Arch Virol (2005) 150: 443–458
DOI 10.1007/s00705-004-0437-1

Archives of Virology
Printed in Austria

Genetic diversity and spread of *Bovine leukaemia virus* isolates in Argentine dairy cattle*

G. Monti[1], **R. Schrijver**[2], and **D. Beier**[3]

[1]Wageningen University and Research Centre, Quantitative Veterinary Epidemiology Group, Wageningen University, Wageningen, The Netherlands
[2]Institute for Animal Science and Health Lelystad, The Netherlands
[3]Federal Research Centre for Virus Diseases of Animals, Institute of Epidemiological Diagnostics, Wusterhausen, Germany

Received December 29, 2003; accepted October 5, 2004
Published online December 10, 2004

Summary. Effective tools for use in control programmes against bovine leukaemia virus (BLV) infections require insight into the relationship between the variant structure of the bovine leukaemia virus and the spatial-temporal interaction of isolates and hosts. Our study showed the presence of two types of BLV isolates – Australian and Argentine – in dairy herds from various parts of Central Argentina; these isolates were characterised by RFLP on PCR amplicons, and some of them were confirmed by sequencing. One genotype (Argentine) was present in all herds, and the Australian genotype was found in two herds.

Phylogenetic analysis indicated four clusters. The first cluster was composed of the Argentine isolates and one from Brazil; the second was composed of several isolates found in European countries and one from Brazil; the third cluster was composed of BLV isolates found in Japan and Germany; the fourth cluster included American and Australian isolates and those from other countries.

The comparison of a number of synonymous and non-synonymous nucleotide substitutions using various BLV genes revealed purifying selection, suggesting that molecular evolution occurred under some functional constraint.

*Nucleotide sequence data reported are available in the GenBank® databases under the following accession numbers: AY178820 (Isolate BLV-ARGSF2); AY178818 (Isolate BLV-ARGSF6); AY178817 (Isolate BLV-ARGSF7); AF485773 (Isolate BLV-ARGSF8); AF485775 (Isolate BLV-ARGSF11); AY178819 (Isolate BLV-ARGSF13) and AF485774 (Isolate BLV-ASF406).

Introduction

Bovine leukaemia virus (BLV) is an exogenous retrovirus that, together with human T-lymphotropic virus (HTLV) and simian T-lymphotropic virus (STLV), belongs to the deltavirus genus of the family *Retroviridae* [8]. Some of the cattle infected with BLV can suffer from lymphomas and/or a proliferation of B-lymphocytes-denominated persistent lymphocytosis, but the majority of BLV-infected cattle are healthy carriers of the virus [23].

The genetic structure of pathogen populations is important in many contexts (such as development of vaccines, resistance and treatments). However, the dynamics of virus strains are also of fundamental biological and epidemiological interest [15]. Moreover, an important goal in studying the genetic diversity of retroviruses is to gain a better understanding of the correlation between retrovirus variation and disease progression [22]. At present, four complete BLV genomes have been reported worldwide: Belgian isolate [28, 29], Japanese isolate [30], Australian isolate [9] and recently, Argentine isolate [11]. Furthermore, knowing whether geographical patterns of variation in a particular trait are adaptations that arose through natural Darwinian selection or purifying selection could give valuable information to understand polymorphism variation and to determine candidate sites of the genotypes for functional studies.

In this study our aim was to characterise BLV isolates and their distribution in dairy herds in Argentina. In addition, we studied whether selective advantages of different genotypes could be introduced by purifying selection.

Materials and methods

Study population

Blood, serum and milk samples were collected from 445 lactating Holstein cows from nine farms located in the central region of Argentina. These herds were selected based on the owners' willingness to collaborate in this study. All lactating cows in five of these herds were sampled, while randomly-selected samples were taken from the other four herds. We also used results from two herds in this study that also were included in a two-year longitudinal study.

Serological assays

Each cow was screened with the ELISA 108 (milk and serum kit), described in [17], which is designed to detect antibodies against gp51. After testing these samples with the PCR, these cows were tested with the Agar Gel Immunodiffusion Test (Fac. Cs. Vet, UNLP) and a commercial ELISA (SERELISA BLV Ab Bi Indirect, Symbiotics).

PCR assay

Oligonucleotide primers for PCR (MWG Biotech) were designed according to sequence data published elsewhere [30]. Primers corresponding to the *env* gene [29] were selected, and env_{5032} 5′-TCT GTG CCA AGT CTC CCA GAT A-3′, and env_{5099} 5′-CCC ACA AGG GCG GCG CCG GTT T-3′ were used as forward primers. The reverse primers were env_{5521r} 5′-GCG AGG CCG GGT CCA GAG CTG G-3′, env_{5608r} 5′-AAC AAC AAC CTC TGG

GAA GGG T-3′. The sets env_{5099} and env_{5521} had been established and described previously [24, 25].

DNA was obtained from frozen blood collected with EDTA and was extracted using the NucleoSpin Blood kit (Machery-Nagel). The first round of nested PCRs [12] was performed using env_{5032}/env_{5608r} as first primers; initial incubation of samples was at 72 °C for 2 min; denaturation took place at 94 °C for 2 min followed by 50 amplification cycles consisting of denaturation at 95 °C, 30 s, primer annealing at 58 °C, 30 s and extension at 72 °C for 1 min; final extension took place at 72 °C for 4 min. The second round of amplifications was conducted using the second pair of primers (env_{5099}/env_{5521r}). The second round was the same as the first round except that the primer annealing temperature was changed to 72 °C. A known positive and negative control DNA sample was included in each test run, and samples showing a band migrating at 444 base pairs (bp) were considered as positive.

Restricted Fragment Length Polymorphism (RFLP) protocol

Integrated BLV provirus present in the DNA of circulating leukocytes from infected animals was compared by restriction endonuclease digestion. Randomly-selected samples (121) from amplicons of cows that tested positive for PCR were analysed by RFLP using a series of triple enzymes (*Bam*HI, *Pvu*II and *Bcl*I; Roche diagnostics GmbH) [3]. The test was done twice on 22 samples to assess repeatability of the results. RFLP was performed using 2.5 µl of the PCR product that was digested using 0.25 µl of each restriction enzyme, 1 µl of buffer and 7 µl of distilled water. This mix was incubated for 2.5 h at 37 °C for *Bam*HI and *Pvu*II and at 52 °C for *Bcl*I.

Nucleotide sequencing and analysis

In addition to RFLP analysis, 13 randomly-selected samples (of the initial 121 samples) that were characterised by RFLP were sequenced using *en*v-forward-primer env_{5099} and *env*-reverse primer env_{5521r}; both primers were modified on the 5′-end with IRD800 fluorescent colour. A sequencing reaction was carried out using the ThermoSequenase fluorescent-labelled primer cycle sequencing kit with 7 deaza-dGTP. The reaction was composed of PCR-amplicons (same primers unmodified) diluted 1:10 as template DNA and sequence primer 1–2 pmol under the following conditions: 25 cycles using 95 °C, 30 sec and 60 °C, 30 s.

Data collection and data analyses were carried out automatically using BaseImageIR™ software (V.4.1). Multiple sequences were aligned using the ClustalW facility of DAMBE, V.4.0.84 [34]. Pairwise comparison of nucleotide sequence and searches for homologous DNA sequences were performed with the program BLAST 2.1 available on the National Center for Biotechnology Information (NCBI) website [2] using all available sequence-database libraries. The sequences we obtained were compared with 15 sequences from GenBank® that were included in Table 1. Four sequences (1272 D/W, 1002 D, 505 DD/W and 1137NZ) belonging to a database from the Institute of Epidemiological Diagnostics in Wusterhausen (Beier, personal communication), were also included in the comparison.

Base composition bias was assessed by evaluating the effective number of codons [26], and the distribution of sequence variation was analysed using mean genetic distances within and between clusters of isolates.

A powerful method of discriminating between positive Darwinian selection and neutral polymorphism is to compare the rates of synonymous (d_s) and non-synonymous (d_n) nucleotide substitutions per site (d_s/d_n) [26]. The null hypothesis is the neutral drift and

Table 1. BLV isolates and genomic sequences

Isolate	Background information	Accession no
	Sequences from GenBank®	
JAP	BLVCG, complete genome, Japan, 1985	K02120
AUST	Australian-BLV isolate, complete genome, Australia, 1990	D00647
ARG38	ARG38 Argentine isolate, complete genome, Argentina, 2000	AF257515
D3	BLV isolate 3, env, partial cds, Germany, 1997	U87872
D15	BLV strain cow 527, env, partial cds, Germany, 1997	AF007764
D18	BLV strain cow 134, env, partial cds, Germany, 1997	AF007763
D23	BLV isolate 23, env, partial cds, Germany, 1997	U87873
VDM	VdM isolate from bovine tumor, env, complete cds, 1990	M35239
LB285B	LB285 isolate from bovine tumor, env, complete cds, Belgium, 1990	M35240,
FLK	FLK-BLV clone pBLV 913, env, complete cds, 1990	M35242
LB59F	LB59 isolate, cell line LB59Lyc, env, complete cds, France, 1990	M35238
IT2	North Italian isolate I2 isolated from blood leukocytes, env, partial cds, Italy, 1996	S83530
Bras1	BLV isolate 8513 isolated from blood leukocytes, env partial cds, Brazil, 2001	AF399702
Bras2	BLV isolate 30, env complete cds, Brazil, 2001	AF399703
Bras3	BLV isolate 384, isolated from blood leukocytes, env complete cds, Brazil, 2001	AF399704
GAGA	BLV from bovine tumour DNA gag, pol and env, complete cds, Belgium, 1985	M10987-K02251
	Sequences from our study	
SF2	Isolate BLV-ARGSF2, env partial cds, Argentina, 1999	AY178820
SF6	Isolate BLV-ARGSF6, env partial cds, Argentina, 1999	AY178818
SF7	Isolate BLV-ARGSF7, env partial cds, Argentina, 1999	AY178817
SF8	Isolate BLV-ARGSF8, env partial cds, Argentina, 1999	AF485773
SF11	Isolate BLV-ARGSF11, env partial cds, Argentina, 1999	AF485775
SF13	Isolate BLV-ARGSF13, env partial cds, Argentina, 1999	AY178819
ASF406	Isolate BLV-ASF406, env partial cds, Argentina, 1999	AF485774

equiprobable nucleotide mutations given the specific codon at each position (neutral theory). We analysed the d_s/d_n for complete alignment of *env* (isolates Bras2, Bras3, Arg38, VdM, FLK, Jap, Aust, GAGA, LB59F and LB285B), *pro*, *pol* and *gag* (isolates Arg38, Jap, Aust and GAGA) genes and some target areas within them. Based on overall and pairwise comparisons, the proportions of possible nucleotide substitutions for all sites for both synonymous and non-synonymous sites were calculated (Pamilo-Bianchi-Li-Method). Evidence of positive, neutral or purifying selection was sought with a Z-test based on the number of synonymous substitutions per synonymous site (K_S) and the number of non-synonymous substitutions per non-synonymous site (K_N). This test evaluated whether the null hypothesis $K_S = K_N$ can be rejected in favour of one of the alternative hypotheses (positive selection for $K_S < K_N$ or purifying selection for $K_S > K_N$). The analyses were conducted using MEGA version 2.1 [19].

Phylogenetic analysis

Only seven of our sequences - ARGSF8, ARGSF11, ASF406, ARGSF2, ARGSF6, ARGSF7 and ARGSF13 – were used for phylogenetic reconstruction because the other six were indentical to these. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 and PHYLIP [14]. Phylogenetic trees were constructed by a neighbour-joining method using the Kimura Two-Parameter model (NJ) of nucleotide substitution [31]. The model assumes that there is independent change at all sites, but allows for a difference between transition and transversion rates. The reliability of the tree was assessed by 1000 bootstrap replications and using a bootstrap confidence value (P_B) of 70 % [26].

Results

Detection of BLV infected cattle

Proviral DNA was detected in whole blood samples by PCR, and antibodies were detected in serum. Positive PCR reactions were found in 311 out of 445 animals (70%), and 280 out of 311 (90%) of these positive reactors also tested positive in a serological test. Moreover, 7% (31 out of 445) of the animals had negative results in the serological test but positive results in the PCR test, probably due to a recent infection.

Table 2 shows the sero-prevalence and PCR results, stratified by county and herd. On most farms, high proportions of adult cows were infected. In addition,

Table 2. Distribution of BLV isolates by RFLP (n = 121) and sequencing (n = 13) according to geographical location and herd

Country	Herd	Prevalence (%)[a]	PCR (n)	RFLP (n = 121)			Sequencing (n = 13)	
				Aust.	Arg.	Indef.	Arg.	Aust.
Rafaela								
	3	60	8	0	7	0	1	0
	5	99	28	0	8	0	1	0
	7	99	2	1	1	0	0	0
G. Roca								
	4	99	89	0	13	0	2	0
	6	82	85	0	30	0	1	0
Susana								
	2	92	68	1*	31	0	4	0
	8	nd	3	0	3	0	nd	nd
B. Italia								
	1	4	25	1	21	1	3	1
Ataliva								
	9	85	3	0	3	0	nd	nd
Total			311	3	117	1	12	1

*Resulting classification did not match with sequencing

[a]Herd prevalence based on results from serum Blocking ELISA 108 results

nd = test not done

we found that some infected animals consistently showed negative results to serological tests after merging with the results of the longitudinal study.

RFLP results

The analysis of digestion products after using three restriction endonucleases from 121 samples showed that three samples were classified as Australian type according to the reference patterns [3] of digestion products (*Bam*HI: 316 bp and 128 bp; *Pvu*II: 444 bp and *Bcl*I: 219 and 225 bp). Furthermore, 117 samples showed a different pattern: (*Bam*HI: 316 bp and 128 bp; *Pvu*II: 280 bp and 164 bp and *Bcl*I: 219 and 225 bp), which, to our knowledge, has not been published before. Using sequencing, we could confirm that they corresponded to Argentine isolates. Only one sample could not be classified as any known genotype based on RFLP.

The results of RFLP and sequencing, stratified according to county and herd, are shown in Table 2. One genotype (Argentine isolates) was present in all herds, but the other genotype (Australian isolate) was also present in two herds. Only a single RFLP result in Herd 2 – showing an Australian pattern but a nucleotide alignment that corresponded to ARG38 isolate – did not match the classification using sequencing. Therefore, correspondence between RFLP and nucleotide sequencing results was 92%. Samples with duplicate results in RFLP (22) matched the same characteristic pattern in both test runs (data not shown).

Sequence analysis

The RFLP patterns were confirmed by sequencing 13 samples characterised by RFLP and comparing them with sequences available in GenBank®. Results from the search in BLAST showed that 12 of the nucleotide sequences we obtained exhibited high homology ($E = 0.0$) with BLV Arg38 isolate; identities ranged from 99 to 100%. The remaining sequence showed high homology ($E = 0.0$) with Australian BLV isolate, also with high identity (99%).

As expected, single base changes were found in some of the isolates, variations among these sequences consisted of 20 nucleotide substitutions in total (6% of the reported sequence), no insertions were observed and they seemed to be distributed randomly. Most of these substitutions were located in the portion of the sequence that belongs to the NH_2 half of the gp51 protein. A few differences in the effective number of codons used between isolates were detected, but these differences were not significant ($P > 0.05$).

The deduced amino acid sequences of some of our isolates and four known BLV isolates are compared in Fig. 1. It shows a conserved segment of gp51 (5154 to 5520) and resulted in four amino acid substitutions (67% of total substitutions); three of them occurred in isolates from the Argentine group and one in our ASF 406. Most of these mutations were located within the peptide region that induces a $CD8^+$ Cytotoxic T-Cell response of the host [11] and the Proline-rich region [18].

One interesting result is that all Argentine isolates shared a substitution of Asparagine 31 (N) (neutral-polar side) for Aspartic Acid (D) (acidic) at the location

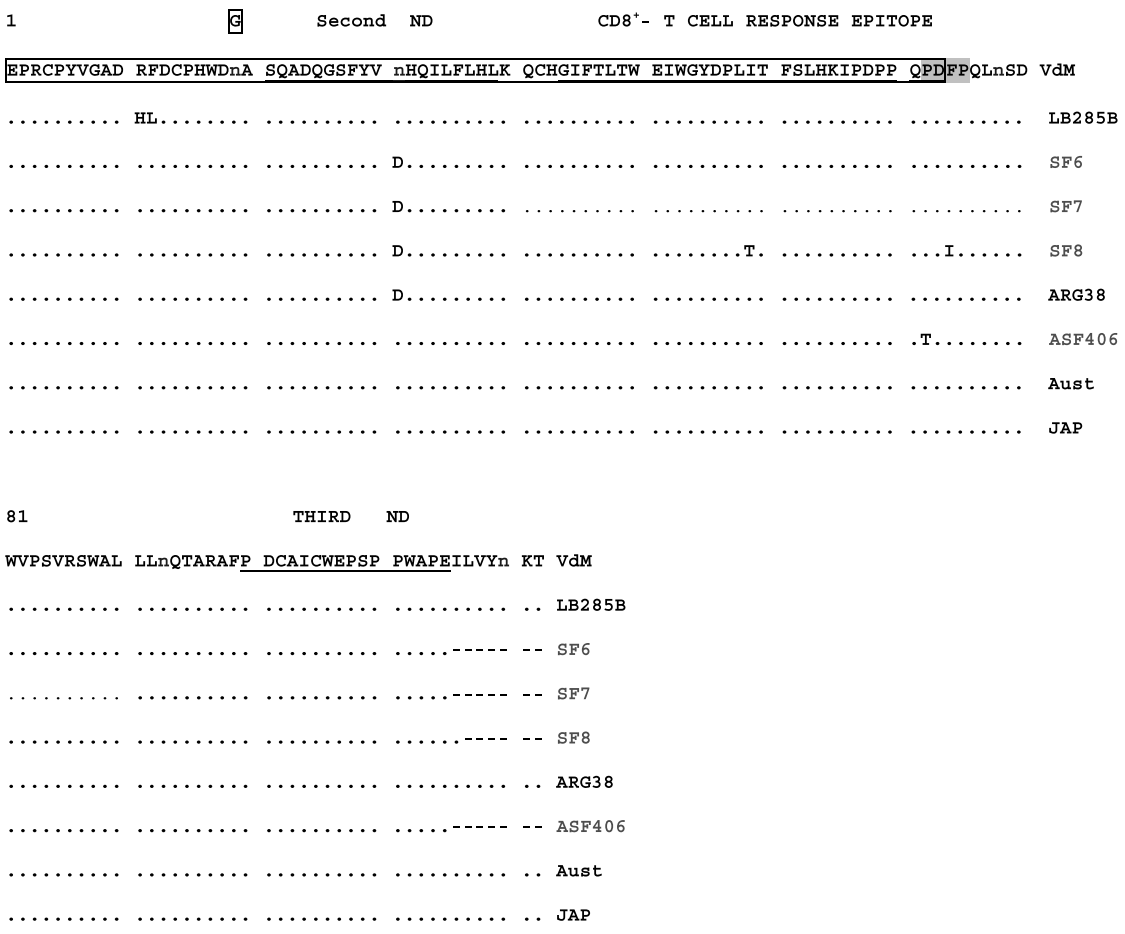

Fig. 1. Deduced amino acid sequences of some *env* proteins from some of our isolates compared to other published BLV strains: VdM – reference at top -, LB285B, ARG38, AUST and JAP. Bullets show identical amino acids with reference sequence; amino acid substitutions are indicated. Missing amino acids are represented by a hyphen (-). N-linked glycosylation sites are marked in small caps. The immunostimulatory epitope for CD8 T cells is underlined to indicate its beginning and end. *ND*: Neutralising Domains, underlines indicate the beginning and end of the NDs. The framed part of the sequence represents the NH_2 half of the gp51 protein. G Conformational epitope G. The shaded background shows the Proline-rich area

of the second neutralising domain. This area was shown to be part of the receptor-binding surface, an area which seems to be critical for receptor binding and infection [16]. The potential glycosylation sites and cysteine residues did not show variations.

Phylogenetic analysis of BLV isolates

The final tree obtained by using the neighbour-joining distance method is represented in Fig. 2A and shows the distance between isolates and the grouping of isolates in 3 main clusters. Figure 2B represents the condensed tree obtained

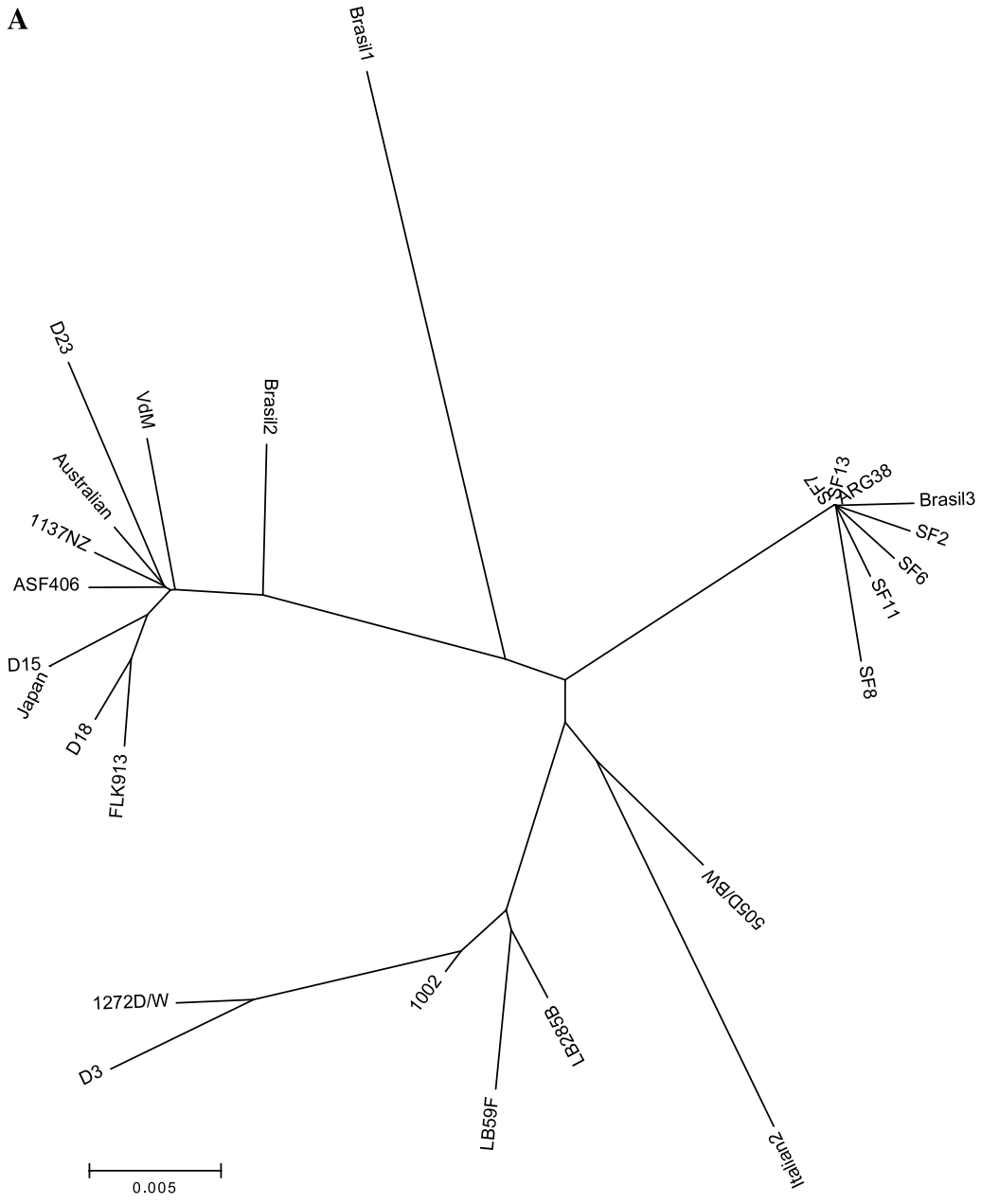

Fig. 2 (*continued*)

after bootstrapping and showed a tree that has low statistical support (P_B) values for some interior branches. With this tree we can therefore emphasize reliable portions of the branching patterns (for topology), but branch lengths are not proportional to the number of nucleotide substitutions. Nevertheless, it shows that clusters observed in the original tree are well-supported. One cluster is composed of the Argentine isolates and one of the isolates from Brazil (Brazil 3). The other cluster is represented by several isolates found in various European countries

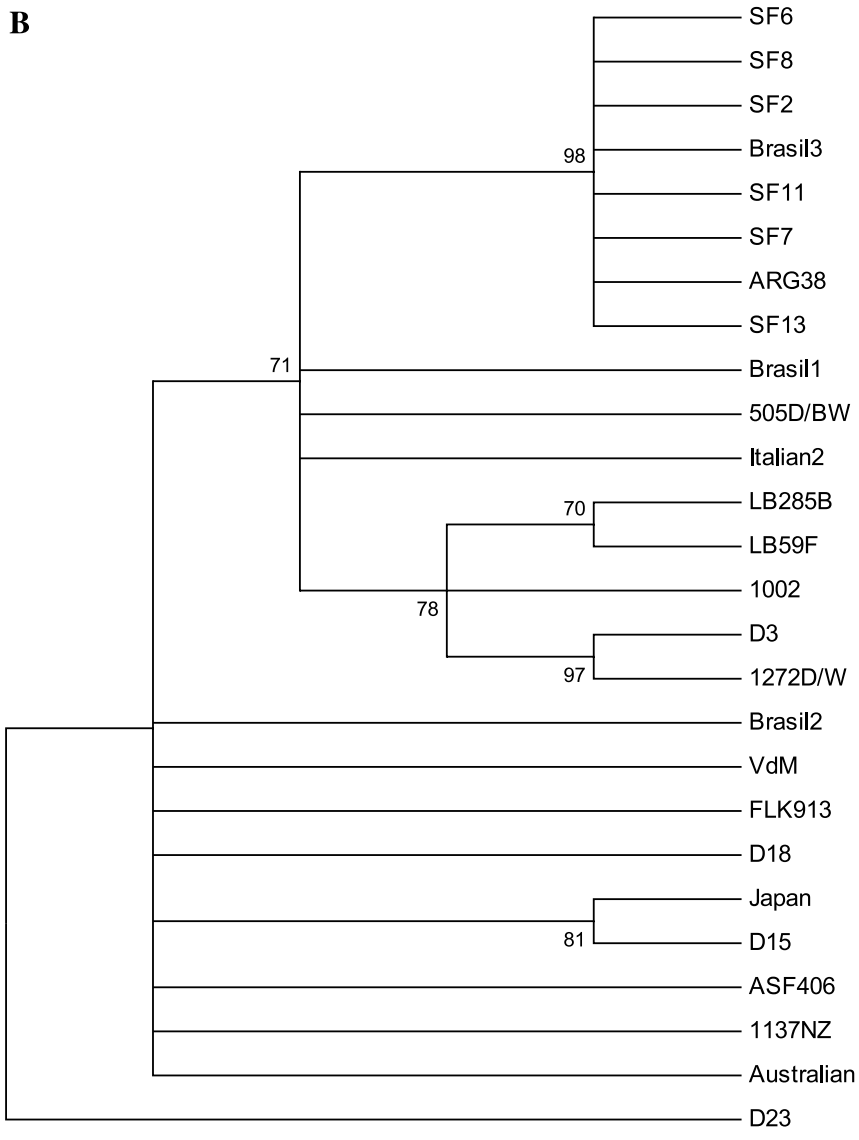

Fig. 2. **A** Phylogenetic tree constructed using the method based on a 350 bp segment of the *env* region of BLV. **B** Neighbour-joining consensus tree generated with MEGA2 software using a Kimura two-parameter model. Bootstrap analysis was applied with 1000 replications; values above 70 out of 100 are shown at the branch points

(Belgium (LB285B), Germany (D3, 1002, 1272 D/W), the northern region of France (LB59F)) and one isolate from Brazil (Brazil1). The third cluster included BLV isolates found in Japan and Germany (D15).

The mean distance between isolates, estimated after 1000 replications of bootstrapping, was 0.028 ± 0.005 (SE). Table 3 summarises variability within clusters. The higher variability of the European cluster seems to be a feature of this group; the larger number of subgroups in Fig. 2B, supported by bootstrapping, also reflects this variability.

Table 3. Mean net genetic distances (d) and SE (between brackets) within clusters of genotypes using Kimura 2-parameter model after 1000 bootstrapping replications

Cluster	d	SE
Argentine	0.004	0.002
Japan	0.000	0.000
Europe	0.014	0.005

Proportion of synonymous and non-synonymous changes in env and pol sequences

Table 4 summarises the results of the analysis for estimating proportion of synonymous (d_s) and non-synonymous (d_n) nucleotide substitutions per site and the Z-test for selection testing for *gag*, *pro*, *pol* and *env* genes and several regions within them. For *gag* sequences, the rates of synonymous nucleotide substitutions per site exceeded the non-synonymous rates; this difference was statistically significant. This pattern indicates purifying selection. The same tendency was

Table 4. Overall number of Synonymous (d_s) and Non-synonymous (d_n) nucleotide substitutions per 100 sites (±SE) calculated by Pamilo-Bianchi-Li Method and Codon-based Z-test for Tests of Selection, using different BLV genes

Gene	Protein	Peptides	d_s	d_n	Z-test
Gag			45.0 ± 3.7	35.1 ± 2.6	**
	p15		12.6 ± 2.8	2.4 ± 0.8	**
	p24		28.5 ± 3.4	14.8 ± 1.9	***
	p12		8.6 ± 3.0	0.74 ± 0.5	**
Pro			8.9 ± 2.1	2.3 ± 0.7	**
Pol			9.6 ± 0.8	1.8 ± 0.2	***
	RNase/RT		9.8 ± 0.9	2.1 ± 0.3	***
	Endonuclease		9.2 ± 1.3	1.2 ± 0.4	***
Env			6.1 ± 0.6	1.3 ± 0.2	***
	Leader Peptide		4.6 ± 2.8	2.3 ± 1.1	NS
	gp51		5.5 ± 0.9	1.3 ± 0.3	***
		Receptor-Binding Domain	5.9 ± 1.3	1.2 ± 0.4	**
		Neutralising Domains	3.8 ± 1.7	0.9 ± 0.5	NS
		CD8-T Cell	9.3 ± 4.9	0.0 ± 0.0	*
		TMHR[a]	4.6 ± 2.6	3.2 ± 1.1	NS
		T-Helper Cell	0.0 ± 2.9	0.0 ± 0.0	NS
	gp30		7.1 ± 1.2	1.2 ± 0.4	***
		GD21	8.9 ± 3.6	0.2 ± 0.2	**

Tests to reject the null hypothesis that $d_s = d_n$: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$
NS = null hypothesis that $d_s = d_n$

[a]Transmembrane Hydrophobic Region

observed for p15, p24 and p12. Similar results were observed for the protease. For *pol*, rates of nucleotide substitution were estimated for the whole gene and for two subunits (RNase/RT and endonuclease), and in each case, the rate of synonymous substitutions per site was significantly greater than non-synonymous substitutions, also indicating purifying selection. The pattern of nucleotide substitution seen in the case of the *env* gene did not differ from the others, although there are subtle differences in the statistical tests considering different subunits. For the main proteins (gp51 and gp30), the results lead to similar interpretations as for the other regions, but for the leader peptide the test indicates that there is no evidence to reject the neutral theory.

Considering that the gp51 protein encompasses several sites that are subject to strong interaction with the host's immunoresponse, several important epitopes were evaluated. The Z-test result from regions that cover the conformational (Receptor-Binding Domain) and the immunomodulator CD8-T cell epitopes also indicates probable purifying selection. In contrast, for the Neutralizing domains, T-Helper cells epitope and Transmembrane Hydrophobic Region (TMHR), there is no evidence that rejects the neutral theory. The TMHR was the only site with rates of non-synonymous nucleotide substitutions per site that were greater than the synonymous rates, but the difference was not statistically significant. The results after conducting the pairwise comparison of all sequences indicate that for *gag*, *pro*, *pol* and *env*, all pairs are submitted to purifying selection. In contrast, the TMHR within the *env* gene showed that for some pairs of sequences the tests of Selection indicate positive selection (Table 5).

Cases of persistently negative serological tests

In Herd 1, after matching results from cross-sectional and longitudinal studies, we found that three animals that had been selected for sequencing (ARGSF7, ARGSF8 and ASF406) were persistently infected but continually tested negative according to serological tests. Moreover, these animals tested negative for milk

Table 5. Pairwise comparison of the number of Synonymous (d_s) and Non-synonymous (d_n) nucleotide substitutions per 100 sites (±SE) calculated by Pamilo-Bianchi-Li Method and Codon-based Z-test for Tests of Selection, for Transmembrane Hydrophobic Region of the BLV *env* gene

Pair of sequences	d_s	d_n	Z-test*
Arg38-LB59B	3.0 ± 3.5	16.1 ± 5.0	**
Arg38-Aust	3.0 ± 3.5	16.1 ± 5.0	**
Arg38-FLK	3.0 ± 3.5	16.1 ± 5.0	**
GAGA-LB285F	3.0 ± 3.5	16.1 ± 5.0	**
Arg38-Jap	3.0 ± 3.5	16.1 ± 5.0	**
Arg38-VDM	3.0 ± 3.5	16.1 ± 5.0	**

*Tests to reject the null hypothesis that $d_s = d_n$: **$P < 0.05$

and serum ELISA 108 one year before and every two months for two years after the sample was taken that appeared positive in the PCR. They were consistently infected because the PCR was positive two years later and the RLFP results of the different samplings were also identical (results not shown).

The analysis of the predicted amino acid showed that two of these three sequences (ARGSF8 and ASF406) differed from the rest of the isolates in the peptide region that induces a $CD8^+$ Cytotoxic T-Cell response of the host. For isolate ASF406, Proline (P) (hydrophobic–inflexible) changed to Threonine (T) (neutral-polar side) in position 84. For ARGSF8, Isoleucine (I) (hydrophobic–aliphatic) changed to Threonine (T) (neutral-polar side) in position 69 and Phenylalanine (F) (hydrophobic–aromatic non-polar) changed to Isoleucine (I) (hydrophobic–aliphatic, large non-polar) in position 82 (Fig. 1).

Discussion

We studied the RLFP and sequence variation of BLV isolates obtained from various cattle herds in Argentina. Our results showed the presence of two types of BLV isolates – Australian and Argentine – in dairy herds from different areas of Central Argentina, which is one of the major dairy regions of South America. Our results complement other studies [11] that report the presence of Argentine isolates in other geographical areas. These findings provide firm evidence that the Argentine isolate is the predominant isolate in cattle populations of Argentina. However, our phylogenetical analysis shows that variation occurs within the cluster of Argentina-like isolates, as indicated by the isolate from Brazil (Brazil 3) [7]. This suggests that Argentine isolates are not geographically restricted to a single country. It is unknown if the presence of the Argentine genotype in Brazil represents a common native variant or if this is the result of introducing infected animals from Argentina. Nevertheless, the phylogenetical tree shows that the Argentine clusters are clearly different from other clusters.

We found only Argentine isolates in most of the herds we analysed except for two herds where both types of isolates were found. The fact that we ascertained both single and multiple genotypes on individual farms agrees with previous findings [7, 12, 20, 27]. Unexpectedly, we found that both farms with dual isolates were closed herds, i.e. they had not purchased replacement animals from other farms, therefore ruling out the possibility of introducing infected cattle. Although another study [11] pointed out the closer relationship between the Argentine and Australian genotypes than reported in GenBank®, this fact cannot be used to attribute the presence of these genotypes to evolutionary linkages after a single introduction of an infected cow. Moreover, introduction of the virus via other vectors such as hematophagous insects from neighbouring farms or contaminated needles during mass inoculations (e.g. vaccination for foot and mouth disease) may also explain the presence of two variants.

It is known that variations in the genotype of retroviruses in general may be associated with changes in disease progression [22] or in differences in infectivity between BLV strains, as was previously hypothesised [6, 11]. In our study, we

found a different distribution of isolates and differences in prevalence. Although our findings showed high rates of BLV infection within Argentine dairy herds, which is in line with previous studies [1, 33], we cannot conclude that the isolates in this study are representative strains that cause a high infection rate. The high prevalence probably reflects the lack of a compulsory control programme and not a high infectivity of the genotype present in the herd.

The phylogenetic data concerning genetic distances, based either on nucleotides or translated amino acids, showed relatively low divergence between various isolates, as previously reported [10, 11, 21, 27]. Nevertheless, the topology of phylogenetical tree clearly shows that Argentine isolates represent a separate and homogeneous group compared to other clusters. Although we could demonstrate a detectable immune response in most infected cattle, we found some persistently infected animals that did not show detectable antibody response. This confirms the findings of other studies [5, 13]. One study [13] suggests that this situation could be related to certain genotypes of the virus, but others [20] identified isolates identical to that study, and those animals showed a detectable immunoresponse. This could be explained by differences in the time after infection when the samples were taken or differences in immunocompetence between the animals. We also detected persistently infected and seronegative animals, but these animals had different BLV genotypes. We therefore concluded that occurrence of seronegative carriers is not related to one specific genotype.

We found one distinctive feature in two of these seronegative animals. The amino acids substitutions were mainly clustered in the segment that is associated with a $CD8^+$ response of the host. Further study will be needed to determine if this affects the three-dimensional structure and subsequently the function of a protein, and to determine the possible relationship with seronegative carriers.

The low rate of non-synonymous substitutions compared to synonymous substitutions found in our analysis supports the hypothesis of purifying selection of *env*, *gag* and *pol* genes and several subunits; consequently, molecular evolution occurs under some functional constraint. Similar results [4, 32] were also found for other retroviruses. These results are consistent with the low divergence between isolates mentioned above and help to explain this distinctive feature of the BLV-HTLV genus. Nevertheless, it appears that the TMHR is a site that might be a good potential candidate for future functional studies.

Another striking outcome is that our interpretation of the results from selection tests applied on the TMHR indicates that – at least for the Argentine isolate – the host influences this site. Additional data on other Argentine genotypes are lacking, making it difficult to evaluate determine if these results are consistent with this genotype cluster. To link selection forces and polymorphism, our results should be complemented with a study of the viral population within a given host that shows fixation of such mutation [32].

In conclusion, we characterised several BLV field isolates and their distribution in dairy herds in Argentina. Most of the isolates were represented in a BLV cluster, but point mutations did occur. We also detected seronegative carrier animals. Finally, we showed that selective advantages of different genotypes seem to be

forced by purifying selection, and that this area could be an interesting topic for future functional studies.

Acknowledgements

The research was supported by the Secretaria de Ciencia y Técnica de la Nación, PICT no. 08-04541, BID 802 OC-Ar, Argentina. The authors would like to thank Dr. F. Conraths, Director of the Institute of Epidemiological Diagnosis, Federal Research Centre for Virus Diseases of Animals, for his support and comments, Dr. L. Geue for his assistance in sequencing and comments and Ms. Schares for her technical assistance in the lab. We would also like to thank Prof. H. F. J. Savelkoul, Dr. A. Groen, Dr. K. Frankena and Dr. T. Van Kempen from Wageningen University for their comments and assistance during analysis of sequences.

References

1. Alejo D, Gutierrez S, Dolcini G, Esteban E, Odeon A, Fernandez Sainz I, Casaro A, Sainz IF (2000) Prevalence of infection with bovine leukemia virus (BLV) in dairy herds in the area of General Pueyrredon and Balcarce. Rev Arg Prod Anim 20: 77–83
2. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402
3. Beier D, Blankenstein P, Marquardt O, Kuzmak J (2001) Identification of different BLV provirus isolates by PCR, RFLPA and DNA sequencing. Berl Munch Tierarztl Wochenschr 114: 252–256
4. Biek R, Rodrigo AG, Holley D, Drummond A, Anderson CR Jr, Ross HA, Poss M (2003) Epidemiology, genetic diversity, and evolution of endemic feline immunodeficiency virus in a population of wild cougars. J Virol 77: 9578–9589
5. Blankenstein P, Bondzio A, Buchel A, Kinder E, Ebner D (1996) Investigations of the Japanese bovine tumour virus (BLV) – its ability to express structural and regulatory BLV proteins. J Vet Med B 43: 343–349
6. Bruck C, Rensonnet N, Portetelle D, Cleuter Y, Mammerickx M, Burny A, Mamoun R, Guillemain B, van der Maaten MJ, Ghysdael J (1984) Biologically active epitopes of bovine leukemia virus glycoprotein gp51: their dependence on protein glycosylation and genetic variability. Virology 136: 20–31
7. Camargos MF, Stancek D, Rocha MA, Lessa LM, Reis JK, Leite RC (2002) Partial sequencing of env gene of bovine leukaemia virus from Brazilian samples and phylogenetic analysis. J Vet Med B Infect Dis Vet Public Health 49: 325–331
8. Coffin JM, Hughes SH, Varmus HE (1997) Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
9. Coulston J, Naif H, Brandon R, Kumar S, Khan S, Daniel RC, Lavin MF (1990) Molecular cloning and sequencing of an Australian isolate of proviral bovine leukaemia virus DNA: comparison with other isolates. J Gen Virol 71: 1737–1746
10. Dube S, Bachman S, Spicer T, Love J, Choi D, Esteban E, Ferrer JF, Poiesz BJ (1997) Degenerate and specific PCR assays for the detection of bovine leukaemia virus and primate T cell leukaemia/lymphoma virus pol DNA and RNA: phylogenetic comparisons of amplified sequences from cattle and primates from around the world. J Gen Virol 78: 1389–1398

11. Dube S, Dolcini G, Abbott L, Mehta S, Dube D, Gutierrez S, Ceriani C, Esteban E, Ferrer J, Poiesz B (2000) The complete genomic sequence of a BLV strain from a Holstein cow from Argentina. Virology 277: 379–386
12. Fechner H, Kurg A, Geue L, Blankenstein P, Mewes G, Ebner D, Beier D (1996) Evaluation of polymerase chain reaction (PCR) application in diagnosis of bovine leukaemia virus (BLV) infection in naturally infected cattle. J Vet Med B 43: 621–630
13. Fechner H, Blankenstein P, Looman AC, Elwert J, Geue L, Albrecht C, Kurg A, Beier D, Marquardt O, Ebner D (1997) Provirus variants of the bovine leukemia virus and their relation to the serological status of naturally infected cattle. Virology 237: 261–269
14. Felsenstein J (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). Cladistics 5: 164–166
15. Foxman B, Riley L (2001) Molecular epidemiology: focus on infection. Am J Epidemiol 153: 1135–1141
16. Gatot JS, Callebaut I, Van Lint C, Demonte D, Kerkhofs P, Portetelle D, Burny A, Willems L, Kettmann R (2002) Bovine leukemia virus SU protein interacts with zinc, and mutations within two interacting regions differentially affect viral fusion and infectivity in vivo. J Virol 76: 7956–7967
17. Gutierrez SE, Dolcini GL, Arroyo GH, Rodriguez Dubra C, Ferrer JF, Esteban EN (2001) Development and evaluation of a highly sensitive and specific blocking enzyme-linked immunosorbent assay and polymerase chain reaction assay for diagnosis of bovine leukemia virus infection in cattle. Am J Vet Res 62: 1571–1577
18. Johnston ER, Albritton LM, Radke K (2002) Envelope proteins containing single amino acid substitutions support a structural model of the receptor-binding domain of bovine leukemia virus surface protein. J Virol 76: 10861–10872
19. Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software. Arizona State University, Tempe, Arizona, USA
20. Licursi M, Inoshima Y, Wu DL, Yokoyama T, Gonzalez ET, Sentsui H (2002) Genetic heterogeneity among bovine leukemia virus genotypes and its relation to humoral responses in hosts. Virus Res 86: 101–110
21. Mamoun RZ, Morisson M, Rebeyrotte N, Busetta B, Couez D, Kettmann R, Hospital M, Guillemain B (1990) Sequence variability of bovine leukemia virus env gene and its relevance to the structure and antigenicity of the glycoproteins. J Virol 64: 4180–4188
22. Mansky LM (1998) Retrovirus Mutation Rates and Their Role in Genetic Variation. J Gen Virol 79: 1337–1345
23. Miller JM, Van der Maaten MJ (1990) Bovine leukosis virus. In: Dinter Z, Morein B (eds) Virus infections of ruminants. Elsevier, Amsterdam, pp 572
24. Naif HM, Brandon RB, Daniel RC, Lavin MF (1990) Bovine leukaemia proviral DNA detection in cattle using the polymerase chain reaction. Vet Microbiol 25: 117–129
25. Naif HM, Daniel RC, Cougle WG, Lavin MF (1992) Early detection of bovine leukemia virus by using an enzyme-linked assay for polymerase chain reaction-amplified proviral DNA in experimentally infected cattle. J Clin Microbiol 30: 675–679
26. Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York
27. Portetelle D, Couez D, Bruck C, Kettmann R, Mammerickx M, Van der Maaten M, Brasseur R, Burny A (1989) Antigenic variants of bovine leukemia virus (BLV) are defined by amino acid substitutions in the NH_2 part of the envelope glycoprotein gp51. Virology 169: 27–33
28. Rice NR, Stephens RM, Couez D, Deschamps J, Kettmann R, Burny A, Gilden RV (1984) The nucleotide sequence of the env gene and post-env region of bovine leukemia virus. Virology 138: 82–93

29. Rice NR, Stephens RM, Burny A, Gilden RV (1985) The gag and pol genes of bovine leukemia virus: nucleotide sequence and analysis. Virology 142: 357–377
30. Sagata N, Yasunaga T, Tsuzuku-Kawamura J, Ohishi K, Ogawa Y, Ikawa Y (1985) Complete nucleotide sequence of the genome of bovine leukemia virus: its evolutionary relationship to other retroviruses. Proc Natl Acad Sci USA 82: 677–681
31. Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425
32. Seibert SA, Howell CY, Hughes MK, Hughes AL (1995) Natural selection on the gag, pol, and env genes of Human immunodeficiency virus (HIV-1). Mol Biol Evol 12: 803–813
33. Trono KG, Perez-Filgueira DM, Duffy S, Borca MV, Carrillo C (2001) Seroprevalence of bovine leukemia virus in dairy cattle in Argentina: comparison of sensitivity and specificity of different detection methods. Vet Microbiol 83: 235–248
34. Xia X, Xie Z (2001) DAMBE: Data analysis in molecular biology and evolution. J Hered 92: 371–373

Author's address: Dr. Gustavo Monti, Wageningen University and Research Centre, Quantitative Veterinary Epidemiology Group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands; e-mail: Gustavo.Monti@WUR.nl