]-gp51         50         ND 100           LS1         Markinskingel Haveslatter LALGEBET THE CLASS LEN QOMMTA YURGER FEBIS DOLLEARING SPECKESPERTLOS WATTER PORTER         LS2           LS3	I         Leader peptide        gpil         50         ND 100           LS1         I         Leader peptide        gpil         50         ND 100           LS2         I         Internet and the period of the per	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3	201 epitope B LLNQTARAF PDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	ILVYNKTISSSGPGLALPD. 	L	GD21 DWLYIRLGFQSLCPTINEPCCPLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPETY	4000 NDS       5000 YSH
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 LS3 LS3 FJ914764 K02120 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 IIRLGDLQPLSQRVSTDWQWPWNWDLGLTAWVRETIHSVLSLFLLALFLLFLAPCLIKCLTSRLKKLRQAPHFPEISLTPKPDSDYQALLPSAPEIYSH LS2 LS3 EF600696	LS3	1         Leader peptide        gp51         50         ND 100           LS3         Michaeleskikkeyeli Medistriktikater Big TrikesLsLangengenkesperstingengenster tigter Strategie Medistriktikater Big TrikesLsLangengen Big TrikesLsLangengen Big TrikesLsLangengen Big TrikesLsLangengengen Big TrikesLangengengengen Big TrikesLangengengengengengengenster Big TrikesLangengengengengengengenster Big TrikesLangengengengengengengengengengengengengen	I         Leader poptide        gpil         50         ND 100           LS1         INCREMENTED INVOLUTIALOUS (INCREMENTATINGERKESISIDQUEENINGSPECKSPRYLDSVNOPPLINPPFORMER (SIS)         INCREMENTATION (INCREMENTATION (INCREMENT)))))))           ISI         1100000000000000000000000000000000000	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 ES1 LS2 LS3 EF600696	201 epitope B LLNQTARAF <mark>PDCAICWEPSPPWAPE</mark> 301  -g <b>g30</b> Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	ILVYNKTISSSGP G. ALSH <u>ORLTSLIHVLEQDQQ</u> Membran LGLTAWVRETIH <i>SVLSLFL</i>	L. L. S. L.A. J50 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark> e span450  →cytoplasmic LALFLLFLAPCLIKCLTSRLLKLLRQA	GD21 DWLYIRLGFOSLCPTINEPCCFLRIQ 	
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 LS3 FJ914764 K02120 AF033818 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 IIRLGDLQPLSQRVSTDWQWPWNWDLGLTAWVRETIHSVLSLFLLALFLLFLAPCLIKCLTSRLKKLRQAPHFPEISLTPKPDSDYQALLPSAPEIYSH LS2 LS3 EF600696 AF033818	Los       EF600696         AF033818	1         Leader peptide        gp1         50         ND 100           LS1         Mediakskikkeyel ikwisteri LACKE BUTWIKCSLSLUKOVQMMTANDEAKSTST IDDULEAHNOSPPCAKSPRTTDEVNGYPHIWPPCORKER           LS2	I         Lender peptide         Constructional epicopes f./g./m         ND 100           LS1         I         Lender peptide         Fig. 50         ND 100           LS2         LS3         ES3         ES4         <	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818	201 epitope B LLNQTARAF <mark>PDCAICWEPSPPWAPE</mark> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	ILVYNKTISSSGP G. 	L. L. S. L.A. 350 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark> e span450  →cytoplasmic LALFLLFLAPCLIKCLTSRLLKLLRQA	GD21 DWLYIRLGFOSLCPTINEPCCFLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPEIN	400 NDS 500 YSH
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>QRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <u>CCFLRIQN</u> SS LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDLGLTAWVRETIH <i>SVLSLFLLALFLLFLAPCLI</i> KCLTSRLKKLRQAPHFPEISLTPKPDSDYQALLPSAPEIYSH LS2 LS3 EF600696 AF033818 FJ914764	L53	1         Leader peptide        gp1         50         ND 100           LS3         Heinikaskukupi Haviantini ALGEB ("WRCSLSLANQOMFIANNORAKPSISTDDILEAHNOSPPCAKSPRTTLDSVNGTPPLIVPPCORKER")           LS3	I         Leader         peptia         50         ND 100           LS1         I         Leader         peptia         50         ND 100           LS2         Immediate Continue Nutrituitation Processing Procesprocessing Procesproces Processing Procesproces Proce	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764	201 epitope B LLNQTARAF <mark>PDCAICWEPSPPWAPE</mark> 301  -g <b>p30</b> Fusion peptide RSPVAALTLGLALSVGLTGINVAVSA 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	ILVYNKTISSSGP G. 	L. .L. .S. .L.A. .I. 350 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark> e span450  →cytoplasmic LALFLLFLAPCLIKCLTSRLIKLLRQA	GD21 DWLYIRLGFOSLCPTINEPCCPLRIQ S. 	400 400 NDS   500 YSH      
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQQRLITAINOTHYNLINVASVVAQNRRGLD</u> WLYIRLGPQSLCPTINEP <i>CCFLRIQNDS</i> LS2 LS3 EF600696 AF033818 FJ914764 AF257515 401 LHR TM Membrane span450  -cytoplasmic domain 500 LS1 IIRLGDLQPLSQRVSTDWQ WPWNWDLGLTAWVRETIHSVLSLFLLALFLLFLAPCLIKCLTSRLLKLLRQAPHFPEISLTPKPDSDYQALLPSAPEIYSH LS2 LS3 EF600696 AF033818 FJ914764 FP H914764 LR K02120 LR LR LR	LS3       EF600696	1         Leder         peptide        gp51         50         ND 100           L63         Michaesanavoi Lavasud TWRCSLSLCNQOWTANUGEAKSIS DOLLEAHNQSPECKSEPPTLDSWOPPKINPPLOGRAME           L63	I         Leader_peptide         Image: point         Image: point <thimage: point<="" th="">         Image: point</thimage:>	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAF <mark>PDCAICWEPSPPWAPE</mark> 301  gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	<u>ILVYNKTISSSGP</u> GLALPD. .G. .G. <u>Membran</u> LGLTAWVRETIH <i>SVLSLFL</i> .G.	L. L. S. L. A. I. 350 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark> e span450  cytoplasmic LALFLLFLAPCLIKCLTSRLLKLLRQA	GD21 	4000 NDS 5000 YSH R
LS1 RSPVAALTEGLALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDQORLITAINOTHYNLLNVASVVAQNRRGLD</u> NLYIRLGPQSLCPTINEP <u>CCPLRIQN</u> DS LS2 	LS3	1         Leder         peptide        gp51         50         ND 100           L83         MEREGRAMMANDELEMPERTITIENDERVERTINGERVERTSDOLLEHINGSPECKESPERTIDESWOPPERTURPER	1         Leader:         pp:1         S0         ND         <	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF03818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  -gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI	ILVYNKTISSSGPGLALPD. .G. .ALSHQRLTSLIHVLEQDQQ Membran LGLTAWVRETIHSVLSLFL .G. .G.	L. L. S. L. A. I. 350 RLITAI <mark>NOTHYNLLNVASVVAQNRRGI</mark> e span450  cytoplasmic LALFLLFLAPCLIKCLTSRLLKLLRQA .ES	GD21 	4000 NDS 5000 YSH R R
LS1       KSPVAALTIGLALSVGLTGINVAVSALSHQRLTSLIHVLEQDQQRLITATIMOTHYNLLNVASVVAQNRGLDWLYIRLGPOSLCPTINEPCCFLRIQNDS         LS2	L53	1         Leader pertide        up51         50         MD 100           LS3         Hernississie/UV is sevent to LAURESISTER/QUERTARY MUGAR/SISTED LLAUROSPECARSPRYTLDS/NUTY RY TYPE/PGGRMT           LS3	1         Leader peptide         50         ND         100           151         Hecksanstructure Lalexter         50         ND         100           162	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS2 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600596 AF033818 FJ914764 K02120 AF257515 LS1 LS1 LS1 LS1 LS1 LS1 LS1 LS1	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI 	ILVYNKTISSSGPGLALPD. 	L. 	GD21 	400 MDS  500 YSH      
LS1       RSPVAALTEGLALSVGLTGINVAVSALSHORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVAQNRRGLDWLYIRLGPQSLCPTINEPCCPLRIQNDS         LS2	LS3	List         Listeder peptide        up51         50         MD 100           L63         Hernamsnikkyl investrittalkostageverksiskage	1         Leader peptide         pp51         50         ND         100           LS1         Heinsmannen om internen interne	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS1 LS2	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTIGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI 	GG	L. L. L. S. L.A. I. 350 RLITAINOTHYNLLNVASVVAQNRRGI e span450  →cytoplasmic LALFLLFLAPCLIKCLTSRLLKLLRQA ES ES	GD21 BWLYIRLGFQSLCPTINEPCCFLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPEIN FP. FA.	400 NDS 500 YSH R
LS1       MSEVAALTIGIALSVGLTGINVAVSALSHORLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLDWLYIRLGFOSLCPTINEPCCFLRIONDS         LS2	LS3	1         Leader peptide        up51         50         MD         100           143         Markanssniku/J. Havatristiku/J. Havatri Havatri Havatristiku/J. Havatri Havatristiku/J. Havatristiku/J.	1         Leader peptide         93         50         ND	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  →gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS2 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI 	GG	L	GD21 DWLYIRLGFOSLCPTINEPCCFLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPET FP. FA.	400 400 NDS  500 YSH       
LS1       KSPVAALTEGLALSVGLTGINVAVSALSHQRLTSLIHVLEQDQQRLITAINOTHYNLLNVASVVAQNRRGLDWLVIRUGFQSLCFTINEPCCPLRTONDS         LS2	LS3	1         Leader peptide        gp51         50         MD 100           LS1         Minimum Semiduri Havart LLAISEN (Ministra Ministra Mini	1         Leader peptide         -gp1 5         50         ND 100           LS1         Imvested propriate interaction of the state	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  -gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVS7 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI 	ILVYNKTISSSGPGLALPD. 	L	GD21 DWLYIRLGFOSLCPTINEPCCPLRIQ S. domain PHFPEISLTPKPDSDYQALLPSAPETY FP. FA. A.	4000 MDS  5000 YSH       
LS1         KSPVAALTIGLALSVGLTGINVAVSALSHORLTSLIHVLEQDQQRLITAINOTHYNLLAVASVAANRRGLDVLYRLGFOSLCPTINEFCCTRTONDS           LS2	LS3	Lie         Justice         Justice         Justice         ND         100         ND         100           L81         Introduces repetide	1         Leader peptide         -951         50         ND         ND         ND           LS1         International control international conteneuron internatinterecon control international cont	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDC <u>AICWEPSPPWAPE</u> 301  gp30 Fusion peptide RSPVAALTLGLALSVGLTGINVAVSA 401 LHR TM <u>IIRLGDLQPLSQRVSTDWQ</u> WPWNWDI 	ILVYNKTISSSGPGLALPD. 	L	GD21 DWLYIRLGFOSLCPTINEPCCPLRIQ S	400 400 MDS  500 YSH       
LS1       KSPYAALTLELALSVGLTGINVAVSALSH <u>ORLTSLIHVLEQDOQRLITAINOTHYNLINVASVVAQNREGLDWLYIRLGFOSLCPTINEPCCPLRIONDS</u> LS2	LS3       EF600696	1         Leader peride        953         50         ND         100           122         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           122         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           123         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           124         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           123         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           124         Development Induced TorREGISTICMQUMPTAYNORARES ISTOCILLARNOOPPECKEPRTIT.DEVNKYPHTYPEPGGENET           125         Conformational epitopes F, G, R[           126         TorREGISTICMUSTICHTURE INDUCED TORREGISTICMUSTICHTURE INDUCED TORREGISTIC INDUCED T	1         Leader peptide         -gp3         50         KD         ND	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDCAICWEPSPPWAPE: 301 	G	L	GD21 	4000 4000 MDS       
LS1         KSPYAATTELELALSVGLTGINVAVSALSHORLTSLIHVLEQDQQRLITAINGTHYNLINVASVVAQNREGIDWLYIRLGEGSLCPTINEPCCLIRIONDS           LS2	LS3	1         Lace:         pp:14         50         MD         MD           12         Lace:         pp:14         50         MD         MD           12         Lace:         pp:14         F         MD	1         Leader peptide         -mp31         50         B0         100           L52         HDERMENDUM HAMMONT LANDEDT WICKSELCKOQUMERANDSAKER LST.DOTLEAHNOSPECAKER PRYTLOS WICKPETTWEPPODERER         F000566         F0000566         F000566         F0005	AF257515 LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF03818 FJ914764 K02120 AF257515 LS1 LS2 LS3 EF600696 AF03818 FJ914764 K02120 AF257515	201 epitope B LLNQTARAFPDCAICWEPSPPWAPE: 301 	G. <u>Membran</u> <u>G.</u> <u>Membran</u> <u>G.</u> <u>G.</u> <u>G.</u> <u>G.</u> <u>G.</u>	L	GD21 	400 400 MDS       
P21 PHILIPIAL CHEPSEAWASPITALINALISSES CHAPPED ALL CHEPSEAWASPITALINAL CHAPTER CHEPSEAWASPITALINAL CHEPSEAWASPITALINAL		1     Leader peptide     -gp51     50     ND 100       LS1     MPKERRSRRPOPI I RWSSTTTLALACRIG TWRCSLSGNQQWMTAYNQEAKFSISIDQI LEAHNQSPFCAKSPRYTLDSVNGPKIYWPPPQGRRRF       LS3	I       Leader peptide       I -gp51       50       ND 100         LS1       MPKERRSPRRPOPTIRWSLILTILLALCRPTOTWRCSLSLEMQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRRF         LS3	AF257515		epitope B'	250 epitopes D D'	TMHR epitope A	300
LS1 LLNQTARAF <mark>PDC<u>AICWEPSPPWAPE</u>ILVYNKTI<b>SSSG</b>PGLALPDAQIFWV<u>NTSSFNTTQGWHHPSQRLLFNVSQGNALLLPPISLVNLSTASSAPPT</u>RVR</mark>		1       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRPOPI IRWVSLTLTLALCRPTOTWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF         LS2	1       Leader peptide      gp51       50       ND 100         LS1       MPKERRSRRPOPI TRWSLTTTLLALCRPTO       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQCRRFF         LS2	7 F 2 F 7 F 1 F		enitope B'	250 enitones D D'	TMUP enitone A	300
201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300 LS1 LLNQTARAF <mark>PDC<u>AICWEPSPPWAPE</u>ILVYNKTISSSGP</mark> GLALPDAQIFWV <u>NTSSFNTTQGWHHPSQR<mark>LLFNVSQGNALLLPPISLVNLSTASSAPPT</mark>RVR</u>	201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300	1       Leader peptide        gp51       50       ND 100         LS1       MPKKERSRRPOPTIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPQGRRRF         LS2	I         Leader peptide        gp51         50         ND 100           LS1         MPKERRSRRPQP1IRWVSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRRF           LS2	K02120		<b>D</b> D			
K02120	K02120	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIERWSLTLTLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF         LS2	I       Leader peptide       -gp51       50       ND 100         LS1       MPKERRSRRPQPI I RWVSLTLTLIALCRPIQ TWRCSLSLGNQQWMTAYNQEAKFS I SI DQ I LEAHNQSPFCAKSPRYTLDSVNGYPK I YWPPPQGRRRF         LS2	FJ914764		DD	G		
FJ914764	FJ914764	1         Leader peptide        gp51         50         ND 100           LS1         MPKERRSRRPOPTIERWSLTLTLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF           LS2	I       Leader peptide       -gp51       50       ND 100         LS1       MPKERRSRRPQPTIRWSULTLTLIALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	AF033818		D			
AF033818	AF033818	1         Leader peptide          -gp51         50         ND 100           LS1         MPKERRSRRPOPTIERWSSLTLELALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPQGRRFF           LS2	I         Leader peptide         -gp51         50         ND 100           LS1         MPKERRSRRPQPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYLDSVNGYPKIYWPPPQGRRRF           LS2	EF600696	· · · · · · · · · · · · · · · · · · ·	<b>.</b>			•••
LS3	LS3	1         Leader peptide          -gp51         50         ND 100           LS1         MPKERRSRRPOPTIERWSSLTLTLAACRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF           LS2	I       Leader peptide       -gp51       50       ND 100         LS1       MPKERRSRRPQPI I RWSULTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFS IS IDQI LEAHNQSPFCAKSPRYTLDSVNGYPK I YWPPPQGRRRF         LS2	LS2	H	· · · · · · · · · · · · · · · · · · ·	.S		• • •
LS2HS. LS3 LS3 AF033818D. FJ914764D. K02120D. AF257515D. 201 epitope B epitope B' 250 epitopes D D' TMHR epitope A 300 LS1 LLNQTARAFPDC <u>AICWEPSPPWAPE</u> ILVYNKTISSSEPGLALPDAQIFWV <u>NTSSFNTTQGWHHPSQRLLPPISLVNLSTASS</u> APPTRVR	LS2	1       Leader peptide        -gp51       50       ND 100         LS1       MPKERRSRRPOPTIEWVSLTLTLALCEPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPQGRRFF         LS2	I       Leader peptide        gp51       50       ND 100         LS1       MPKERRSRRPQDTIRWSULTLILALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTDSVNGYPKIYWPPPQGRRRF         LS2	LS1	<mark>GARAMV</mark> TYDCEPRCPYVGADRFDCPH	HWDNA <mark>SQANQGSFYVNHQI</mark>	LFLHLKQCH <mark>GIFTLTWEIW<u>GYDP</u>L<i>ITF</i></mark>	<mark>SLHKIPDPPQPD</mark> FPQLN <mark>SDWVPSVRSV</mark>	WAL
LS1 GARAMV TYDCEPRCPYVGADRFDCPHWDNA <mark>SQANQGSFYVNHQILFLHL</mark> KQCHGIFTLTWEIWGYDE ITFSLHKIPDPPOH FPQLNSDWVPSVRSWAL LS2	LS1       GARAMUTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHGIFTLTWEIWGYDELITFSLHKIPDPPOPIFPQLNSDWVPSVRSWAL         LS2	1         Leader peptide          -gp51         50         ND 100           LS1         MPKERRSRRPOPTIERWSLTLTLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF           LS2	I       Leader peptide       I -gp51       50       ND 100         LS1       MPKERRSRRPQPI I RWSULTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFS IS IDQI LEAHNQSPFCAKSPRYTLDSVNGYPK I YWPPPQGRRRF         LS2		101	ND	150CD8 <sup>+</sup> -T Cell epitop	e epitope E	200
101       ND       150CD8*-T Cell epitope       epitope E       200         LS1       GARAMVTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHGIFTUTWEIWGXDELITESLHKIFDPPOPULFPQLNSDWVPSVRSWAL       S	101       ND       150 CD8*-T Cell epitope       epitope E       200         LS1       GARAMVTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHLKQCHGIFTITWEIWGXDELITESLHKIPDPPORFFPQLNSDWVPSVRSWAL       S	1       Leader peptide       →gp51       50       ND 100         LS1       MPKERRSRRPQPTIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2	I       Leader peptide       I -gp51       50       ND 100         LS1       MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF         LS2			Conformation	nal epitopes F,G,H $_{\leftarrow} $	4	
Conformational epitopes F,G,H-          101       ND       150CD8*-T Cell epitope       epitope E       200         LS1       GARAMVTDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQILFLHL LS2	Conformational epitopes F,G,H-          101       ND       150CD8*-T Cell epitope       epitope E       200         LS1       GARAMVTYDCEPRCPYVGADRFDCPHWDNASQANQGSFYVNHQLLFLHLKQCHGIFTLTWBIWGYDDLTFFSLHKIPDPPOPT       FPQLNSDWVPSVRSWAL         LS2	1         Leader peptide        gp51         50         ND 100           LS1         MPKERRSRRPOPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF           LS2	I       Leader peptide       -gp51       50       ND 100         LS1       MPKERRSRRPQPIIRWSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	AF257515		QF		F	
AF257515	AF257515	1         Leader peptide         ¬gp51         50         ND         100           LS1         MPKERRSRRPOPTIRWSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF           LS2	I       Leader peptide       →       ND 100         LS1       MPKERRSRRPQPIIRWSLTLTLLALCRPIQTWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	FJ914764		Q.T		F	• • •
FJ914764	FJ 914764	1     Leader peptide      -gp51     50     ND 100       LS1     MPKERRSRRPQPIIRWVSLTLTLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF       LS2	I       Leader peptide         ¬gp51       50       ND 100         LS1       MPKERRSRRPQPTIRWVSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	AF033818		Q	T	.PRF	
AF033818	AF033818	1     Leader peptide      -gp51     50     ND 100       LS1     MPKERRSRRPQPIIRWVSLTLTLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	LS2 LS3	EF600696	K				
EF600696      K.         AF033818	EF600696      K.         AF033818	1     Leader peptide     →gp51     50     ND     100       LS1     MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	I     Leader peptide           S0     ND 100       LS1     MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ     TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF       LS2	LS3		<b>.</b>			
LS3      K.         BEF600696      K.         AF033818	LS3      K.         AF033818	1         Leader peptide         →gp51         50         ND         100           LS1         MPKERRSRRPQPIIRWVSLTLTLLALCRPIQ         TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRF	I Leader peptide  -gp51 50 ND 100 LS1 MPKERRSRRPQPIIRWSLTLTLLALCRPIQTWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRFF	LS2		<u></u> . <b>.</b>			
LS2         LS3         EF600696        K.         AF033818        K.         FJ914764         K02120         AF257515        Conformational epitopes F,G,H          I01       ND         150CD8'-T Cell epitope epitope E         200         LS3         EF600696	LS2         LS3         EF600696         AF033818        K.         FJ914764         K02120         AF257515        L         Q.T.        L         K02120         AF257515        L         Q.T.        L         K02120         AF257515        L         Q.T.        L        L         AF257515        L         Q.T.        L	1 Leader peptide $\rightarrow qp51$ 50 ND 100	1 Leader peptide  →qp51 50 ND 100	LS1	MPKERRSRRRPQPIIRWVSLTLTLL	ALCRPIQTWRCSLSLGNQQ	WMTAYNQEAKFSISIDQILEAHNQSPF	CAKSPRYTLDSVNGYPKIYWPPP <mark>QGRI</mark>	<u>RRF</u>
LS1       MPKERRSRRRPOPI IRWVSLTLTLLALCRPIQ       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRP         LS2	LS1       MPKERRSRRRPOPI IRWVSLTLTLLALCRPIG       TWRCSLSLGNQQWMTAYNQEAKFSISIDQILEAHNQSPFCAKSPRYTLDSVNGYPKIYWPPPQGRRRF         LS2	•	→ conformational epicopes r,g,n		1 Leader peptide	–  →qp51	50	ND	100

**Figure 4 Alignment of amino acid sequences of Envelope (Env) proteins of BLV isolates.** Amino acid residues corresponding to gp51 (SU) and gp30 (TM) proteins are shown on the top of the alignment. Leader peptide of SU is shown in green. Structural strong turn motif GYPD is shown in bold, italics and double underlined. Conformational epitope region is indicated on top of the alignment. Linear epitopes are shown in italics and underlined [41,42]. Receptor-binding domain (RBD) [43] residues are delimited by two triangles. Second strong turn, SSSG, is shown in bold and italics. Amino acids involved in neutralization domains are shown in yellow. CD8 + -T epitope is shown in red. SU transmembrane hydrophobic region (TMHR) is shown in light blue. TM fusion peptide is shown in bold, italics and underlined. BLV leash and  $\alpha$ -helical region (LHR) [35] is shown in bold, italics and underlined. Epitope GD21 is shown in magenta. TM membrane-spanning region is shown in bold and italics. The cytoplasmic domain is indicated on top of the alignment. The rest is the same as in Figure 1.

(A) tax

. . . . . . . . .

LS1 LS2

LS3 EF600696 AF033818

LS1

LS2

LS2 EF600696

LS1

LS2

LS3

LS1

LS2 LS3

FJ914764 K02120

AF257515

AF033818

FJ914764

AF257515

EF600696

AF033818

FJ914764

EF600696 AF033818

FJ914764 K02120 AF257515

K02120 AF257515

K02120

X	
Zn finger	
MASVVGWGPHSLHACPALVLSNDVTIDAWCPLCGPHERLQFERIDTTLTCETHRINWTADGRPCGLNGTLFPRLHVSETRPQGPRRLWINCPL	PAVRAQP
•••••••••••••••••••••••••••••••••••••••	
······································	
157 Leucine-rich activation domain	
GPVSL <mark>E</mark> PFERSPFQPYQCQLPSASSDGCPIIGHGLLPWNNLVTHP <b>V</b> LGKVLILNHM <mark>ANFSLLPSFDTLLVDPLRLSVFAPDTRGAIRYLSTLL</mark>	<mark>TLC</mark> PATC
······································	
	••••
Q	· · · · V · ·
Q	
V O DN	
v	
261	
240 247 251 255 287	297
ILPLGEPFSPNVPICRFPRDSNEPPLSEFELPLIQTPGL <mark>SWSVPAIDLFLTGPPSPCDRLHVWSS</mark> PQALQRFLHDPTLTWSELV <b>A</b> SRKIRLD <mark>S</mark>	PLK <b>L</b> QLL
· · · · · · · · · · · · · · · · · · ·	
· · · · · · · · · · · · · · · · · · ·	
L	
PP	
L	
303	
ENEWLSRLF	
T	

## (B) Rex

LS1	MPKERRSRRRPQPIIRWRQNYFLSFKQVLLVGGPTLYMPARPWFCPMMSPSMPGAPSAGPMSDSNSKGSTPRSPARPTVSTGPPMDDLAASMERCSLDCM
LS2	
LS3	
EF600696	ĸ
AF033818	м рт. н
RT014764	
FJ914764	······P·····P·····P·····P·····P·····P····
K02120	
AF257515	
	NLS
LS1	<b>NLS</b> SPRPAPKGPDDSGSTAPFRPFALSPA <mark>RFH</mark> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT
LS1 LS2	<b>NLS</b> SPRPAPKGPDDSGSTAPFRPFALSPA <mark>RFH</mark> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT
LS1 LS2 LS3	<b>NLS</b> SPRPAPKGPDDSGSTAPFRPFALSPA <mark>RFH</mark> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT
LS1 LS2 LS3 EF600696	NLS SPRPAPKGPDDSGSTAPFRPFALSPA <mark>RFH</mark> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT
LS1 LS2 LS3 EF600696 AF033818	NLS SPRPAPKGPDDSGSTAPFRPFALSPA <b>RFH</b> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT 
LS1 LS2 LS3 EF600696 AF033818 FJ914764	NLS SPRPAPKGPDDSGSTAPFRPFALSPARPH FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT 
LS1 LS2 EF600696 AF033818 FJ914764 K02120	NLS SPRPAPKGPDDSGSTAPFRPFALSPA <mark>RFH</mark> FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT 
LS1 LS2 LS3 EF600696 AF033818 FJ914764 K02120 AF257515	NLS           SPRPAPKGPDDSGSTAPFRPFALSPARFH           FPPSSGPPSSPTNANCPRPLATVAPLSGTAFFPGTT

Figure 5 Alignment of amino acid sequences of Tax and Rex proteins of BLV isolates. In (A) the alignment of Tax sequences is shown. The region between amino acids 240 and 265, in which missense mutations influence the transactivation activity of the Tax protein [46] is shown in light blue. A putative zinc finger domain is shown in grey, a leucine-rich activation domain in yellow and sites of phosphorylation are indicated in red. Position 303, where a previously described substitution E303K gave rise to a replication-deficient virus [47] is shown in green. Positions where substitutions have been previously reported to have an effect on transactivation activity are shown in bold. In (B) alignment of Rex sequences is shown. Nuclear export signal (NES) is shown in yellow and the nuclear localization signal (NLS) is shown in green. The rest is the same as Figure 1.

NEC

(amino acids 30 to 53), a transactivating domain (amino acids 157 to 197) and two phosphorylation sites (amino acids 106 and 293) [44] (see Figure 5A). A series of BLV Tax mutants with strikingly more ability to stimulate BLV LTR-directed transcription in comparison with wild-type Tax have been previously described. All these mutants have substitutions between amino acid 240 and 265 [45]. Amino acid changes previously related to higher transcriptional activity as well as changes in the previously described phosphorylation sites were not observed in all the BLV isolates examined in the present study (see Figure 5A). Only one substitution was found in strain LS1 outside the leucine-rich activation domain (Figure 5A). Whether this amino acid substitution has an effect on BLV transcription is currently unknown.

Previous studies revealed that silencing is critical for tumor progression and distinct genetic and epigenetic mechanisms were identified for complete suppression of BLV Tax expression.

Conservation of sites involved in suppression of viral expression may be an important factor for the uncontrolled proliferation of BLV-infected tumor cells [5].

The Rex proteins of Deltaretroviruses act to facilitate the export of intron-containing viral RNA [48]. The Rex proteins shuttle between nucleus and cytoplasm using the nuclear localization signal (NLS) and nuclear export signal (NES) (see Figure 5B). No significant substitutions were found in Rex proteins of all BLV strains enrolled in this study.

G4 protein amino acid sequence includes an aminoterminal stretch of hydrophobic residues (amino acids 1 to 24) followed by potential proteolytic cleavage sites and an arginine-rich region (amino acids 58 to 72) located in the middle of the protein [13] (see Figure 6A). This latter region is required for the interaction of G4 with cellular protein farnesyl pyrophosphate synthetase (FPPS), (phosphorylation) [49]. The biological relevance of G4-FPPS interaction has been previously demonstrated in cellular transformation. Mutations in the arginine-rich  $\alpha$ -helix of G4 abrogate primary cell immortalization and induction of tumors in nude mice [49]. Therefore, disruption of the interaction between G4 and FPPS could interfere with the oncogenic process.

No amino acid substitutions were found in the argininerich  $\alpha$ -helix of G4 protein of the previously sequenced BLV isolates examined in this work (see Figure 6A). Nevertheless, an amino acid substitution (A29V) can be observed in G4 of all BLV LSI.

Interestingly, premature stop codons were observed in R3 of two of the three LS BLV isolates (Figure 6B). Previous studies on BLV infection using sheep provide insight on the molecular genetic and epigenetic modulation of viral expression [50]. These studies show that the deletion of the region that expands from the end of the env gene to the splice acceptor site of the tax/rex mRNA does not impair infectivity [21]. These sequences correspond to the third and second exons of R3 and G4, respectively, revealing that these sequences may not be essential for infectivity in vivo. Although previous studies have shown that deletions in R3/G4 interferes with the efficiency of BLV propagation and restricts pathogenesis [14,15,46,49], another study has shown that one out of 20 sheep infected with a R3/G4 mutant developed a lymphoma after 7.5 years of latency, suggesting that the deleted sequences may not be strictly required for



pathogenesis [51]. Further studies will be needed to address the biological significance of these findings in studies using the cow as a model for BLV infection.

In summary, although the genome of BLV is highly conserved in our isolates and in isolates from other sources previously described, variations can be observed in some genome regions. It is thought that silencing of viral expression is a multi-step process leading to the uncontrolled growth of a transformed B-cell clone and the onset of disease [5] and is critical for tumor progression and proliferation of BLV-infected tumor cells [5], as well as escaping recognition by the host immune response [4]. In that sense, the substitution found in the GRE site of the 5'LTR of all BLV strains isolated from the lymphosarcomas might contribute to these factors, since previous studies have shown that substitutions in GRE site significantly reduces basal LTR transcription activity [52] (see Figure 1). Moreover, all amino acid substitutions in Tax previously found to be related to stimulate high transcriptional activity of 5'LTR were not found in this study (see Figure 5A). Genetic and epigenetic mechanisms have been recently proposed for BLV suppression of viral gene expression [53]. The results of the present report, using full-length genome sequences, suggest that point mutations along the whole genome may also be needed to allow BLV provirus to achieve silencing.

## **Additional files**

Additional file 1: Primers used for amplification and sequencing of the BLV genome. List of primers used for amplification and sequencing of full-length genome sequences of the BLV strains.

Additional file 2: Strategy for amplification and sequencing of fulllength BLV genomes. A scheme shows the strategy for amplifying the BLV genome in two long PCRs and the strategy used for sequencing fulllength BLV genomes. A scheme of BLV genome is shown on top of the figure and the relative position of LTR and BLV proteins in the BLV genome can be seen by the bar undelying the genome scheme. BLV was amplified in two long PCR shown in yellow using appropriate primers shown bellow. The position of the primers used for sequencing is also shown on the bottom of the figure.

#### **Competing interests**

All authors declare that they have no competing interests.

### Authors' contributions

GM and JC conceived the study; GM and SF have made substantial contributions to the acquisition of the data and the analysis. GO, LT, GR, FC and SB made substantial contribution to the interpretation of the data. OP has made substantial contribution to interpretation of data and has been involved in critically revising the manuscript for important intellectual content and made important contributions to the interpretation and discussion of the results found in this work. JC and GM participated in the different analysis and wrote the paper. All authors read and approved the final manuscript.

### Acknowledgements

We acknowledge support from Programa de Desarrollo de Ciencias Básicas (PEDECIBA), Agencia Nacional de Investigación e Innovación (ANII) and Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay. We acknowledge anonymous reviewer's for important suggestions and contributions to improve the quality of this work.

#### Author details

<sup>1</sup>Unidad de Biofísica de Proteínas, Institut Pasteur de Montevideo, Mataojo 2020, Montevideo 11400, Uruguay. <sup>2</sup>Laboratorio de Virología Molecular, Centro de Investigaciones Nucleares, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay. <sup>3</sup>Departamento Básico de Medicina, Hospital de Clínicas, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay. <sup>4</sup>Departamento de Inmunobiología. Facultad de Medicina, Universidad de la República, Av. General Flores 2125, Montevideo 11800, Uruguay. <sup>5</sup>Laboratorio de Técnicas Nucleares, Facultad de Veterinaria, Universidad de la República, Lasplaces 1550, Montevideo, Uruguay.

Received: 30 August 2012 Accepted: 30 January 2013 Published: 18 March 2013

### References

- Burny A, Cleuter Y, Kettmann R, Mammerickx M, Marbaix G, Portetelle D, Van den Broeke A, Willems L, Thomas R: Bovine leukaemia: facts and hypothesis derived from the study of an infectious cancer. *Cancer Surv* 1987, 6:139–159.
- Kettmann R, Deschamps J, Cleuter Y, Couez D, Burny A, Marbaix G: Leukemogenesis by bovine leukemia virus: proviral DNA integration and lack of RNA expression of viral long terminal repeat and 3' proximate cellular sequences. Proc Natl Acad Sci U S A 1982, 79:2465–2469.
- Kettmann R, Cleuter Y, Mammerickx M, Meunier-Rotival M, Bernardi G, Burny A, Chantrenne H: Genomic integration of bovine leukemia provirus: comparison of persistent lymphocytosis with lymph node tumor from of enzootic. Proc Natl Acad Sci U S A 1980, 77:2577–2581.
- 4. Pierard V, Guiguen A, Colin L, Wijmeersch G, Vanhulle C, Van Driessche B, Dekoninck A, Blazkova J, Cardona C, Merimi M, Vierendeel V, Calomme C, Nguyen T, Nuttinck M, Twizere JC, Kettmann R, Portetelle D, Burny A, Hirsch I, Rohr O, Van Lint C: DNA cytosine methylation in the Bovine Leukemia Virus promoter is associated with latency in a lymphoma-derived B-cell line. J Biol Chem 2010, 285:19434–19449.
- Merimi M, Klener P, Szynal M, Cleuter Y, Bagnis C, Kerkhofs P, Burny A, Martiat P, Van den Broeke A: Complete suppression of viral gene expression is associated with the onset and progression of lymphoid malignancy: observations in Bovine Leukemia Virus-infected sheep. *Retrovirology* 2007, 4:e51.
- Lagarias DM, Radke K: Transcriptional activation of bovine leukemia virus in blood cells from experimentally infected, asymptomatic sheep with latent infections. J Virol 1989, 63:2099–2107.
- Radke K, Sigala T, Grossman D: Transcription of bovine leukemia virus in peripheral blood cells obtained during early infection in vivo. *Microb Pathog* 1992, 12:319–331.
- Gaudray G, Gachon F, Basbous J, Biard-Piechaczyk M, Devaux C, Mesnard JM: The complementary strand of the human T-Cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. J Virol 2002, 76:12813–12822.
- Sagata N, Yasunaga T, Ohishi K, Tsuzuku-Kawamura J, Onuma M, Ikawa Y: Comparison of the entire genomes of bovine leukemia virus and human T-cell leukemia virus and characterization of their unidentified open reading frames. *EMBO J* 1984, 3:3231–3237.
- Rice NP, Stephens RM, Couez D, Deschamps J, Kettmann R, Burny A, Gilden RV: The nucleotide sequence of the env gene and post-env region of bovine leukemia virus. *Virology* 1984, 138:82–93.
- Derse D: Trans-acting regulation of bovine leukemia virus mRNA processing. J Virol 1988, 62:1115–1119.
- Gillet N, Florins A, Boxus M, Burteau C, Nigro A, Vandermeers F, Balon H, Bouzar AB, Defoiche J, Burny A, Reichert M, Kettmann R, Willems L: Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in humans. *Retrovirology* 2007, 4:18–49.
- Alexandersen S, Carpenter S, Christensen J, Storgaard T, Viuff B, Wannemuehler Y, Belousov J, Roth JA: Identification of alternatively spliced mRNAs encoding potential new regulatory proteins in cattle infected with bovine leukemia virus. J Virol 1993, 67:39–52.
- Kerkhofs P, Heremans H, Burny A, Kettmann R, Willens L: In vitro and in vivo oncogenic potential of bovine leukemia virus G4 protein. J Virol 1998, 72:2554–2559.

- Willems LD, Kerkhofs P, Dequiedt F, Portetelle D, Mammerickx M, Burny A, Kettmann R: Attenuation of bovine leukemia virus by deletion of R3 and G4 open reading frames. *Proc Natl Acad Sci U S A* 1994, 91:11532–11536.
- Derse D: Bovine leukemia virus transcription is controlled by a virusencoded trans- acting factor and by cis-acting response elements. *J Virol* 1987, 61:2462–2471.
- Dube S, Abbott L, Dube DK, Dolcini G, Gutierrez S, Ceriani C, Juliarena M, Ferrer J, Perzova R, Poiesz BJ: The complete genomic sequence of an in vivo low replicating BLV strain. *Virol J* 2009, 6:120.
- 18. DNA Data Bank of Japan. [http://arsa.ddbj.nig.ac.jp]
- Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res 1994, 22:4673–4680.
- Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007, 24:1596–1599.
- Adam E, Kerkhofs P, Mammerickx M, Kettmann R, Burny A, Droogmans L, Willems L: Involvement of the cyclic AMP-responsive element binding protein in bovine leukemia virus expression in vivo. J Virol 1994, 68:5845–5853.
- Nguyen TL, de Walque S, Veithen E, Dekoninck A, Martinelli V, de Launoit Y, Burny A, Harrod R, Van Lint C: Transcriptional regulation of the bovine leukemia virus promoter by the cyclic AMP-response element modulator tau isoform. J Biol Chem 2007, 282:20854–20867.
- 23. Calomme C, Dekoninck A, Nizet S, Adam E, Nguyen T, Van Den Broeke A, Willems L, Kettmann R, Burny A, Van Lint C: Overlapping CRE and E Box motifs in the enhancer sequences of the Bovine Leukemia Virus 5' long terminal repeat are critical for basal and acetylation-dependent transcriptional activity of the viral promoter: implications for viral latency. J Virol 2004, 78:13848–13864.
- Dekoninck A, Calomme C, Nizet S, de Launoit Y, Burny A, Ghysdael J, Van Lint C: Identification and characterization of a PU.1/Spi-B binding site in the bovine leukemia virus long terminal repeat. Oncogene 2003, 22:2882–2896.
- Xiao J, Buehring GC: In vivo protein binding and functional analysis of cis-acting elements in the U3 region of the bovine leukemia virus long terminal repeat. J Virol 1998, 72:5994–6003.
- Kiss-Toth E, Unk I: A downstream regulatory element activates the bovine leukemia virus promoter. Biochem Biophys Res Commun 1994, 202:1553–1561.
- 27. Kiermer V, Van Lint C, Briclet D, Vanhulle C, Kettmann R, Verdin E, Burny A, Droogmans L: An interferon regulatory factor binding site in the U5 region of the bovine leukemia virus long terminal repeat stimulates Taxindependent gene expression. J Virol 1998, 72:5526–5534.
- 28. Niermann GL, Buehring GC: Hormone regulation of bovine leukemia virus via the long terminal repeat. *Virology* 1997, **239**:249–258.
- Wang H, Norris KM, Mansky LM: Involvement of the matrix and nucleocapsid domains of the bovine leukemia virus Gag polyprotein precursor in viral RNA packaging. J Virol 2003, 77:9431–9438.
- 30. Beasley BE, Hu WS: Cis-Acting elements important for retroviral RNA packaging specificity. J Virol 2002, 76:4950–4960.
- D'Souza V, Melamed J, Habib D, Pullen K, Wallace K, Summers MF: Identification of a high affinity nucleocapsid protein binding element within the Moloney murine leukemia virus Psi-RNA packaging signal: implications for genome recognition. J Mol Biol 2001, 314:217–232.
- Sperka T, Miklossy G, Tie Y, Bagossi P, Zahuczky G, Boross P, Matuz K, Harrison RW, Weber IT, Tozser J: Bovine leukemia virus protease: comparison with human T- lymphotropic virus and human immunodeficiency virus protease. J Gen Virol 2007, 88:2052–2063.
- Bagossi P, Sperka T, Feher A, Kadas J, Zahuczky G, Miklossy G, Boross P, Tozser J: Amino acid preferences for a critical substrate binding subsite of retroviral proteases in type 1 cleavage sites. J Virol 2005, 79:4213–4218.
- Wallin M, Ekstrom M, Garoff H: Isomerization of the intersubunit disulphide-bond in Env controls retrovirus fusion. EMBO J 2004, 23:54–65.
- Lamb D, Schuttelkopf AW, ven Aalten DMF, Brighty DW: Charge-surrounded pockets and electrostatic interactions with small ions modulate the activity of retroviral fusion proteins. *PLoS Pathog* 2011, 7:e1001268.
- Portetelle D, Gouez D, Bruck C, Kettmann M, Mammerickx M, van der Maaten M, Brasseur R, Burny A: Antigenic variants of bovine leukemia virus (BLV) are defined by amino acid substitutions in the NH2 part of the envelope glycoprotein gp51. *Virology* 1989, 169:27–33.

- Bruck C, Mathot S, Portetelle D, Berte C, Franssen JD, Herion P, Burny A: Monoclonal antibodies define independent antigenic regions of the bovine leukemia virus (BLV) envelope glycoprotein gp51. *Virology* 1982, 122:342–352.
- Gallaher WR, Ball JM, Garry RF, Martin-Amedee AM, Montelaro RC: A general model for the surface glycoproteins of HIV and other retroviruses. *AIDS Res Hum Retroviruses* 1995, 11:191–202.
- Ban J, Czene S, Altaner C, Callebaut I, Krchnak V, Merza M, Burny A, Kettmann R, Portetelle D: Mapping of sequential epitopes recognized by monoclonal antibodies on the bovine leukemia virus external glycoprotein expressed in Escherichia coli by means of antipeptide antibodies. J Virol 1992, 373:2457–2464.
- Moratorio G, Obal G, Dubra A, Correa A, Bianchi S, Buschiazzo A, Cristina J, Pritsch O: Phylogenetic analysis of bovine leukemia viruses isolated in South America reveals diversification in seven distinct genotypes. *Arch Virol* 2010, 155:481–489.
- Mamoun RZ, Morisson M, Rebeyrotte N, Busetta B, Couez D, Kettmann R, Hospital M, Guillemain B: Sequence variability of bovine leukemia virus env gene and its relevance to the structure and antigenicity of the glicoproteins. J Virol 1990, 64:4180–4188.
- Ban J, Czene S, Altaner C, Callebaut I, Krchnak V, Merza M, Burny A, Kettmann R, Portetelle D: Mapping of sequential epitopes recognized by monoclonal antibodies on the bovine leukaemia virus external glycoproteins expressed in Escherichia coli by means of antipeptide antibodies. J Gen Virol 1992, 73:2457–2461.
- 43. Gatot JS, Callebaut I, Van Lint C, Demonte D, Kerkhofs P, Portetelle D, Burny A, Willems L, Kettmann R: Bovine leukemia virus SU protein interacts with Zinc, and mutations within two interacting regions differently affect viral fusion and infectivity in vivo. J Virol 2002, 76:7956–7967.
- 44. Willems LD, Kettmann R, Burny A: The amino acid (157–197) peptide segment of bovine leukemia virus p34tax encompass a leucine-rich globally neutral activation domain. *Oncogene* 1991, 6:159–163.
- Tajima S, Aida Y: The region between amino acids 245 and 265 of the bovine leukemia virus (BLV) Tax protein restricts transactivation not only via the BLV enhancer but also via other retrovirus enhancers. J Virol 2000, 74:10939–10949.
- Powers MA, Grossman D, Kidd LC, Radke K: Episodic occurrence of antibodies against the bovine leukemia virus Rex protein during the course of infection in sheep. J Virol 1991, 65:4959–4965.
- Van den Breoeke A, Bagnis C, Ciesiolka M, Cleuter Y, Gelderblom H, Kerkhofs P, Griebel P, Mannoni P, Burny A: In vivo rescue of a silent tax-deficient bovine leukemia virus from a tumor-derived ovine B-cell line by recombination with a retrovirally transduced wild-type tax gene. J Virol 1999, 73:1054–1065.
- Choi EA, Hope TJ: Mutational analysis of bovine leukemia virus Rex: identification of a dominant-negative inhibitor. J Virol 2005, 79:7172–7181.
- Lefèbvre L, Vanderplasschen A, Ciminale V, Heremans H, Dangoisse O, Jauniaux JC, Toussaint JF, Zelnik V, Burny A, Kettmann R, Willems L: Oncoviral bovine leukemia virus G4 and human T-cell leukemia virus type 1 p13(II) accessory proteins interact with farnesyl pyrophosphate synthetase. J Virol 2002, 76:1400–1414.
- Hajj HE, Nasr R, Kfoury Y, Dassouki Z, Nasser R, Kchour G, Hermine O, de Thé H, Bazarbachi A: Animal models on HTLV-1 and related viruses: what did we learn? Front Microbiol 2012, 3:333.
- Florins A, Gillet N, Boxus M, Kerkhofs P, Kettmann R, Willens L: Even attenuated bovine leukemia virus proviruses can be pathogenic in sheep. J Virol 2007, 81:10195–10200.
- Merezak C, Pierreux C, Adam E, Lemaigre F, Rousseau GG, Calomme C, Van Lint C, Christophe D, Kerkhofs P, Burny A, Kettmann R, Willems L: Suboptimal enhancer sequences are required for efficient bovine leukemia virus propagation in vivo: implications for viral latency. J Virol 2001, 75:6977–6988.
- Colin L, Dekoninck A, Reichert M, Calao M, Merimi M, Van den Broeke A, Vierendeel V, Cleuter Y, Burny A, Rohr O, Van Lint C: Chromatin disruption in the promoter of Bovine Leukemia Virus during transcriptional activation. Nucleic Acids Res 2011, 39:9559–9573.

### doi:10.1186/1297-9716-44-19

Cite this article as: Moratorio *et al.*: A detailed molecular analysis of complete Bovine Leukemia Virus genomes isolated from B-cell lymphosarcomas. *Veterinary Research* 2013 44:19.