Genetics and Molecular Research

Genetic variability and phylogeny of the 5' long terminal repeat from Brazilian bovine leukemia virus

C. Hirsch[1], M.F. Camargos[2], E.F. Barbosa-Stancioli[3], A.A. Fonseca Júnior[2], D.S. Rajão[1], M.B. Heinemann[1], J.K.P. Reis[1] and R.C. Leite[1]

[1]Laboratório de Retroviroses, Departamento de Medicina Veterinária Preventiva, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil
[2]Laboratório de Biologia Molecular, Laboratório Nacional Agropecuário, Pedro Leopoldo, MG, Brasil
[3]Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

Corresponding author: J.K.P. Reis
E-mail: jenner@ufmg.br

Genet. Mol. Res. 14 (4): 14530-14538 (2015)
Received January 14, 2015
Accepted June 25, 2015
Published November 18, 2015
DOI http://dx.doi.org/10.4238/2015.November.18.16

ABSTRACT. We conducted a phylogenetic analysis of 22 strains of bovine leukemia virus obtained by polymerase chain reaction to amplify a 582-base pair fragment of the transcriptional regulatory region 5' long terminal repeat (LTR). Twenty-two samples of proviral DNA from peripheral blood mononuclear cells containing bovine leukemia virus from naturally infected bovine from 4 distinct geographic regions in Brazil were investigated. The products obtained by polymerase chain reaction were subjected to direct sequencing and sequence alignment. Fragments of 422 nucleotides were obtained, located between positions -118 and +303 base pairs of the 5'LTR. These fragments corresponded to 80% of the LTR region and included 56% of sub-region U3, 100% of R, and 82.5% of U5. Phylogenetic analysis of these sequences showed a high conservation degree in the 5'LTR region, with 5

well defined groups. However, a hotspot occurrence in the R-U5 region was also observed, which contained 40% of all nucleotide variability observed.

Key words: Bovine leukemia virus; Long terminal repeat; Phylogeny

INTRODUCTION

Enzootic bovine leukosis (EBL) is a viral disease that causes hematologic and neoplastic disorders. The disease is caused by bovine leukemia virus (BLV), which belongs to the Retroviridae family. Progression of BLV infection in the host is characterized by a prolonged incubation period, during which viral transcription regulation mechanisms suppress immune recognition, maintaining the infected animal as a reservoir (Kabeya et al., 2001). Clinical EBL occurs in 1-5% of infected bovine and is characterized by lymphoproliferative neoplasia in multiple systems of the host (Camargos et al., 2007). Thus, even animals without clinical signs undergo a humoral immune response against viral proteins. Of these animals, 30% show increased circulating lymphocytes as a result of B cell clonal expansion, a condition known as persistent lymphocytosis (Debacq et al., 2002).

BLV is highly disseminated in Brazilian dairy herds (Camargos et al., 2002). A high infection rate leads to a higher probability of generating mutant viral strains, which may cause development of variants with different pathogenic characteristics, such as antigenic variability, cellular tropism, oncogenicity, and interspecies transmission. This variability is associated with the low replicative fidelity of retroviruses reverse transcriptase (RT) (Avidan et al., 2002), and results in a high evolutional rate and adaptation of strains to different locations.

Phylogenetic analysis of the 5'long terminal repeat (LTR) region of human T lymphotropic virus has been used as a sample differentiation tool, allowing for geographic group characterization, evolutional rate determination, and epidemiologic chain establishment, as well as used for association analysis of specific pathologies (Van Dooren et al., 2007). However, there is little information regarding the BLV LTR region. Zhao et al. (2007) found significant variability in studied sequences of the 5'LTR from BLV strains from North America, Costa Rica, Japan, and Belgium, identifying 7 genetic groups and demonstrating that the transcriptional regulation region can be analyzed in phylogenetic studies of the virus. Furthermore, LTR region variability may contribute to viral transcription regulation, a process that is directly related to BLV virulence and pathogeny (Moratorio et al., 2013). In addition, small nucleotides polymorphisms identified in the LTR region may be responsible for the emergence of the 2 different infection profiles (Juliarena et al., 2013).

Here, we describe 5'LTR region variability in Brazilian BLV strains and provide partial characterization of BLV genetic diversity based on the transcriptional regulation region.

MATERIAL AND METHODS

Genomic DNA was extracted from peripheral blood mononuclear cells from 22 naturally BLV-infected bovine (Camargos et al., 2002). Animals were previously tested using the agar gel immunodiffusion test (AGID/gp 51) and polymerase chain reaction (PCR) according to Camargos et al. (2002), and showed positive results in both tests. Samples were obtained from 4 distinct geographic regions in Brazil: 9 from Londrina, PR, 1 from Corumbá, MS, 4 from Coronel Pacheco, MG, and 8 from Pedro Leopoldo, MG.

PCR amplified a 582-base pair portion of the BLV proviral DNA 5'LTR region according to the sequence published by Sagata et al. (1985). Primer-pairs were designed using Oligo Primer

Analysis version 3.3 (Rychlik and Rhoads, 1989) and had the following sequences: forward primer 5'-GCCACCGCCCCGTAAACCAG-3' and reverse primer 5'-AGAGACCCAAAATGCCGCG-3'. Reactions contained reaction buffer, 1.5 mM $MgCl_2$, 0.4 µL 0.2 M dNTPs, 0.1 µL 5 U/µL Taq DNA polymerase, 20 picomoles of each primer, and 2.0 µL DNA, at concentrations of 10-100 ng/µL, for a final volume of 20 µL. Reagents were from PCR Core System I kit (Promega, Madison, WI, USA). Positive (FLK cells persistently infected with BLV) and negative controls were used in each PCR experiment.

PCRs were carried out in a thermocycler PT100 (MJ Research Inc., Waltham, MA, USA), under the following conditions: 95°C for 5 min, followed by 35 cycles at 95°C for 30 s, 61°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 7 min.

PCR products were sequenced according to Camargos et al. (2007). Edited sequences were aligned using Clustal W (Jeanmougin et al., 1998) and analyzed using MEGA 5.2 (Tamura et al., 2007) for phylogeny reconstruction using 3 methods of tree construction/test (nucleotide substitution models chosen according to the AKAIKE INFORMATION CRITERION test using the same software): maximum likelihood (substitution model HKY, 1000 bootstrap replication), neighbor-joining (maximum composite likelihood, 1000 bootstrap replications), and maximum parsimony. Other methods were also tested to verify the clusters. We used Splitstree 4 and Network to construct a phylogenetic network to analyze the genetic profiles and groupings. Splitstree was used to build a phylogenetic network using Neighbornet to verify the presence of conflicting signals in the phylogeny. The Network software can be used to reconstruct a network using haplotypes by median joining (Bandelt et al., 1999). The term haplotype is used in genetics indicates a combination of alleles (DNA sequences) at adjacent locations (loci) on a chromosome and are inherited together. As viruses do not have chromosomes, the genetic profile of each unique genetic sequence found in the analysis is analyzed. In addition to the 22 Brazilian isolates obtained in this study, 57 samples from GenBank were used for phylogenetic analysis, as described in Table 1.

Table 1. Identification, accession number, and genomic region from GenBank sequences used to obtain consensus sequences and to generate phylogenetic analysis of 5'LTR/BLV region.

Locus identification	Accession No.	Country of origin	Sequence length	Reference
BLVCG	K02120	Japan	8.714	Sagata et al., 1985
BLVLTR3	K01617	USA	792	Couez et al., 1984
BLVGPE	D00647	Australia	7.933	Coulston et al., 1990
BLVLTRA3	K03103	USA	612	Derse et al., 1985
BOVLTRX1	M12527	Japan	606	Sagata et al., 1984
BLVLTRX2	M19208	Japan	614	Sagata et al., 1984
AF257515	AF257515	Argentina	8.588	Dube et al., 2000

RESULTS

A fragment of 422 base pairs was sequenced from each of the 22 BLV samples (Table 2). The sequences corresponded to 79.4% of the BLV 5'LTR region, representing 56% of U3, 100% of R, and 82.5% of U5. Phylogenetic analysis showed a high conservation degree between all 22 samples, with 394 positions conserved, representing 93.1% conservation. Variable positions were 6.85% from the total of 422 positions. From the mutation positions, the leading were transitions (22), followed by 7 transversions, 4 deletions, and 2 insertions.

Table 2. Brazilian BLV LTR samples used in this study, with GenBank accession numbers and city of origin.

Locus identification	Accession No.	Country of origin	Reference
BLVCG	K02120	Japan	Sagata et al., 1985
BLVLTR3	K01617	USA	Couez et al., 1984
BLVGPE	D00647	Australia	Coulston et al., 1990
BLVLTRA3	K03103	USA	Derse et al., 1985
BOVLTRX1	M12527	Japan	Sagata et al., 1984
BLVLTRX2	M19208	Japan	Sagata et al., 1984
AF257515	AF257515	Argentina	Dube et al., 2000
US-IA	DQ288172.1	USA	Zhao et al., 2007
US-WI	DQ288173.1	USA	Zhao et al., 2007
US-PA	DQ288175.1	USA	Zhao et al., 2007
UCD1326	DQ288175.1	USA	Zhao et al., 2007
UCD1400	UCD1326	USA	Zhao et al., 2007
UCD1402	DQ288180.1	USA	Zhao et al., 2007
UCD1453	DQ288182.1	USA	Zhao et al., 2007
UCD1571	DQ288183.1	USA	Zhao et al., 2007
UID1K	DQ288184.1	USA	Zhao et al., 2007
JP-K1	DQ288189.1	Japan	Zhao et al., 2007
JP-K2	DQ288190.1	Japan	Zhao et al., 2007
JP-K3	DQ288191.1	Japan	Zhao et al., 2007
JP-K4	DQ288192.1	Japan	Zhao et al., 2007
JP-K5	DQ288193.1	Japan	Zhao et al., 2007
JP-M1	DQ288194.1	Japan	Zhao et al., 2007
JP-M2	DQ288195.1	Japan	Zhao et al., 2007
JP-M3	DQ288196.1	Japan	Zhao et al., 2007
JP-M4	DQ288197.1	Japan	Zhao et al., 2007
JP-M5	DQ288198.1	Japan	Zhao et al., 2007
JP-A1	DQ288199.1	Japan	Zhao et al., 2007
JP-A2	DQ288200.1	Japan	Zhao et al., 2007
JP-A3	DQ288201.1	Japan	Zhao et al., 2007
JP-A4	DQ288202.1	Japan	Zhao et al., 2007
JP-A5	DQ288203.1	Japan	Zhao et al., 2007
JP-E1	DQ288204.1	Japan	Zhao et al., 2007
JP-E2	DQ288205.1	Japan	Zhao et al., 2007
JP-E3	DQ288206.1	Japan	Zhao et al., 2007
JP-E4	DQ288207.1	Japan	Zhao et al., 2007
JP-E5	DQ288208.1	Japan	Zhao et al., 2007
JP-H1	DQ288209.1	Japan	Zhao et al., 2007
JP-H2	DQ288210.1	Japan	Zhao et al., 2007
JP-H3	DQ288211.1	Japan	Zhao et al., 2007
JP-H4	DQ288212.1	Japan	Zhao et al., 2007
JP-H5	DQ288213.1	Japan	Zhao et al., 2007
JP-F1	DQ288214.1	Japan	Zhao et al., 2007
JP-F2	DQ288215.1	Japan	Zhao et al., 2007
JP-F3	DQ288216.1	Japan	Zhao et al., 2007
JP-F4	DQ288217.1	Japan	Zhao et al., 2007
CR-L1-1	DQ288218.1	Costa Rica	Zhao et al., 2007
CR-L1-2	DQ288219.1	Costa Rica	Zhao et al., 2007
UCD1205K	DQ288220.1	USA	Zhao et al., 2007
CR-L1-4	DQ288221.1	Costa Rica	Zhao et al., 2007
CR-A1-1	DQ288222.1	Costa Rica	Zhao et al., 2007
CR-A1-3	DQ288224.1	Costa Rica	Zhao et al., 2007
CR-G1-1	DQ288225.1	Costa Rica	Zhao et al., 2007
CR-L2-1	DQ288226.1	Costa Rica	Zhao et al., 2007
CR-A2-1	DQ288227.1	Costa Rica	Zhao et al., 2007
CR-A2-2	DQ288228.1	Costa Rica	Zhao et al., 2007
CR-A2-3	DQ288229.1	Costa Rica	Zhao et al., 2007
BELG-2	DQ288230.1	Belgium	Zhao et al., 2007

The phylogenetic trees created with the 78 sequences using all methods showed distinct genetic groups for the LTR, numbered here from 1-5; similar groups were described by Zhao et al. (2007). A phylogenetic tree built using maximum likelihood is illustrated in Figure 1. The division into

5 genetic groups was based on 2 factors. First, these groups were present in maximum likelihood and neighbor joining trees. We found 7 maximum parsimony trees, 5 of which formed the same groups, but the subgroups were divided into different clades. Additionally, the genetic distance between groups was larger than the genetic distance within groups (Table 3).

Table 3. Mean nucleotide distances in 400-base pair fragment of envelope gene within (intra-genotype) and among (inter-genotype) BLV genotypes.

	Gp1	Gp2	Gp3	Gp4	Gp5
Gp1	0.005 (0.001)*				
Gp2	0.020 (0.005)	0.011 (0.003)*			
Gp3	0.019 (0.006)	0.028 (0.007)	0.012 (0.005)*		
Gp4	0.018 (0.005)	0.028 (0.006)	0.021 (0.006)	0.014 (0.004)*	
Gp5	0.030 (0.006)	0.034 (0.007)	0.033 (0.007)	0.033 (0.006)	0.023 (0.006)*

*Intra-group distance.

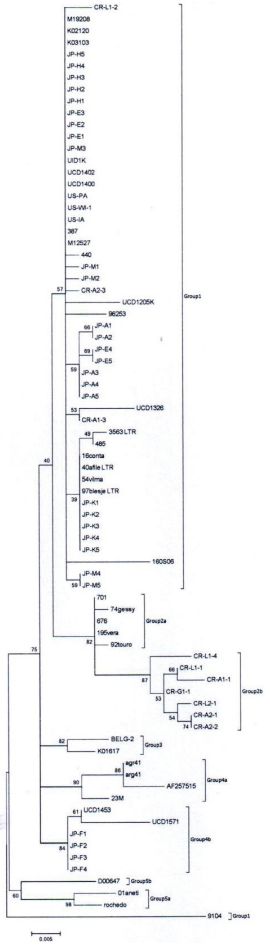

Figure 1. Phylogenetic tree constructed by the neighbor-joining method, using the maximum likelihood method, with 1000 bootstraps. The tree shows 5 distinct groups supported by bootstrap values over 70. Brazilian sequences are marked with a black circle.

Phylogenetic networks showed the same 5 groups but indicated conflicting signals, particularly in group 1, where most sequences were found. Splitstree and Network revealed incompatible and ambiguous signals in a data set represented by the parallel edges, rather than the single branches in the phylogenetic tree. The 78 sequences were grouped in 40 genetic profiles, 30 of which were unique. These genetic profiles were used to construct a phylogenetic network in which Group 1 formed the base of the network, connecting all other groups (Figure 2).

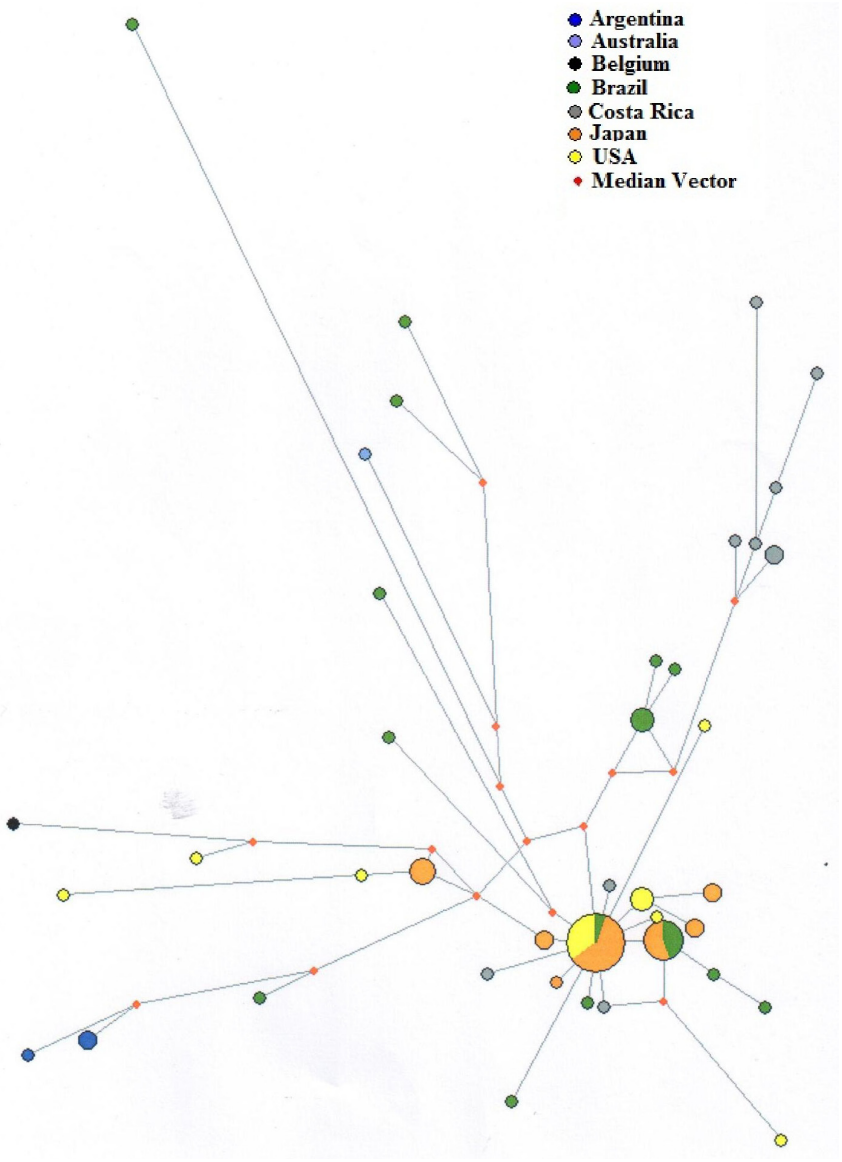

Figure 2. Phylogenetic network constructed by median joining using LTR sequences. Circle sizes represent are proportional to the number of sequences with that specific genetic profile. Median vectors represent missing intermediates.

DISCUSSION

Forty percent of all variable positions observed in this study were located in the R-U5 region, representing 20.8% of the investigated sequence and only 16.7% of LTR. This finding is important because it may be directly associated with the regulation of BLV transcription. This variability may also be related to viral cellular tropism, as cellular proteins that recognize proviral DNA at the 5'LTR position are important for promoting transcription, and thus are correlated with genetic variations in this viral region (Reed-Inderbitzin and Maury, 2003). Different nucleotide substitutions in these regions may lead to better silencing of viral transcription in the lymphosarcoma strains as a strategy to avoid recognition by the host immune response (Xiao and Buehring, 1998).

Genomic analysis of BLV Brazilian strains was used to evaluate partial cyclic AMP-responsive element sequences, which are present as triplets at the BLV 5'LTR U3 region. This is considered to be a highly conserved region with few nucleotide substitutions even in viruses isolated from different cell types such as peripheral blood mononuclear cells, tumor cells, experimentally infected sheep, and cell lines (FLK) or B-cell lymphosarcomas (Moratorio et al., 2013). Merezak et al. (2001) demonstrated that minor variations in the cyclic AMP-responsive elements result in negative transcription regulation of BLV and may contribute to viral escape from the immune system and establishment of persistent infection. These cyclic AMP-responsive element sequences were highly conserved in the Brazilian strains, suggesting that these regions are important for viral viability and, therefore, are less prone to mutations.

Brazilian BLV-LTR sequences were distributed in all 5 genetic groups except group 3. Group 1 was formed by sequences from various countries. There was more than one subgroup present in group 1, but only one appeared to be formed exclusively from sequences from one country (Japan). Group 2 was formed by sequences from 2 groups and showed 2 specific clusters, one from Brazil and another from Costa Rica. Group 3 included only 2 sequences, one from the USA and another from Belgium. We opted to denominate this cluster as a group, despite such few sequences, because it was genetically distinct from other sequences and may have contained 2 sequences only because there were no other sequences from Europe in the database. Group 4 was formed by Brazilian, Japanese, and North American sequences. Group 5 was formed by Australian and Brazilian sequences.

Most genetic differences within a group exist because of clusters formed by sequences from the same countries. There was a very clear division between these clusters, except in group 1, where there was only one cluster specific for sequences from only one country. The subgroups within these major groups could be also divided into new groups, but few samples were available for LTR, and thus we did not divide the BLV into 9 phylogenetic groups. Group 5 is a clear example, as it was divided into 2 clusters (one formed by 2 Brazilian sequences and one by an Australian sequence) and showed the greatest genetic distance within groups (even greater than the genetic distance between the other groups). This result indicates the presence of 2 genetic groups, but a limited number of sequences were available, and thus we used only 5 genetic groups. The addition of other sequences support the formation of a new group, as observed in other studies examining the *env* gene (Rola-Łuszczak et al., 2013).

Camargos et al. (2007) found 4 groups when comparing the BLV *env* gene of Brazilian strains and from other countries; however, strains were positioned differently. Strains D0067 and K02120 were positioned in different groups, but were clustered closer. Notably, Brazilian strain 23M clustered in a separate group from others of the same location, but together with strains from Argentina, ARG41 and AF257515, which are characterized by low and high replication rates, respectively (Dube et al., 2000).

Moratorio et al. (2010) identified a higher degree of genetic variability, with 7 different BLV genotypes circulating in the South American region and 2 genotypes of BLV in Brazilian herds using the *env* gene for phylogenetic analysis. Rola-Łuszczak et al. (2013) also found higher variability in *env* sequences with inter-group distances varying from 2.5-4.1%. Despite the differences in genetic variation, LTR and *env* phylogenetic analysis showed that close related sequences clustered in specific subgroups within the 5 possibly 6 groups. The major groups were not specific for any region or country, possibly because of cattle commerce and the high prevalence of the disease in many countries (Rola-Łuszczak et al., 2013).

BLV appears to have evolved similarly to human and simian T-cell leukemia/lymphotropic viruses. Both viruses show slow evolution in LTR region and present 5-6 major genetic groups (Van Dooren et al., 2001). Among human T lymphotropic virus-1 genetic groups, a cosmopolitan group is found in most continents (Sonoda et al., 2011), similarly to what is observed for BLV Group 1. Human T lymphotropic virus-II also exhibits low phylogeographic concordance and low levels of transcontinental genetic differentiation (Slattery et al., 1999).

Our results indicate that BLV phylogeny may differ dramatically according to the viral DNA region used. Camargos et al. (2007) used a coding BLV portion with elevated immune system selective pressure. The LTR region does not encode proteins associated with the immune response and, thus, is not subject to the same direct pressure. However, this region is associated with viral genetic regulation. Both regions evolve differently, and it is likely that LTR is under neutral or even negative selective pressure, as reported for other retroviruses (Blackard et al., 2000). BLV LTR phylogeny is marked by a major group (Group 1), with viruses from multiple continents and minor groups (Groups 2-5) divided into clusters formed by sequences from specific countries.

In conclusion the 5'LTR PCR developed in this study was specific and was used to successfully sequence the BLV 5'LTR region amplification products. BLV Brazilian strains were highly conserved in the 5'LTR region, although 5 distinct groups were identified. These observations suggest that the genetic transcription regulation region can be used for phylogenetic analysis of BLV.

The occurrence of hypervariable and less variable sequences, or even highly conserved sequences, in the LTR of BLV Brazilian strains in transcription regulation segments reinforce that the virus may use genetic variability as a transcriptional control process, in parallel with other mechanisms.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), INCT Pecuária and Fundação de Apoio ao Ensino, Pesquisa e Extensão (FEPE). M.B. Heinemann, J.K.P. Reis and R.C. Leite had fellowships from CNPq.

REFERENCES

Avidan O, Meer ME, Oz I and Hizi A (2002). The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus. *Eur. J. Biochem*. 269: 859-867.

Bandelt HJ, Forster P and Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol*. 16: 37-48.

Blackard JT, Renjifo B, Chaplin B, Msamanga G, et al. (2000). Diversity of the HIV-1 long terminal repeat following mother-to-child transmission. *Virology* 274: 402-411.

Camargos MF, Stancek D, Rocha MA, Lessa LM, et al. (2002). Partial sequencing of the env gene of bovine leukaemia virus from braziliam samples and phylogenetic analysis. *J. Vet. Med. B Infect. Dis. Vet. Public Health* 49: 325-331.

Camargos MF, Pereda A, Stancek D, Rocha MA, et al. (2007). Molecular characterization of the env gene from Brazilian field isolates of Bovine leukemia virus. *Virus Genes* 34: 343-350.

Debacq C, Asquith B, Kerkhofs P, Portetelle D, et al. (2002). Increased cell proliferation, but not reduced cell death, induces lymphocytosis in bovine leukemia virus-infected sheep. *Proc. Natl. Acad. Sci. U. S. A.* 99: 10048-10053.

Dube S, Dolcini G, Abbott L, Mehta S, et al. (2000). The complete genomic sequence of a BLV strain from a Holstein cow from Argentina. *Virology* 277: 379-386.

Dube S, Abbott L, Dube DK, Dolcini G, et al. (2009). The complete genomic sequence of an *in vivo* low replicating BLV strain. *Virol. J*. 6: 120.

Jeanmougin F, Thompson J, Gouy M, Higgins DG, et al. (1998). Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* 23: 403-405.

Juliarena MA, Lendez PA, Gutierrez SE, Forletti A, et al. (2013). Partial molecular characterization of different proviral strains of bovine leukemia virus. *Arch. Virol*. 158: 63-70.

Kabeya H, Ohashi K and Onuma M (2001). Host immune responses in the course of bovine leukemia virus infection. *J. Vet. Med. Sci*. 63: 703-708.

Merezak C, Pierreux C, Adam E, Lemaigre F, et al. (2001). Suboptimal enhancer sequences are required for efficient bovine leukemia vírus propagation *in vivo*: implications for viral latency. *J. Virol.* 75: 6977-6988.

Moratorio G, Obal G, Dubra A, Correa A, et al. (2010). Phylogenetic analysis of bovine leukemia viruses isolated in South America reveals diversification in seven distinct genotypes. *Arch. Virol*. 155: 481-489.

Moratorio G, Fischer S, Bianchi S, Tomé L, et al. (2013). A detailed molecular analysis of complete bovine leukemia Virus genomes isolated from B-cell lymphosarcomas. *Vet. Res*. 44: 19.

Reed-Inderbitzin E and Maury W (2003). Cellular specificity of HIV-1 replication can be controlled by LTR sequences. *Virology* 314: 680-695.

Rychlik W and Rhoads RE (1989). A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro DNA amplification. *Nucleic Acids. Res*. 17: 8543-8551.

Rola-Łuszczak M, Pluta A, Olech M, Donnik I, et al. (2013). The molecular characterization of bovine leukaemia virus isolates from Eastern Europe and Siberia and its impact on phylogeny. *PLoS One* 8: e58705.

Sagata N, Yasunaga T, Tsuzuku-Kawamura J, Ohishi K, et al. (1985). Complete nucleotide sequence of the genome of bovine leukemia virus: Its evolutionary relationship to other retroviruses. *Proc. Natl Acad. Sci. U. S. A.* 82: 677-681.

Slattery JP, Franchini G and Gessain A (1999). Genomic evolution, patterns of global dissemination, and interspecies transmission of human and simian T-cell leukemia/lymphotropic viruses. *Genome Res*. 9: 525-540.

Sonoda S, Li HC and Tajima K (2011). Ethnoepidemiology of HTLV-1 related diseases: ethnic determinants of HTLV-1 susceptibility and its worldwide dispersal. *Cancer Sci*. 102: 295-301.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol*. 24: 1596-1599.

Van Dooren S, Salemi M and Vandamme AM (2001). Dating the origin of the African human T-cell lymphotropic virus type-i (HTLV-I) subtypes. *Mol. Biol. Evol*. 18: 661-671.

Van Dooren S, Verschoor EJ, Fagrouch Z and Vandamme AM (2007). Phylogeny of primate T lymphotropic virus type 1 (PTLV-1) including various new Asian and African non-human primate strains. *Infect. Genet. Evol*. 7: 374-381.

Xiao X and Buehring GC (1998). *In vivo* protein binding and functional analysis of cis-acting elements in the U3 region of bovine leukemia virus long terminal. *J. Virol.* 72: 5994-6003.

Zhao X, Jimenez C, Sentsui H and Buehring GC (2007). Sequence polymorphisms in the long terminal repeat of bovine leukemia virus: evidence for selection pressures in regulatory sequences. *Virus Res*. 124: 113-124.